Lambda Lysozyme: Overexpression, Ligand Studies, and Incorporation of Trifluoromethionine

by

Henry Steven Duewel

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Chemistry

Waterloo, Ontario, Canada, 1997

© Henry S. Duewel 1997

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced with the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-21342-0

Canada

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

Lambda Lysozyme: Overexpression, Ligand Studies, and Incorporation of Trifluoromethionine

Lysozymes are widely distributed in nature and function to cleave the glycan chains of the bacterial peptidoglycan. The lysozyme from bacteriophage lambda (lambda lysozyme; LaL) is the late gene product of the lambda *R* gene. The lambda *R* gene has been incorporated into several plasmids for the intended high level expression of LaL in *Eschericia coli*. An efficient procedure permitting purification of LaL from three of the four expression systems was established. The highest level of overproduction of LaL, producing 25-30 mg of purified LaL per litre of culture, was achieved by the expression of LaL from *E. coli* BL21(λDE3) harboring plasmid pLR102.

A turbidimetric assay utilizing chloroform-sensitized *E. coli* substrate cells has been optimized to assess the bacteriolytic activity of LaL. Several preparations of the substrate cells have indicated a "batch to batch" variation with respect to their susceptibility to LaL. Chitooligosaccharides ($(GlcNAc)_n$; oligomers of β-1,4-N-acetyl-D-glucosamine) were shown to inhibit the bacteriolytic activity of LaL but were not cleaved by the enzyme. Differential scanning calorimetry revealed that the temperature at which the thermal denaturation of LaL occurred increased in the presence of $(GlcNAc)_3$ and $(GlcNAc)_5$. Lambda lysozyme was also expressed in an *E. coli* strain auxotrophic for methionine, allowing for the incorporation of [*methyl*-^{13}C]methionine in the enzyme. An alteration of the [^{1}H-^{13}C]HMQC NMR spectra of the labelled enzyme was observed in the presence of $(GlcNAc)_5$. Chemical modification studies suggested the importance of acidic residue(s) for the bacteriolytic activity of LaL and chitooligosaccharides were shown to protect carboxyl group modification. The observed pattern of inhibition and carboxyl group protection by chitooligosaccharides and the effect of chitooligosaccharides on the intrinsic fluorescence of LaL suggested a possible difference in the binding modes for $(GlcNAc)_{n\leq 3}$ and for $(GlcNAc)_{n\geq 4}$ to LaL. The results indicate that LaL has specific interactions with oligosaccharides of N-acetylglucosamine but is incapable of cleaving the glycosidic bonds of these saccharides.

Two small peptides (*L*-Ala-*D*-*iso*-Glu-*L*-Lys-$(D\text{-Ala})_n$, n=1 or 2) mimicking the peptide portion of *E. coli* peptidoglycan were prepared and their interactions with LaL were evaluated. Affinity capillary electrophoresis suggested the possibility of a complex formation between peptide and LaL. Interestingly, the bacteriolytic activity of LaL was found to increase in the presence of the peptides. In addition, evidence is presented suggesting that only in the presence of exogenous peptide is LaL capable of cleaving a

chitooligosaccharide substrate. The synthesis of *p*-nitrophenyl-β-GlcNAc (β-PNP-GlcNAc) and some substituted derivatives of β-PNP-N-acetylmuramic acid is described. The muramic acid analogues did not serve as substrates for LaL or hen's egg white lysozyme.

Much interest is currently focused on understanding the detailed contribution that particular amino acid residues make in protein structure and function. Fluorinated amino acids have been utilized successfully to probe protein structure and dynamics and have pointed to the importance of specific residues to biological function. The importance of the amino acid methionine in proteins has been examined by the successful incorporation of *L*-S-trifluoromethylhomocysteine (*L*-trifluoromethionine; *L*-TFM) into lambda lysozyme, an enzyme containing three methionine residues. The *L* isomer of TFM was synthesized in an overall yield of 33% from N-acetyl-*D*, *L*-homocysteine thiolactone and trifluoromethyl iodide. The expression plasmid pLR102 was transformed into an *E. coli* strain auxotrophic for methionine permitting the expression of LaL in the presence of *L*-TFM. The analogue would not support growth of the auxotroph and was found to be inhibitory to cell growth. However, cells that were initially grown in a methionine rich media followed by protein induction under careful control of the respective concentrations of *L*-methionine and *L*-TFM in the media, were able to overexpress TFM-labelled LaL (TFM-LaL) at both high (70 and 74%) and low (31%) levels of TFM incorporation. TFM-LaL at both levels of incorporation exhibited activity analogous with that of the wild type enzyme and were inhibited by chitooligosaccharides indicating that replacement of the methionine residues by TFM did not hinder enzyme function. Interestingly, the ^{19}F solution NMR spectra of the TFM labelled enzymes consisted of four sharp resonances spanning a chemical shift range of 0.9 ppm, with three of the resonances showing very modest shielding changes on binding of $(GlcNAc)_5$. The ^{19}F NMR analysis of TFM-LaL at both high and low levels of incorporation suggested that one of the methionine positions gives rise to two separate resonances. The intensities of these two resonances were influenced by the extent of incorporation which was interpreted as an indication that subtle conformational changes in protein structure accompany the incorporation of TFM. The similarities and differences between Met and TFM were analyzed using semi-empirical and *ab initio* molecular orbital calculations. The TFM-labelled enzymes were found to be unreactive towards chemical cleavage with cyanogen bromide. Mass spectral analysis has strongly suggested the specific formylation of the hydroxyl groups of serine, threonine and C-terminal homoserine residues during the course of the cyanogen bromide reaction. The methodology presented offers promise as a new approach to the study of protein-ligand interactions as well as for future investigations into the functional importance of methionine in proteins.

Acknowledgements

I would like to extend my heartfelt gratitude to Dr. John Honek for being an equally exceptional supervisor, educator and valued friend. My gratefulness for his endless and devoted support, spirited enthusiasm and selfless guidance throughout my studies is exceeded only by my appreciation for his kind and continual regard for my personal welfare and prosperity. Thank you, John, for everything.

I also wish express my sincere appreciation to Dr. Elisabeth Daub for her co-supervision of this project and for her dedication of valuable time towards the instrumental cloning studies presented in this thesis.

I would also like to thank Martha Cox, Walter Giust, Denise Miedema and Robert Ruman for their contributions to the work presented here. I am grateful to Dr. Sandra Mooibroek for her assistance performing HMQC NMR experiments and to Valerie Robinson of the University of Guelph for the acquisition of ^{19}F NMR spectra. I am indebted to Lorne Taylor for his prominent technical assistance and instruction with mass spectrometry. I would also like to acknowledge the excellent services provided by all the staff members in the Department of Chemistry.

I am honoured to have studied along side my former and present colleagues. I thank John Barnard, Dr. Anna Crivici, Dr. Mike Houston, Robert Kinach, Laura Marrone, Dr. Hamzah Mohd. Salleh, Peter Sampson, Joey Smith and Dr. Gerry Wright for their stimulating discussions (at times, approaching topics related to science), essential assistance and friendship. In particular, I would like to thank Dr. Ken Sokoll for his boundless support both in and out of the lab.

I am extremely grateful to Dr. Thammaiah Viswanatha for his constant encouragement, valued discussions and use of laboratory equipment throughout the last seven years. Many thanks to Dr. Gilles Lajoie for his advice on peptide synthesis and unrestricted access to the MALDI and electrospray mass spectrometers. I appreciate the generous assistance and helpful discussions of Drs. Harold Frey and James Lepock with differential scanning calorimetry experiments.

I gratefully acknowledge the Natural Sciences and Engineering Research Council of Canada for a graduate scholarship and for financial support from the University of Waterloo.

Finally, this work would not have been possible without the understanding, patience and encouragement of my family and friends.

For my parents, Jüergen and Elisabeth

Table of Contents

Abstract	iv
Acknowledgements	vi
Dedication	vii
Table of Contents	viii
List of Tables	xiv
List of Figures	xvi
List of Abbreviations	xxi

CHAPTER 1

Introduction

1.1. Classification of Lysozymes	1
1.2. The Lysozyme Substrate: Peptidoglycan and the Cell Wall	6
1.3. Bacteriophage Lambda Lysozyme	
1.3.1. Phage Induced Lysis and the λ *R* Gene	12
1.3.2. What is Known about Lambda Lysozyme	13
1.4. *E. coli* Lytic Transglycosylases	16
1.5. The Phillips Model for Hexasaccharide Binding and the Lysozyme Mechanism	19
1.5.1. Evidence Supporting the Phillips Mechanism	22
1.5.2. Alternatives to the Phillips Mechanism	27
1.5.3. Mechanism for Lambda Lysozyme and Slt70	35
1.6. Structural Homologies between Lysozymes	36
1.6.1. Homologies Existing with Lambda Lysozyme	43
1.7. Theoretical Predictions of Hexasaccharide Binding	48
1.8. Analysis of the Contributions to Binding of Saccharides to the Subsites of HEWL	53
1.9. Lysozyme/Chitooligosaccharide Interactions	
1.9.1. Cleavage Patterns and Transglycosylation Reactions	59
1.9.2. Productive and Nonproductive Complexes	62
1.9.3. Conformational Changes	68
1.10. Considerations of Distortion	71
1.11. Biological Properties of Peptidoglycan and Muramyl Peptides	82
1.12. General Objectives	86

CHAPTER 2

Overexpression, Purification and Characterization of Lambda Lysozyme

2.1. INTRODUCTION 87

2.2. EXPERIMENTAL

2.2.1. General Experimental 88

2.2.2. Materials 88

2.2.3. Plasmids and Bacterial Strains 89

2.2.4. General Recombinant DNA Techniques 89

2.2.5. Cloning of the λ *R* Gene and Construction of Expression Vectors

2.2.5.1. *E. coli* DH5α/pLR1 92
2.2.5.2. *E. coli* TG-1/pLR7 92
2.2.5.2.1 Sequence determination of the pLR7 insert 93
2.2.5.3. *E. coli* OR1265/pLcIR18 94
2.2.5.4. *E. coli* TG-1/pHDM10 95
2.2.5.5. *E. coli* BL21(λDE3)/pLR102 95

2.2.6. Activity Determinations 97

2.2.7. SDS-PAGE Determination of LaL Expression Levels 98

2.2.8. General SDS-PAGE 98

2.2.9. Protein Purification

2.2.9.1. General Purification Procedures 99
2.2.9.2. Purification from *E. coli* TG-1/pLR7 101
2.2.9.3. Purification from *E. coli* OR1265/pLcIR18 103
2.2.9.4. Purification from *E. coli* TG-1/pHDM10 104
2.2.9.5. Purification from *E. coli* BL21(λDE3)/pLR102 106

2.2.10. Protein Quantitation Methods

2.2.10.1. Bradford Analysis 107
2.2.10.2. Quantitation by Mass Following Dialysis and Lyophilization 107
2.2.10.3. Determination of the Molar Extinction Coefficient 109

2.2.11. Electrospray Mass Spectrometry 112

2.2.12. Estimation of the Isoelectric Point (pI) 114

2.2.13. Crystallization Attempts 114

2.3. RESULTS AND DISCUSSION

2.3.1. Cloning of the λ *R* Gene and Purification of LaL from Strains Harboring Recombinant LaL Expression Vectors

2.3.1.1. Construction of pLR7 and Purification from *E. coli* TG-1/pLR7 115
2.3.1.2. Construction of pLcIR18 and Purification from *E. coli* OR1265/pLcIR18 125
2.3.1.3. Construction of pHDM10 and Purification from *E. coli* TG-1/pHDM10 131

2.3.1.4. Construction of pLR102 and Purification from *E. coli* BL21(λDE3)/pLR102 143

2.3.2. Characterization of LaL Purified from *E. coli* TG-1/pHDM10 and *E. coli* BL21(λDE3)/pLR102

2.3.2.1. Nucleotide Level 154
2.3.2.2. Protein Level: N-Terminal Analysis, ESMS and pI 155

2.3.3. Guanidine Induced Unfolding and Molar Extinction Coefficient of LaL 161

2.3.4. Comparison of Protein Quantitation Methods 164

2.3.5. Distinctive Properties of LaL Relevant to its Manipulation

2.3.5.1. Maintaining Lysozyme Preparations 166
2.3.5.2. Difficulties Encountered 167

2.3.6. Crystallization Efforts 175

2.3.7. Summary and Future Work 177

CHAPTER 3

Turbidimetric Assay, Interactions of Saccharides and Peptides with LaL and Synthesis and Evaluation of Alternative Substrates

3.1. INTRODUCTION 179

3.1.1. Methods to Demonstrate Lysozyme Activity

3.1.1.1. Bacteriolytic Assays 179
3.1.1.2. Insoluble Peptidoglycan Substrates 181
3.1.1.3. Soluble Chemically-Defined Substrates 182
3.1.1.3.1. Class I: Substrates Derived from Peptidoglycan 183
3.1.1.3.2. Class II: Substrates Derived from Chitin 184

3.1.2. Substrate Induced Evolution 190

3.1.3. Comparative Inhibitory Behaviour 193

3.1.4. Supporting Evidence for Additional Lysozyme/Substrate Interactions 194

3.1.5. Observations on Activity with use of Different Substrates 198

3.1.6. Concluding Statements 202

3.2. EXPERIMENTAL

3.2.1. Materials 203

3.2.2. Analytical Techniques 204

3.2.3. Assays Involving Substrate Cells (Turbidimetric Assay)

3.2.3.1. Preparation of Bacterial Substrate Cells 207

3.2.3.2. Characterization of the Response of Substrate Cell Preparations to LaL Concentration 208
3.2.3.3. Inhibition of Bacteriolytic Activity Assays 210
3.2.3.4. Effect of Buffer and Ionic Strength on the Turbidimetric Assay 210

3.2.4. Biophysical Studies on the Interactions of $(GlcNAc)_n$ with LaL

3.2.4.1. Fluorescence 211
3.2.4.2. Differential Scanning Calorimetry 212
3.2.4.3. NMR Studies of [*methyl*-^{13}C]Methionine Labelled LaL
3.2.4.3.1. Synthesis of [*methyl*-^{13}C]methionine 213
3.2.4.3.2. Requirement of Methionine Supplementation for the Growth of *E. coli* DH93/pHDM10 214
3.2.4.3.3. Preparation of [*methyl*-^{13}C]Methionine Labelled LaL 214
3.2.4.3.4. NMR Spectroscopy 216

3.2.5. Peptide Synthesis and Purification

3.2.5.1. Solid Phase Peptide Synthesis 217
3.2.5.2. Resin Cleavage and Purification 221

3.2.6. Synthesis of Substrates

3.2.6.1. General 226
3.2.6.2. β-PNP-GlcNAc 227
3.2.6.3. β-PNP-MurNAc and Derivatives 230
3.2.6.4. Substituted β-PNP-MurNAc Derivatives 235

3.2.7. Enzymic Assays with Commercial and Synthetic Substrates

3.2.7.1. Chitooligosaccharides $(GlcNAc)_n$ and Nitrophenyl Glycosides 243
3.2.7.2. β-PNP-MurNAc Analogues 245

3.2.8. Chemical Modification 247

3.3. RESULTS AND DISCUSSION 248

3.3.1. Substrate Cells and the Turbidimetric Assay

3.3.1.1. Chloroform Treatment and Evaluation of Substrate Cells 248
3.3.1.2. Ionic Strength Effects on the Turbidimetric Assay 255

3.3.2. Interactions of Chitooligosaccharides with LaL

3.3.2.1. Kinetic Observations 261
3.3.2.2. Fluorescence Studies 269
3.3.2.3. Differential Scanning Calorimetry 275
3.3.2.4. NMR Studies on [*methyl*-^{13}C]Methionine Labelled LaL 280
3.3.2.5. Chitooligosaccharide Protection of Carboxyl Modification 286

3.3.3. Interactions of Peptides with LaL 294

3.3.3.1. Peptide Synthesis and Characterization 294
3.3.3.2. Evaluation of the Interactions of Peptides with LaL 309

3.3.4. Synthesis of β-*p*-Nitrophenyl-Muramic Acid Analogues 319

3.3.5. Evaluation of the β-PNP-MurNAc Analogues as Substrates 337

3.3.6. Summary and Future Work 341

CHAPTER 4

Incorporation of Trifluoromethionine into Lambda Lysozyme at High and Low Levels: Characterization, ^{19}F NMR and Cyanogen Bromide Reaction of the Labelled Proteins

4.1. INRODUCTION

4.1.1. Methionine in Proteins 344

4.1.2. ^{19}F NMR of Proteins 348

4.1.3. Scope and Objectives 351

4.2. EXPERIMENTAL

4.2.1. General Experimental and Materials 352

4.2.2. Synthesis of *L*-S-Trifluoromethylhomocysteine 352

4.2.3. Expression and Purification of TFM-LaL 354

4.2.4. Optimization of Conditions Yielding Maximal TFM-LaL Expression and Desired Incorporation Levels 355

4.2.5. Effect of TFM on Cell Growth 356

4.2.6. ^{19}F NMR Spectroscopy 356

4.2.7. Methionine Aminopeptidase Activity on LaL 357

4.2.8. Molecular Modelling and NMR Calculations 358

4.2.9. Cyanogen Bromide Reactions

4.2.9.1. Cyanogen Bromide Cleavage of LaL and TFM-LaL Proteins 359
4.2.9.2. Synthesis of N-Ac-Met and N-Ac-TFM 361
4.2.9.3. Treatment of N-Ac-Met and N-Ac-TFM with CNBr 362
4.2.9.4. Treatment of Serine, Threonine and Tyrosine with CNBr 363

4.3. RESULTS AND DISCUSSION

4.3.1. Synthesis of *L*-TFM 364

4.3.2. Optimal Conditions to Produce TFM-Labelled LaL 364

4.3.2.1. High Level TFM-LaL 366
4.3.2.2. Low level TFM-LaL 369

4.3.3. TFM Toxicity and Metabolism 369

4.3.4. Effect of *L*-TFM Incorporation on Activity and Propensity to Inhibition by Chitooligosaccharides 373

4.3.5. ^{19}F NMR Studies of TFM-LaL

4.3.5.1. ^{19}F NMR Spectra of High and Low level TFM-LaL and the Effect of $(GlcNAc)_5$ Binding on the Spectra 375
4.3.5.2. Possible Nature of the C/D "Double Resonance" 382

4.3.6. Molecular Modelling Calculations 387

4.3.7. Cyanogen Bromide Treatment of LaL and TFM-LaL Proteins 392
4.3.7.1. Effect of CNBr reaction on N-Ac-Met and N-Ac-TFM 410
4.3.7.2. Formylation of CNBr Reaction Products 417
4.3.8. Summary and Future Work 425

Appendix A: Media and Buffer Preparation and Description 428
Appendix B: Crystal Screen™ Reagent Components 433
Appendix C: Sample Calculations 435
Appendix D: Determination of the Level of TFM Incorporation 438
REFERENCES 441

List of Tables

Chapter 1 **Page**

1.1. Classification of lysozymes of bacteriophage origin 4

1.2. Comparison of Gram-negative and Gram-positive cell walls 7

1.3. Catalytic acid residues suggested for several lysozymes 22

1.4. Comparison of the association of various saccharides to HEWL 54

1.5. Estimated binding contributions of the subsites in HEWL 57

1.6. Binding and cleavage data of chitooligosaccharides with HEWL 59

1.7. Values of k_{cat}/K_m ($M^{-1}s^{-1}$) for the HEWL catalyzed hydrolysis of XylNAc and GlcNAc nitrophenyl-glycosides 79

Chapter 2

2.2. Silver staining of protein gels 100

2.2. Representative purification of LaL from a 6 L culture growth scale of *E. coli* TG-1/pHDM10 142

2.3. Purification of LaL from a 6 L culture growth scale of *E. coli* BL21(λDE3)/pLR102 154

2.4. Masses of the molecular ion series for LaL 159

2.5. Calculation of the molar extinction coefficient for LaL 163

2.6. Comparison of methods to quantitate LaL 166

Chapter 3

3.1. Values of k_{cat}/K_m for the hydrolysis of nitrophenyl-glycosides by HEWL 187

3.2. Substrate specificity requirements of lysozymes 191

3.3. Bacterial substrate specificity of Gram-negative and Gram-positive phage lysozymes 192

3.4. Inhibitory properties of lysozymes 193

3.5. Inhibition of T4L by various muropeptides 195

3.6. Comparison of the response of different substrate cell preparations 255

3.7. Inhibition of the bacteriolytic activity of LaL by selected saccharides 267

3.8. Effect of saccharides on the fluorescence of LaL and HEWL 270

3.9. Effect of LaL and HEWL on the fluorescence of $(GlcNAc)_3$-4MU 274

3.10. Transition temperatures obtained for LaL in the presence and absence of selected saccharides 277

3.11. Effect of $(GlcNAc)_5$ and maltotetraose on the chemical shifts of the ^1H- and ^{13}C-nuclei of the methionine methyl groups in ^{13}C-Met-LaL 284

3.12. Summary of peptide syntheses 295

3.13. ^1H and ^{13}C resonance assignments for peptide **2** 303

3.14. Effect of peptide **2** and **3** on the bacteriolytic activity of LaL 309

3.15. Evaluation of compounds **9**, **10** and **12** with LaL and HEWL 339

Chapter 4

4.1. Gradient conditions for separation of CNBr generated fragments 360

4.2. Gradient conditions for analysis of reaction products of CNBr with Ac-Met and Ac-TFM 362

4.3. Effect of various concentrations of *L*-met or *L*-TFM alone or in concert on the growth of *E. coli* B834 (DE3) harboring pLR102 370

4.4. ^{19}F NMR data for high and low level TFM-LaL 377

4.5. Effect of $(GlcNAc)_5$ binding on the ^{19}F NMR resonances of TFM-LaL 378

4.6. Structural and Electronic Characteristics of *n*-Butane and Analogues 389

4.7. Summary of HPLC results of CNBr generated fragments of *wt* LaL 399

4.8. Summary of HPLC results of CNBr generated fragments of High level TFM-LaL 402

4.9. Summary of HPLC results of CNBr generated fragments of Low level TFM-LaL 406

4.10. Integration data for the reactions of *N*-Ac-Met and *N*-Ac-TFM with CNBr 415

4.11. Amino acid composition of CNBr generated fragments of LaL 420

4.12. ESMS data for Ser, Thr and Tyr treated under various conditions 420

List of Figures

		Page
Chapter 1		
1.1.	Structure of peptidogylcan	8
1.2.	Space-filling model of the peptidoglycan	10
1.3.	Amino acid sequence and composition of LaL	14
1.4.	1,6-Anhydro-muramic acid production by LaL	15
1.5.	Places of action of *E. coli* peptidoglycan hydrolases	17
1.6.	Structure of the *E. coli* soluble lytic transglycosylase Slt70	18
1.7.	Proposed Phillips model for hexasaccharide binding to HEWL	19
1.8.	The catalytic mechanism for HEWL as proposed by Phillips	21
1.9.	Possible transition state for the HEWL catalysed reaction	26
1.10.	Proposed double displacement mechanism for HEWL	28
1.11.	Participation of the 2-acetamido carbonyl oxygen in oxocarbonium ion stabilization	30
1.12.	Single displacement mechanism proposed for T4L	32
1.13.	Proposed endocyclic cleavage mechanism for HEWL	34
1.14.	Possible mechanism for lambda lysozyme and Slt70	35
1.15.	X-ray structures of lysozymes	38
1.16.	Schematic of the structural relationship between HEWL, GEWL and T4L	41
1.17.	Superposition of the structures of HEWL, GEWL, and T4L	42
1.18.	Schematic of the inferred saccharide-protein interactions for HEWL, T4L and GEWL	44
1.19.	Binding of MurNAc-GlcNAc-MurNAc to the B, C, and D sites of HEWL	45
1.20.	Structure of C-Slt70 and conserved features with lysozymes	47
1.21.	Predicted structures describing left-and right-sided binding of hexasaccharides to HEWL	49
1.22.	Schematic model representing the enzyme forms important for the lysozyme catalyzed cleavage of chitohexaose	63
1.23.	Regions of HEWL that experience conformational changes on substrate binding	69
1.24.	Comparison of elements in the active site of T4L and HEWL	71
1.25.	Possible conformations of a pyranose ring	73
1.26.	Distorted site D MurNAc residues observed in crystal structures of HEWL and T4L	74
1.27.	Change in standard free energy for an enzymatically catalyzed reaction	78

1.28. Possible reaction coordinates describing the HEWL catalysed reaction with XylNAc or GlcNAc containing substrates 80

1.29. Structures of some biologically active muramyl peptides 84

Chapter 2

2.1. Construction of plasmid pLR1 116

2.2. Construction of plasmid pLR7 117

2.3. Schematic of the cloning of the pLR7 λ DNA fragment into M13 vectors for sequencing 118

2.4. Schematic of the *tac* promoter and polylinker regions of pTTQ18 119

2.5. Mono-S chromatogram from purification of LaL from *E. coli* TG-1/pLR7 121

2.6. Phenyl-Superose chromatogram from purification of LaL from *E. coli* TG-1/pLR7 122

2.7. SDS-PAGE analysis of the stages of purification of LaL from *E. coli* TG-1 harboring pLR7 123

2.8. Sequence of the *R* gene ribosome binding site 125

2.9. Construction of plasmid pLcIR18 126

2.10. Mono-S chromatogram from purification of LaL from *E. coli* OR1265/pLcIR18 128

2.11. Phenyl-Superose chromatogram from purification of LaL from *E. coli* OR1265/pLcIR18 129

2.12. SDS-PAGE analysis of the stages of purification of LaL from *E. coli* OR1265 harboring pLcIR18 130

2.13. Construction of plasmid pHDM10 132

2.14. Schematic of the insertion mode of the λ DNA fragment of pHDM10 into the *Bam*HI site of pTTQ18 133

2.15. S-Sepharose Fast Flow chromatogram from purification of LaL from *E. coli* TG-1/pHDM10 137

2.16. Mono-S chromatogram from purification of LaL from *E. coli* TG-1/pHDM10 138

2.17. Phenyl-Superose chromatogram from purification of LaL from *E. coli* TG-1/pHDM10 139

2.18. Phenyl-Superose chromatogram for the concentration of LaL from *E. coli* TG-1/pHDM10 140

2.19. SDS-PAGE analysis of the stages of purification of LaL from *E. coli* TG-1 harboring pHDM10 141

2.20. Schematic of the polymerase chain reaction for the amplification of the lambda *R* gene from pHDM10 144

2.21. Construction of plasmid pLR102 146

2.22. Schematic of the transcription and expression region of pET-22b 147

2.23. Expression of LaL directed by *E. coli* BL21(λDE3)/pLR102 and pLR103 147
2.24. S-Sepharose Fast Flow chromatogram from purification of LaL from *E. coli* BL21(λDE3)/pLR102 149
2.25. Mono-S chromatograms from purification of LaL from *E. coli* BL21(λDE3)/pLR102 150
2.26. Phenyl-Superose chromatograms from the purification of LaL from *E. coli* BL21(λDE3)/pLR102 151
2.27. SDS-PAGE analysis of the stages of purification of LaL from *E. coli* BL21(λDE3) harboring pLR102 152
2.28. MaxEnt reconstructed ESMS spectra of LaL purified from *E. coli* BL21/pLR102 and TG-1/pHDM10 157
2.29. Multiply charged raw data ESMS spectra of LaL purified from *E. coli* BL21/pLR102 and TG-1/pHDM10 158
2.30. ESMS spectrum of non-desalted LaL illustrating the potassium and water adducts 160
2.31. Guanidine induced unfolding curve of LaL 162
2.32. NMR spectra demonstrating contamination resulting from dialysis tubing 170
2.33. NMR spectra obtained for LaL and HEWL following dialysis 172

Chapter 3

3.1. Class I soluble substrates for lysozymes 183
3.2. Class II soluble substrates for lysozymes 185
3.3. Reaction of $(GlcNAc)_n$-PNP with HEWL and human urinary lysozyme 189
3.4. Structure of the mutant T4L/peptidoglycan complex 196
3.5. Models describing substrate-assisted catalysis 200
3.6. Chloroform treatment of *E. coli* MG1655 249
3.7. Response of substrate cells to LaL concentration (1) 253
3.8. Response of substrate cells to LaL concentration (2) 254
3.9. Effect of potassium phosphate and Tris buffer concentration on the turbidimetric assay 257
3.10. Effect of KCl and $(NH_4)_2SO_4$ concentration on the turbidimetric assay 258
3.11. Spontaneous lysis of *E. coli* substrate cells 260
3.12. Action of LaL and HEWL on $(GlcNAc)_5$ and $(GlcNAc)_6$ 262
3.13. Inhibition of the bacteriolytic activity of LaL by $(GlcNAc)_5$ 265
3.14. Inhibition of the bacteriolytic activity of LaL by $(GlcNAc)_n$ 266
3.15. Alignment of Trp containing regions of LaL and HEWL 272
3.16. Typical DSC curve for the thermal denaturation of LaL 276

3.17. Representative DSC scans observed for LaL in the presence of selected saccharides 278
3.18. DSC scans for HEWL in the presence and absence of $(GlcNAc)_3$ 278
3.19. DSC scans demonstrating the partial reversibility of the thermal denaturation of LaL 279
3.20. Dependence on methionine supplementation for the growth of *E. coli* DH93 harboring pHDM10 282
3.21. [^1H-^{13}C]HMQC spectrum of [*methyl*-^{13}C]methionine labelled LaL 283
3.22. ESMS spectra of ^{13}C-Met-LaL purified from *E. coli* B834/pLR102 285
3.23. Inactivation of LaL by EDC and EAC 287
3.24. Comparison of the inactivation of LaL by EAC and EDC 288
3.25. Order of the inactivation of LaL by EAC and EDC 290
3.26. Protection of LaL from EDC inactivation by $(GlcNAc)_n$ 291
3.27. Protection of LaL from EAC inactivation by $(GlcNAc)_5$ 291
3.28. Possible side reactions reducing yields during synthesis of peptides 296
3.29. Capillary electrophoresis chromatograms of peptides **1**, **2**, and **3** 298
3.30. Mass spectra of peptides **1**, **2**, and **3** 300
3.31. ^1H NMR spectrum of peptide **1** 302
3.32. ^1H NMR spectrum of peptide **2** 304
3.33. ^1H (A) and ^{13}C (B) NMR spectra of peptide **3** 307
3.34. Affinity capillary electrophoresis of LaL with peptide **3** 311
3.35. Analysis of the co-incubation of LaL with $(GlcNAc)_5$-PNP and peptide **3** 314
3.36. ^1H NMR of the β-PNP-MurNAc analogues **13**, **14** and **15** 331
3.37. MALDI MS spectrum of compound **15** 335
3.38. Capillary electrophoresis chromatograms of compound **15** 336
3.39. UV spectra of selected compounds 338

Chapter 4

4.1. Conformations of *n*-butane and ethyl methyl sulfide and the χ_3 torsion angle of methionine 345
4.2. Methionines found in the carbohydrate binding sites of some proteins 347
4.3. Dependence on methionine supplementation for the growth of *E. coli* B834 harboring pLR102 365
4.4. Comparison of the expression efficiency of LaL and TFM-LaL under various conditions 367
4.5. Reconstructed electrospray mass spectra of high level TFM-LaL, low level TFM-LaL and wild type LaL 368

4.6.	Comparison of the activity and inhibition of *wt* and TFM-LaL	374
4.7.	Comparison of the ^{19}F NMR spectra of *L*-TFM and high level TFM-LaL	375
4.8.	^{19}F NMR spectra of high and low level incorporated TFM-labelled LaL	376
4.9.	Effect of $(GlcNAc)_5$ binding on the ^{19}F NMR resonances of TFM-labelled LaL	379
4.10.	Two-dimensional EXSY spectra of high level TFM-LaL	383
4.11.	Methionine amino peptidase activity on *wt* LaL	385
4.12.	Reverse phase chromatography of high level TFM-LaL	391
4.13.	Mechanism of CNBr cleavage of methionyl peptides	392
4.14.	Possible products from the CNBr cleavage of *wt* LaL and TFM-LaL	395
4.15.	RP-HPLC chromatograms monitoring the reaction of *wt* LaL and high level TFM-LaL with CNBr	396
4.16.	RP-HPLC chromatograms monitoring the reaction of low level TFM-LaL with CNBr	404
4.17.	Mechanisms of the reaction pathways during treatment of methionyl peptides with CNBr	410
4.18.	Diagram of the reaction pathways during treatment of *N*-acetylmethionine with CNBr	411
4.19.	RP-HPLC chromatograms monitoring the reaction of *N*-Ac-Met and *N*-Ac-TFM with CNBr	412
4.20.	ESMS spectrum of CNBr generated fragment B demonstrating the extent of formylation	417
4.21.	ESMS spectra for fragment B obtained from *wt* LaL illustrating the higher degree of formylation with time	418
4.22.	ESMS spectra of CNBr generated fragments E' and C	422
4.23.	ESMS spectrum of LaL treated in 70% formic acid for 24 h illustrating its formylation	423

List of Abbreviations

ACE	affinity capillary electrophoresis
*anh*MurNAc	1,6-anhydro-N-acetyl-muramic acid
BOC	*t*-butoxycarbonyl
BOP	benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorate
bp	base pair
^{13}C-Met-LaL	[*methyl*-^{13}C]methionine labelled LaL
ChL	*Chalaropsis* lysozyme
CZE	capillary zone electrophoresis
Da	dalton
ms-DAP	*meso*-2,6-diaminopimelic acid
DCC	dicyclohexylcarbodiimide
DCM	dichloromethane, methylene chloride
DDW	distilled deionized water
DIPC	N,N'-diisopropylcarbodiimide
DIPEA	diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	N,N-dimethylformamide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
ddNTP	dideoxynucleotide triphosphate
DTT	dithiothreitol
EAC	1-(4-azonia-4,4-dimethylpentyl)-3-ethyl-carbodiimide iodide
EDC	1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide
EDTA	N,N,N',N'-ethylenediaminetetraacetic acid
ESMS	electrospray (ionization) mass spectrometry
equiv.	equivalent
Fmoc	9-fluorenylmethoxycarbonyl
FPLC	fast protein liquid chromatography
Gdn·HCl	guanidine hydrochloride
GEWL	goose egg white lysozyme
GlcNAc	*N*-acetyl-D-glucosamine
$(GlcNAc)_n$	β(1→4)-linked *n*-mer of GlcNAc
$(GlcNAc)_3$-4MU	4-methylumbelliferyl-N-acetyl-chitotriose
HEWL	hen egg white lysozyme
HOBt	1-hydroxybenzotriazole
HPLC	high performance (pressure) liquid chromatography
HuL	human lysozyme
HMQC	heteronuclear multiple-quantum correlation
IPA	isopropanol
IPTG	isopropyl-β-*D*-thiogalactopyranoside
K_A	association constant
K_d	dissociation constant
kb	kilobase
kDa	kilodalton
KPB	potassium phosphate buffer
LaL	lambda lysozyme
LB	Luria-Bertani medium (see Appendix A)
LB_{amp}	LB medium supplemented with ampicillin
LT	a medium (see Appendix A)
M9	M9 minimal salts medium (see Appendix A)

$M9_{Glu,amp}$	M9 minimal medium supplemented with glucose (Glu) and ampicillin (Amp)
mequiv.	*milliequivalent*
MALDI MS	matrix assisted laser desorption ionization mass spectrometry
Met	methionine
MGM	β(1→4)-linked MurNAc-GlcNAc-MurNAc
MQW	Milli-Q™ Water
mRNA	messenger RNA
4MU	4-methylumbelliferone
MurNAc	*N*-acetyl-muramic acid
MWCO	molecular weight cut off
NMM	N-methylmorpholine
NMR	nuclear magnetic resonance
OD_λ	optical density measured and wavelength λ (nm)
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
PVDF	polyvinylidene difluoride
psi	pounds per square inch
RBS	ribosome binding site
RF	replicative form
RNA	ribonucleic acid
RNase	ribonuclease
RP-C18	reverse phase-carbon 18
SC	substrate cells
SD	Shine-Dalgrano sequence
SDS	sodium dodecylsulfate
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
Slt	soluble lytic transglycosylase
SOC	a medium (see Appendix A)
SRP	signal recognition particle
STET	a buffer (see Appendix A)
TACA	the free acid of TACL
TACL	tetra-N-acetylchitotetraose δ-lactone
TBE	Tris-borate-EDTA buffer (see Appendix A)
TE	Tris-EDTA buffer (see Appendix A)
TEMED	N,N,N',N'-tetra-methylethylenediamine
TEWL	turkey egg white lysozyme
TFA	trifluoroacetic acid
TFM	S-trifluoromethylhomocysteine, trifluoromethionine
TFM-LaL	TFM labelled LaL
TFMSA	trifluoromethanesulfonic acid
TLC	thin layer chromatography
TSP	2,2,3,3-tetradeutero-3-trimethylsilyl-propionic acid sodium salt
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
UV	ultraviolet
V	volt
vol.	volume
wt	wild type
X-gal	5-bromo-4-chloro-3-indolyl-β-*D*-galactopyranoside
XylNAc	2-acetamido-2-deoxy-D-xylose

CHAPTER 1

Introduction

Lysozyme has had an exceptional history paralleled by few enzymes. At one time when Sir Alexander Fleming was stricken with a cold, he inadvertently discovered a substance in his own nasal mucus capable of lysing certain bacteria. Fleming subsequently found that similar agents were present in many biological tissues and secretions. His evidence indicated that the lytic agents were enzymes and he applied the term *lysozyme* for *lytic enzyme* (Fleming, 1922). The primary structure of the lysozyme from hen's egg white (hen's egg white lysozyme; chicken lysozyme; HEWL) was elucidated independently by two groups in 1963 and became the first protein sequenced containing together the twenty naturally occurring amino acids (Canfield, 1963; Jollès et al., 1963). In 1965, David Phillips and his colleagues reported the three dimensional structure of HEWL which was the first enzyme whose structure was visualized from X-ray diffraction analysis (Blake et al., 1965). From inspection of the structure of HEWL with substrate analogues, HEWL later became the first enzyme for which a detailed mechanism of action was proposed (Phillips, 1966; Blake et al., 1967; Vernon, 1967), a mechanism widely depicted in current biochemistry texts and exemplifying the principles of enzymatic catalysis. The publication of the structure and mechanism of HEWL focused the attention of many scientists onto this enzyme and, in particular, HEWL has served as a model protein for many biochemical and structural studies. The modern era of lysozyme research has lost little momentum.

1.1. Classification of Lysozymes

Lysozymes are typically small molecular weight, basic proteins and are defined as mucopeptide β-1,4-N-acetylmuramylhydrolases (EC 3.2.1.17) and function to cleave the glycosidic linkage between N-acetylmuramic acid and N-acetylglucosamine of the bacterial peptidoglycan (the peptidoglycan is discussed in section 1.2). Some lysozymes have also demonstrated chitinase activity (1,4-(2-acetamido-2-deoxy)-β-D-glucoside glycanohydrolase; EC 3.2.1.14). Lysozymes are ubiquitous enzymes which are widespread in nature and have been isolated from a variety of organisms (reviewed by Jollès & Jollès (1984) and Osserman et al. (1974)). The role of lysozymes in plants, vertebrates and invertebrates has been most closely associated with one of providing defense against

bacterial or fungal pathogens (Jollès & Jollès, 1984). Additional physiological roles for lysozymes of higher vertebrates might exist and are discussed in section 1.11.

Structural, enzymatic and physical characterization of lysozymes has demonstrated the existence of four main lysozyme families; the *c*, *g*, *p* and *ch* types. Other lysozyme or lysozyme-like enzymes have been identified (Jollès & Jollès,1984) but the discussion that follows will center on the aforementioned groups.

i) c type Lysozymes

The *c* type (or chicken type) lysozymes are the largest and the best characterized of the lysozyme families. Members of this family include the lysozymes isolated from avian sources (hen, quail, pheasant, guinea fowl, turkey, duck, chachalaca), from mammals (human, baboon, rat, bovine, equine) and from reptiles (tortoise) and some insect species (Jollès & Jollès, 1984). The structures of several of the *c* type lysozymes have been analyzed by X-ray crystallography including HEWL (Blake et al., 1965, 1967; Ford et al., 1974; Kelly et al., 1979; Strynadka & James, 1991), human lysozyme (Artymiuk & Blake, 1981; Song et al., 1994; Muraki et al., 1996), rainbow trout lysozyme (Karlsen et al., 1995) and the lysozymes obtained from the egg whites of tortoise (Aschattenburg et al., 1980), turkey (Sarma & Bott, 1977; Howell et al., 1992), pigeon (Yao et al., 1992), partridge (Turner & Howell, 1995), quail (Chitarra et al., 1993) and guinea and pheasant (Lescar et al., 1994).

The *c* type lysozymes are typified by HEWL (MW 14305; Ganem et al., 1991), a protein containing 129 amino acids and four disulfide linkages. All the *c* type lysozymes that have been compared have a molecular weight of approximately 14-15 kDa, contain 129 or 130 amino acids and analogous positions for the four disulfides are observed. Based on the sequence of HEWL, 28 amino acid positions have been found to be invariant and approximately 40 positions are highly conserved among the known *c* type lysozymes (Nitta & Sugai, 1989; Teahan et al., 1991). A high level of homology (80-98%) exists between HEWL and other avian *c* type lysozymes while the sequence homology between HEWL and mammalian, reptilian and insect lysozymes is somewhat lower (40-60%). The crystal structures that have been determined for the *c* type lysozymes have indicated a strong similarity in the overall protein fold, secondary structural elements and in the positions of the conserved amino acids amongst the *c* type lysozymes. In the case of human lysozyme, a recent report has implicated that point mutations of two highly conserved amino acids may be a cause for amyloidosis in humans, a rare yet fatal hereditary disorder (Pepys et al., 1993).

Evidence has accumulated to suggest that α-lactalbumin is homologous with *c* type lysozymes and that these proteins have diverged from a common ancestor. Similarities have been found in the genetic organization, amino acid sequences and three-dimensional structures of the α-lactalbumin and *c* type lysozyme families (Acharya et al., 1991 and references therein). α-Lactalbumin regulates mammalian lactose biosynthesis by controlling the activity of galactosyltransferase (Hill & Brew, 1975; Kronman, 1989). Albeit very weak, lysozyme activity has been detected for α-lactalbumin (McKenzie & White, 1987; Kumangai et al., 1992a; White et al., 1993).

iii) g type Lysozymes

Fewer members of the *g* type (or goose type) lysozymes have been characterized than the *c* type lysozymes. The *g* type lysozymes are typified by the lysozyme from Emden goose (goose egg white lysozyme; GEWL) and other members identified include the egg white lysozymes from the black swan and ostrich (Jollès & Jollès, 1984). The *g* type lysozymes have a molecular weight of approximately 21 kDa and contain 185 amino acids. The primary sequences of the goose, black swan and ostrich lysozymes have been reported (see Jollès & Jollès, 1984) and a strong homology exists between the sequences (97% for goose and swan; 84% for goose and ostrich; 82% for swan and ostrich). The X-ray structures of GEWL (Grütter et al., 1983; Weaver et al., 1995) and the black swan lysozyme (Isaacs et al., 1985) have been reported and comparison of the two structures has indicated that these lysozymes are very similar (Weaver et al., 1985a).

Immunological studies have shown that the *g* type lysozymes have a much broader taxonomic distribution in bird egg-whites than have the *c* type lysozymes and the two groups do not cross-react immunologically (Arnheim, 1973a; Prager et al., 1974). Both a *g* and a *c* type lysozyme have been detected in the egg white of the black swan (Jollès & Jollès, 1984). The more frequent occurrence of the *g* type than the *c* type lysozymes suggests that the *g* type lysozymes are more commonly present in bird egg whites in order to provide protection to the developing embryo against bacterial infection.

iii) p type Lysozymes

The *p* type (or phage type) lysozymes include those that originate from bacteriophages. Not all bacteriophage lysozymes necessarily belong to the *p* type family as some have been more closely associated with the *ch* type lysozymes. Bacteriophage T4 lysozyme (T4L) is the paradigmatic member of the *p* type family, a protein having a molecular weight of 18.7 kDa and containing 164 amino acids (Tsugita & Inouye, 1968). Other bacteriophage lysozymes that are related by sequence homology and that have been

Table 1.1. Classification of lysozymes of bacteriophage origin.

Phage	Size		Reference†
	MW (kDa)	No. aa	
***P* type**			
T4 (gene e; soluble)	18.7	164	Tsugita & Inouye (1968)
T4bp (baseplate associated; gene 5)		575	Mosig et al. (1989)
P22 (gene 19)	16.1	145	Rennell & Poteete (1985, 1989) Koteswara Rao & Burma (1971)†
ϕ29 (gene 15)	28.0	258	Garvey et al. (1986); Saedi et al. (1987)†
21 (gene *R*)		165	Bonovich & Young (1991)
PA-2 (gene *R*)		165	Bonovich & Young (1991)
PRD1 (P15)	17.3	149	Pakula et al. (1989); Caldenetey et al. (1994)†
***ch* type**			
Cp family		338	García et al. (1988b); Sanz et al. (1993)†
mv1		195	Boizet et al. (1990)
LL-H lysin	34	297	Vasala et al. (1995)†
Tuc2009		428	Arendt et al. (1994)
Others‡			
SF6		316	Verma (1986)
9NA	21		Verma & Siddiqui (1986)†
PL-1	37		Hayashida et al. (1987)†
ϕvML3		226	Shearman et al. (1994)

† References cited pertain primarily to sequence data except for those indicated which also include (some) characterization of the particular lysozyme.
‡ Sequence comparison with other lysozyme types not reported (or found).

included in the *p* type family (Poteete et al., 1992; Jespers et al., 1992) are listed in Table 1.1. The homology between the *p* type lysozymes is weaker than observed between the members belonging to the *c* or to the *g* type families. Although the PA-2 and 21 lysozymes are closely related in sequence to each other (97% identical), the sequence homology amongst the different members (in which gaps have been included during the alignments) ranges from approximately 20-40% (Jespers et al., 1992). Of the 164 amino acids of T4L, there are 14, 9 and 15 which are 100, 80 and 60% conserved respectively among the *p* type lysozymes listed in Table 1.1 (excluding PRD1 lysozyme which was not included in

the analysis; Poteete et al., 1992). To date, only the X-ray structure for T4L has been reported (Matthews & Remington, 1974; Anderson et al., 1981; Weaver & Matthews, 1987). The P22 lysozyme has been modelled onto the structure of T4L (Weaver et al., 1985b).

Over the past twenty years, T4L has been the focus of intensive investigations by many groups but most notably, by Brian Matthews and his colleagues as a model system to study the roles of hydrophobic and electrostatic interactions, hydrogen bonding, and Van der Waals forces as determinants in protein stability and folding (for a review, see Matthews, 1993). To this end, numerous mutant forms of T4L have been prepared and many have been examined by X-ray crystallography (for examples, see Brennan et al., 1988; Rennell et al., 1991; Morton & Matthews, 1995). These contributions have greatly furthered our appreciation of protein structure and function.

iv) ch type Lysozymes

The *ch* type (or *Chalaropsis* type) lysozymes represent a group of lysozymes preliminarily exemplified by the lysozyme from the fungus species *Chalaropsis* (ChL). ChL contains 211 amino acids, has a molecular weight of 22.4 kDa and has been partially characterized (Hash, 1974; Fouche & Hash, 1978). Crystals of the *Chalaropsis* lysozyme have been prepared but the structure itself has not been reported (Lyne et al, 1990). Although ChL possesses no sequence homology to *c*, *g* or *p* type lysozymes (Chang et al., 1979), significant homology has been found between ChL and the lysozymes from the bacteria *Streptomyces erythraeus* (partial sequence known, ≈185 aa, 20 kDa; Morita et al., 1978; Sarma et al., 1979), *Steptomyces globisporus* (217 aa; Lichenstein et al., 1990) and *Clostridium acetobutylicum* (324 aa; Croux & García, 1991). In addition to the usual lysozyme activity, the *Chalaropsis* and *S. erythraeus* enzymes also possess a β-1,4-N,6-*O*-diacetylmuramidase activity.

In recent years, several bacteriophage lysozymes have also been associated with the *ch* type lysozymes (see examples in Table 1.1). The *ch* lysozymes differ substantially in size and there is strong support that the N-terminal domains of the proteins are responsible for the catalytic activity whereas the C-terminal domains are involved in substrate recognition (Sanz & García, 1990; Sanz et al., 1993). Sequence comparisions between the bacteriophage, bacterial and fungal *ch* lysozymes has indicated strong homology between the N-terminal regions (approximately the first 200 amino acids) but no similarity is observed in the C-terminal regions of the proteins (Croux & García, 1991; Vasala et al., 1995). The partial X-ray structure of the *S. erythraeus* lysozyme has been

reported (Harada et al., 1981) and is currently the only structure of a member of the *ch* type family known.

1.2. The Lysozyme Substrate: Peptidoglycan and the Cell Wall

The natural substrate for lysozyme was first demonstrated by Salton (Salton 1952a, b) to be a polysaccharide component of the bacterial cell wall. The cell wall or envelope can be viewed as a multi-component, three-dimensional macrostructure which surrounds the bacterial cell above the plasma or cytoplasmic membrane. Unlike the cytoskeleton of eukaryotic cells, the cell wall is considered to be the prime stress-bearing structure maintaining bacterial shape (Koch, 1983; Beveridge, 1988). Bacteria have been broadly differentiated into two groups, the Gram-positive and Gram-negative bacteria, since the introduction of the classical Gram reaction staining procedure (Gram, 1884). The sensitivity of bacteria to the Gram-reaction was attributed to differences in the basic composition of the cell walls of the two groups (Salton, 1953; Salton, 1963). The primary physiological importance of the cell wall is to withstand the high cellular pressure within bacteria. The estimated turgor pressure is considerable and can range from 2-5 atmospheres in Gram-negative cells and can be 5 to 10 times higher for Gram-positive bacteria (Koch, 1983; Beveridge, 1988; Mendelson & Thwaites, 1989). Any loss in the structural integrity of the cell wall either by physical or enzymatic processes will ultimately result in the lysis of the bacterial cell.

The primary component of the cell wall is the peptidoglycan, a highly water-insoluble heteropolymer consisting of glycan chains interconnected by short peptides. This giant macromolecule forms a net-like structure completely surrounding the cell thereby conferring the majority of the physical strength to the cell wall. A variety of names have been used to designate this macromolecule (eg. glycopeptide, mucopeptide, murein, sacculus) but it is customarily referred to as the peptidoglycan as this name more fittingly describes its polymeric composition (Schleifer & Kandler, 1972). The peptidoglycan exhibits considerable variation among the bacteria (described below) with close to 100 different chemical forms identified (Schleifer & Kandler, 1972; Seidl & Schleifer, 1986). As will be described later, these variations in the peptidoglycan structure appear to influence the activities of lysozymes.

Other differences in the cell walls of Gram-positive and Gram-negative bacteria are listed in Table 1.2. Most notable of these is the presence of an outer membrane in Gram-negative bacteria. The outer-membrane and the peptidoglycan are secured to one another by the lipoprotein, providing further stabilization to the cell wall. This small 58 amino

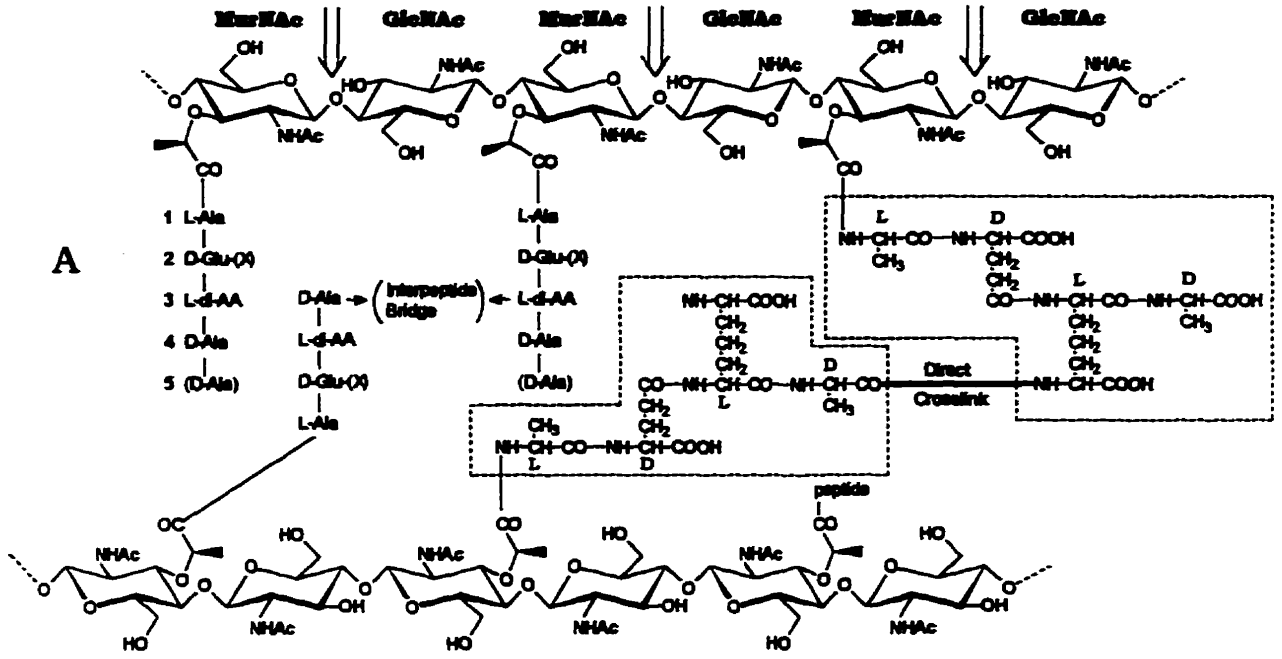
Table 1.2. Comparison of Gram-negative and Gram-positive cell walls. Adapted from Salton (1994).

	Gram-positive	Gram-negative
Peptidoglycan	Present (thick)	Present (thin)
- % weight of wall	40-90%	10-20%
- teichoic acids	present in many	absent
- teichuronic acids	present in some	absent
- lipopolysaccharides	absent†	present
- lipoprotein	absent	present†
Outer Membrane	absent	present
Surface Layers‡	present in some	present in some
Turgor Pressure	20 atm	< 10 atm
Isoelectric Points	pH 2-3	pH 5-6

† Exceptions do exist. ‡ These include capsules, external slimes, sheaths and S layers (reviewed by Beveridge & Graham, 1991).

acid protein has been identified in *E. coli* and other Gram-negative species (Braun & Rehn,1969; Braun et al., 1970). The lipoprotein is covalently linked by the ε-amino of its C-terminal Lys to the carboxyl group of every 10th to 12th diaminopimelic acid (DAP) residue of the peptidoglycan and is anchored into the inner leaflet of the outer membrane through the fatty acid chains covalently attached to the N-terminus of the protein (Braun & Bosch, 1972; Braun, 1975).

The chemical constituents of the glycan and peptide moieties of the peptidoglycan (Cummins & Harris, 1956; Work, 1957) and the compositional arrangement of these constituents (Park & Hancock, 1960; Primosigh et al., 1961; Ghuysen & Strominger, 1963; Ghuysen, 1968; Schleifer & Kandler, 1972; Glauner, 1988; Glauner et al., 1988; Harz et al., 1990) have been determined for several bacteria and has led to a firm description of the structure of the peptidoglycan (Fig. 1.1 A). The glycan segment of the peptidoglycan is rather invariant in the microbial world and is a heteropolymer consisting of alternating β(1→4) linked N-acetylglucosamine (GlcNAc, 2-acetamido-2-deoxy-β-D-glucopyranose) and *N*-acetylmuramic acid (MurNAc, 2-*acetamido*-3-*O*-(D-1-carboxyethyl)-β-D-glucopyranose) residues. Muramic acid is the 3-*O*-D-lactic acid ether of GlcNAc and the peptide chains are bound through their N-termini to the carboxyl group of the lactyl ether. It is the cleavage of the glycoside bond between MurNAc and GlcNAc that is associated with lysozyme activity. The glycan reveals only few variations, including phosphorylation



B

MurNAc

1 **L-Ala** Gly, L-Ser

2 **D-Glu-(X)** 3-Hyg-(X) $\xrightarrow{\alpha}$ X = OH, NH_2, Gly, $GlyNH_2$, D-Ala-NH_2, cadaverine

3 **L-di-AA** *ms*-DAP, L-Lys, L-Orn, LL-DAP, *ms*-HyDAP, L-Dab
L-HyL, L-Hsr, L-Ala, L-Glu, L-Lan

4 **D-Ala**

5 **(D-Ala)**

Figure 1.1. Structure of peptidogylcan. Adapted from Seidl et al. (1986) and Schleifer & Kandler (1972).

(A) Schematic of the fundamental structure of peptidogylcan. The substituents in brackets may or may not be present. Variation of the interpeptide bridge are discussed in the text. The enclosed dashed boxes are representative of the structure of the peptide moieties and an interpeptide bridge characteristic of *E. coli*. Arrows indicate the site of action for lysozyme.

(B) Variations of the structure of the peptide subunit. Shown in bold type is the basic peptide composed of alternating *L-* and *D-* amino acids which may be replaced by the amino acids shown beside them. The substituent X denotes possible substitutions on the α-carboxyl of amino acid at position 2.

Abbreviations: di-AA, diamino acid; 3-Hyg, 3-hydroxyglutamic acid; DAP, *meso*-2,6-diaminopimelic acid; HyDAP, 3-hydroxy-DAP; Dab, diaminobutyric acid; HyL, hydroxy-Lys; Hsr, homoserine; Lan, lanthionine.

(Liu & Gotschlich, 1967; Araki et al., 1971) or acetylation (Ghuysen & Strominger, 1963; Ghyusen, 1968; Clarke & Dupont, 1992) of the C6-hydroxyl group of MurNAc; replacement of MurNAc by N-glycolylmuramic (N-$COCH_2OH$ replaces the N-acetyl group) acid (Azuma et al., 1970) or mannomuramic acid (Hoshino et al., 1972); and the presence of muramic acid δ-lactams found in the peptidoglycan of bacterial spores (Moir & Smith, 1990). The teichoic acids of Gram-positive species are covalently attached via a phosphodiester bond to the C6-hydroxyl of the MurNAc residues (Pooley & Karamata, 1994).

In contrast to the uniform structure of the glycan strands, the peptide moiety reveals considerable variation amongst different genus of bacteria. The general structure of the peptide subunit is a tetrapeptide comprising alternating *L* and *D* amino acids (Fig. 1.1 B). Usually *L*-Ala is bound to the muramic acid, followed by *D*-Glu, which is linked by its γ-carboxyl group (an *iso*-peptide bond) to a *L*-diamino acid, and finally *D*-alanine terminates the peptide. The greatest variation occurs for the diamino acid in position 3. The most widely distributed diamino acid is *meso*-2,6-diaminopimelic acid (*ms*-DAP), which is present in probably all Gram-negative bacteria while *L*-Lys is the other most frequently occurring amino acid at this position (Schleifer & Kandler, 1972; Seidl & Schleifer, 1986). Other possible replacements are illustrated in Fig. 1.1 B.

Individual peptide subunits can be crosslinked to one another. The most common type of cross-linkage (Group A peptidoglycan) forms between the omega amino group of the diamino acid in position 3 of one peptide subunit to the carboxyl group of *D*-Ala in position 4 of another. This cross-linkage can be direct (the primary type in *E. coli* and as illustrated in Fig. 1.1 A) or can involve a variety of interpeptide bridges. The types of interpeptide bridges are numerous but are usually made of between one and six amino acids and can be both homogeneous and heterogeneous in amino acid composition (Schleifer & Kandler, 1972). The second type of cross-linkage (Group B peptidoglycan) is much less frequent and involves a diamino acid forming a bridge between the α-carboxyl of Glu at position 2 with the carboxyl of *D*-Ala at position 4 of an adjacent peptide subunit. In all cases, peptide crosslinking involves removal of the *D*-Ala at position 5. More recently, a new type of direct crosslinkage, first described in *E. coli* (Glauner, 1988; Glauner et al., 1988) and now in other Gram-negative species (Quintela et al., 1995b) has been identified. The crosslink forms as a *L*-*D*-peptide bond between the omega amino (the *D* centre) of one DAP residue and the α-carboxyl (the *L* centre) of another DAP residue which would involve displacement of the *D*-Ala at position 4.

Original models proposed for the peptidoglycan described a relatively flat and straight arrangement of the glycan chains in which alternating sugar monomers were rotated approximately 180° to each other (similar to that of cellulose and chitin) and the peptides extended essentially in the same direction from the glycan axis (Formanek et al., 1974; Oldmixon et al., 1974). In light of diffraction, spectroscopic and modelling studies (Burge et al., 1977; Labischinski et al., 1979; Barnickel et al., 1979; Naumann et al., 1982) a model has been suggested in which the bulky lactyl ether induces the glycan strands to adopt a helical conformation (Fig. 1.2). The peptides along the glycan chain consecutively extend approximately up, left, down and right from the helix axis. The peptidoglycan is now considered to have relatively flexible and elastic properties (Labischinski & Maidhof, 1994). The peptidoglycan of Gram-positive bacteria are multilayered and thicker than that of Gram-negative bacteria (Labischinski & Maidhof, 1994). There is some debate as to whether the peptidoglycan of Gram-negative bacteria is mono- or multilayered (Wientjes et al., 1991). In *E. coli*, it has been suggested that the majority of the peptidoglycan surface (75-80%) is monolayered except for some localized regions where the peptidoglycan is thicker and multi-layered (Wientjes et al., 1991; Labischinski & Maidhof, 1994).

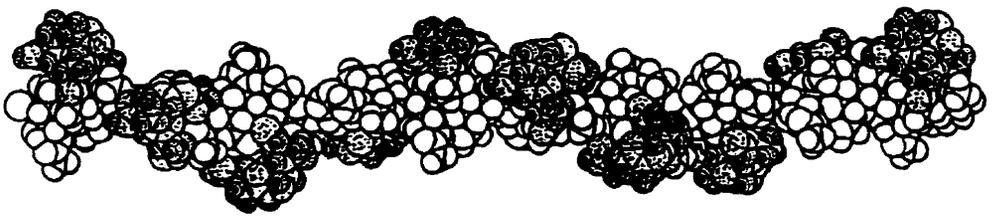


Figure 1.2. Space-filling model of the peptidoglycan. Shown are 10 disaccharide (GlcNAc-MurNAc) peptide units. Sugar atoms appear white while peptide atoms are dotted. Reproduced from Labischinski & Maidhof (1994).

The length of the glycan chains in Gram-positive bacteria can range from 30-40 (Henze et al., 1993) to as high as 100-200 disaccharide units (Rogers et al., 1980). The length appears shorter in Gram-negative bacteria. In *E. coli*, the chain length is quite variable but the majority of the peptidoglycan is made from glycan strands of 5-10 disaccharide units (69.3% of glycan strands have an average length of 8.9 disaccharides while 30.7% have an average length of 45.1 disaccharide units yielding an average chain

length of 21 disaccharide units; Harz et al., 1990). The average chain length in *E. coli* shows some dependency on the growth conditions (Glauner et al. (1988) found the average chain length to be 33 disaccharide units). The average glycan length of the Gram-negative bacterium *Thermus thermophilus* was also determined to be 30 disaccharide units (Quintela et al., 1995a).

In addition, the peptidoglycans of most bacteria have virtually all of their muramic acid residues substituted with peptide, the known exceptions being *Micrococcus lysodiekticus* and some related *Micrococcaceae* in which less than 50% of the muramic acids are substituted (Ghuysen, 1974). These bacteria, of which *M. lysodiekticus* has been intensively investigated with HEWL, are also unique in that they are the only example in which crosslinking is achieved by an interpeptide bridge that has the same amino acid sequence as the peptide units that substitute the muramic acid residues (i.e. *L*-Ala-*D*-*iso*-Glu(Gly)-*L*-Lys-*D*-Ala). This bridge is formed between the terminal *D*-Ala carboxyl of one MurNAc substituted peptide and the ε-amino of Lys of another MurNAc substituted peptide with the amino and carboxyl termini of the peptide bridge respectively. The fact that over half of the MurNAc residues are unsubstituted in *M. lysodiekticus* was indeed fortunate in the early studies with HEWL allowing for the isolation and characterization of peptide-free $(GlcNAc\text{-}MurNAc)_n$ oligomers after digestion of peptidoglycan from this bacteria by HEWL (Sharon, 1967; Ghuysen, 1968).

The helical arrangement of the glycan chains and the spatial arrangement of the peptide subunits restricts crosslinking to only those peptides which are close in space. The mostly mono-layered structure of Gram-negative bacteria and scarcity of interpeptide bridges leads to a relatively low degree of crosslinking of 24-30% (this is typical for *E. coli*) with values as high as 60% reported (Schleifer & Kandler, 1972; Glauner, 1988; Glauner et al., 1988; Quintela et al., 1995a,b). The multilayered arrangement of peptidoglycan in Gram-positive species and an increased presence of interpeptide bridges results in higher levels of crosslinking of up to 70% (Labischinski & Johannsen, 1986). In *Staphylococcus aureus*, 90% of the peptides are crosslinked which is facilitated by their long and flexible penta-Gly interpeptide bridges which can span larger distances (Labischinski & Maidhof, 1994). Crosslinking in both Gram-negative and Gram-positive species can occur between peptides on the same (intra-strand, involving 2 consecutive disaccharide-peptide units) or different (inter-strand) glycan chains, although the latter is more common.

In recent years, our understanding of the intricate structure of peptidoglycan has expanded with the development of analytical methods to identify components produced from the digestion of isolated peptidoglycan with specific peptidoglycan hydrolases. A

recent report exemplifies the judicious application of these methods in the compositional analysis and structural determination of the peptidoglycan from *T. thermophilus* (Quintela et al., 1995a). Unfortunately, the higher variation in the muropeptide composition of Gram-positive bacteria has made their separation more difficult although reports have appeared (Garcia-Bustos et al., 1988; De Jong et al., 1993).

Similar analyses on the *E. coli* peptidoglycan have revealed a remarkable complexity in which over 80 different muropeptides have been separated and many identified (Glauner, 1988; Glauner et al., 1988; Höltje & Glauner, 1990). These studies have indicated that about 55% of the MurNAc residues are substituted with the tetra-peptide *L*-Ala-*D*-*iso*-Glu-*ms*-DAP-*D*-Ala and that 70% of these are free (i.e., 30% are involved in crosslinks). The free tetra-peptide is then the major substituent of the *E. coli* peptidolgycan. In addition, the glycan strands in *E. coli* peptidoglycan terminate with a 1,6-anhydro-muramic acid (*anh*MurNAc), a consequence of the normal metabolism of the peptidoglycan during cell growth by the action of the lytic transglycosylases (see section 1.4). It now appears that 1,6-anhydro-muropeptides are normal components of the Gram-negative peptidoglycan as they have been found in several of these species (Sinha & Rosenthal, 1980; Quintela et al., 1995a,b).

1.3. Bacteriophage Lambda Lysozyme

1.3.1. Phage Induced Lysis and the λ R Gene

Jacob and Fuerst (1958) initially demonstrated that lysates of *E. coli* induced by bacteriophage lambda (λ) contained a protein, designated as λ endolysin that is capable of degrading peptidoglycan. Endolysin is a generic name given to phage-encoded enzymes that exhibit peptidoglycan degrading activity causing host cell lysis. The λ endolysin was identified as the product of the λ *R* gene (Campbell, 1961) although the specificity of the endolysin remained uncertain. Two activities were associated with the λ endolysin; i) an endopeptidase responsible for the cleavage of the peptide crosslinkages present in the peptidoglycan (Taylor, 1971); and ii) a hitherto unknown type of N-acetylmuramidase or lysozyme activity (Taylor & Gorazdowska, 1974; Taylor et al., 1975). Mutagenesis experiments indicated that both the endopeptidase and lysozyme were found to contribute to the lytic activity of λ lysates and had also indicated the involvement of the λ *Rz* gene product (Young et al., 1979). Further studies had established that the lysozyme was indeed the λ *R* gene product (Bienkowska-Szewczyk et al., 1981).

During the course of the lytic cycle of phage λ and other bacteriophage, lysis of the host cells is promoted at the end of the lytic cycle by the breakdown of the peptidoglycan.

For phages of both Gram-negative and Gram-positive bacteria, it is believed that lysis involves a system including not only enzymes which act on the peptidoglycan, but also the requirement of proteins, which have been termed holins, to create lesions in the cytoplasmic membrane for the access of the endolysins to the peptidoglycan (for a review on the general mechanism of bacteriophage lysis, see Young, 1992; for a review on holins, see Young & Bläsi, 1995).

The λ *R* gene is transcribed as part of a cistron (the lysis cassette) which like all the late genes required for host cell lysis, phage DNA packaging and morphogenesis, is under control of the λ promoter p'_R. The λ lysis cassette consists of three genes, the *S*, *R*, and *Rz* genes which form an overlapping cluster (i.e. 5' *SRRz*). The genetic organization of the lysis cassette in some other bacteriophage is very similar to that of phage λ (Young, 1992). The λ *S* gene product is a holin and permits the diffusion of accumulated λ endolysin at the onset of lysis across the inner membrane to the periplasm and the peptidoglycan (Garrett & Young, 1982; Young, 1992). Two mature forms of the S protein, S105 and S107 (the two S proteins differ only at the N-terminal in which S105 lacks the first two amino acids of S107) are synthesized at the onset of lysis (Chang et al., 1995). S105 is the actual holin while S107 inhibits holin function (Chang et al., 1995). The function of the *Rz* gene product is unknown but it is believed to encode for the endopeptidase activity of λ endolysin (Bienkowska-Szewczyk & Taylor, 1980). Recently, the *Rz* gene product has been overexpressed in *E. coli* but was not characterized (Hanych et al., 1993). However, it was found that the overexpression of the *Rz* gene product in *E. coli* was toxic to the cells (Hanych et al., 1993). Endolysins from *Pseudomonas* phage ϕ6 (Caldentey & Bamford, 1922) and *Listeria* bacteriophages (Loessner et al., 1995) have been identified and characterized as endopeptidases, the activity of the latter being the first description of an *L*-alanyl-*D*-glutamate peptidase.

1.3.2. What is Known about Lambda Lysozyme (LaL)

Bacteriophage lambda lysozyme (lambda lysozyme; LaL; λL) was first purified to apparent homogeneity and partially characterized by Black and Hogness (1969a,b,c). The primary structure was determined from protein sequencing but the sequence reported included a minor error (Imada & Tsugita, 1971). The determination of the complete nucleotide sequence of λ DNA (Sanger et al., 1982) has provided the correct and accepted amino acid sequence for the enzyme (Fig. 1.3). LaL is a small protein containing 158 amino acids and having a molecular weight of 17825 Da. The basic character of the protein is suggested by the difference between the number of basic (24) and acidic (20)

```
1          5          10         15         20         25         30
M V E I N N Q R K A F L D M L A W S E G T D N G R Q K T R N
31         35         40         45         50         55         60
H G Y D V I V G G E L F T D Y S D H P R K L V T L N P K L K
61         65         70         75         80         85         90
S T G A G R Y Q L L S R W W D A Y R K Q L G L K D F S P K S
91         95         100        105        110        115        120
Q D A V A L Q Q I K E R G A L P M I D R G D I R Q A I D R C
121        125        130        135        140        145        150
S N I W A S L P G A G Y G Q F E H K A D S L I A K F K E A G
151        155   158
G T V R E I D V
```

Ala:	13	**Cys:**	1	**His:**	3	**Met:**	3	**Thr:**	6
Arg:	12	**Gln:**	9	**Ile:**	9	**Phe:**	5	**Trp:**	4
Asn:	6	**Glu:**	7	**Leu:**	14	**Pro:**	5	**Tyr:**	5
Asp:	13	**Gly:**	15	**Lys:**	12	**Ser:**	9	**Val:**	7

Figure 1.3. Amino acid sequence and composition of LaL.

amino acid residues. Sedimentation analysis on LaL provided a molecular weight of 17.9 kDa (Black & Hogness, 1969a) indicating that the enzyme is monomeric and the single cysteine (Cys120) is present in the reduced form. The three-dimensional structure of LaL is currently unknown.

In 1971, Taylor reported that *E. coli* peptidoglycan subjected to a partially purified preparation of LaL resulted in the production of low molecular weight products, evidence that the peptidoglycan had been fragmented, yet no reducing groups in the products could be detected (Taylor, 1971). Separation and isolation of the peptidoglycan fragments resulted in two main products, which were termed muropeptides CA and CB (Taylor & Gorazdowska, 1974). Acid hydrolysis of CA and CB indicated that they contained MurNAc, GlcNAc, Ala, Glu, and DAP yet it remained difficult to explain the absence of any reducing sugar in CA and CB.

The cause clarifying the lack of reducing groups came upon elucidation of the structures of CA and CB by mass spectral analysis (Taylor et al., 1975). Both CA and CB were found to contain a 1,6-anhydro-N-acetylmuramic acid (*anh*MurNAc) at their respective reducing positions (Fig. 1.4). The products produced by other lysozymes such

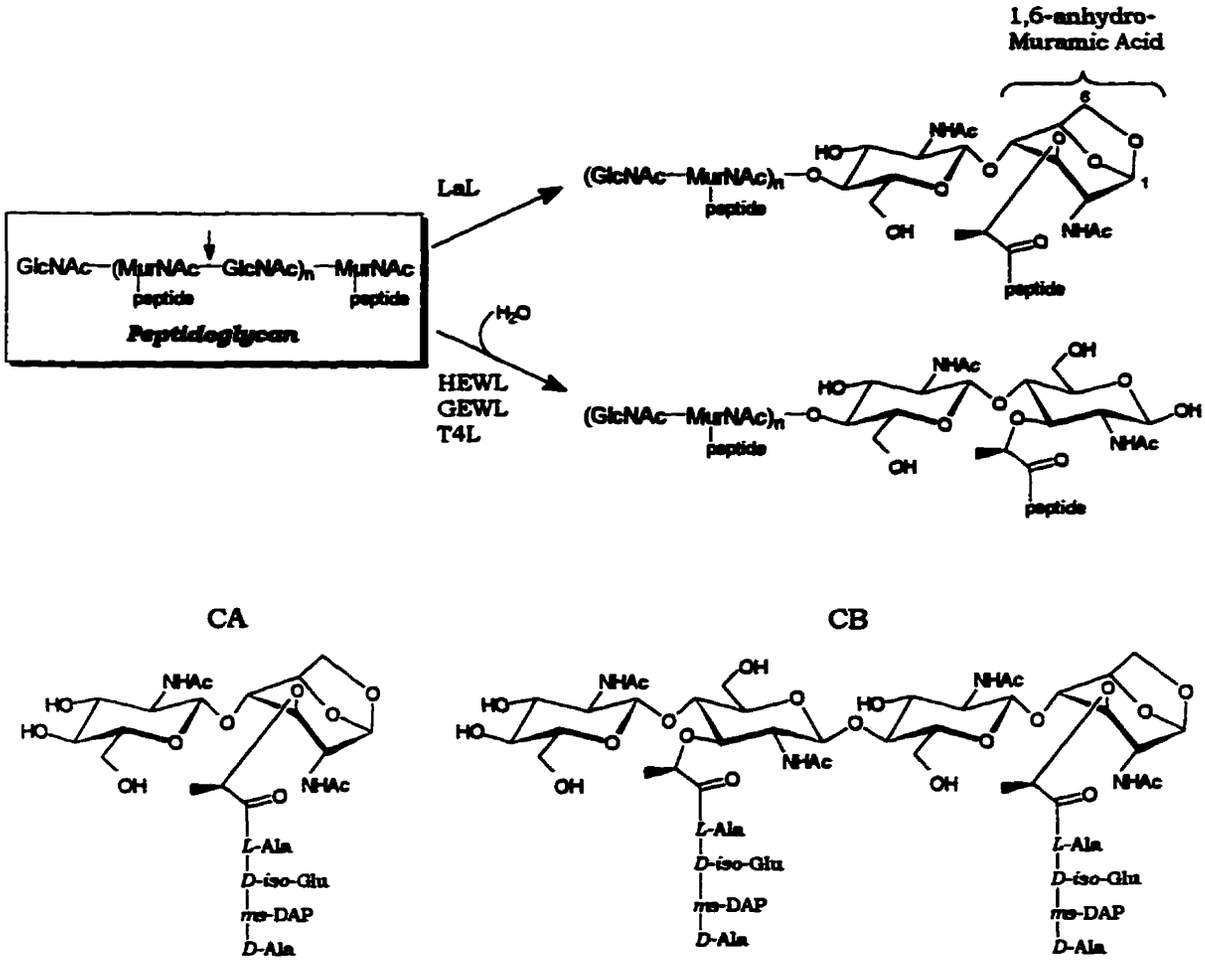


Figure 1.4. 1,6-Anhydromuramic acid production by LaL. Lambda lysozyme cleavage of the glycosidic bond between MurNAc and GlcNAc results in the formation of a 1,6-anhydro-muramic acid at the reducing end of the products. The structures of muropeptides CA and CB obtained from the digestion of *E. coli* peptidoglycan by LaL are shown.

as HEWL, GEWL and T4L contain free reducing ends as these enzymes act as hydrolases and employ water to hydrolyse the glycosidic linkages of the peptidoglycan (Fig. 1.4). Lambda lysozyme appears to differ in its peptidoglycan cleaving activity. Instead of employing water as the nucleophile, LaL promotes cleavage of the glycosidic linkages by way of an internal transglycosylation reaction utilizing the C6-hydroxyl group as the nucleophile. In this sense, LaL is a transglycosylase and therefore, does not elaborate the true definition of lysozyme activity, i.e. β-1,4-N-acetylmuramylhydrolase. However, those that have been engaged in the study of this enzyme have recognized this distinction yet continue, as have we, to refer to the *R* gene product as a lysozyme because of its specific

peptidoglycan cleaving activity. The degradation of peptidoglycan into fragments containing *anh*MurNAc has also been reported for the endolysin produced by phage Vi II (Taylor & Gorazdowska, 1974; Taylor et al., 1975) in addition to the family of *E. coli* lytic transglycosylases (discussed in section 1.4).

Apart from the studies indicated above, little is known concerning the distinctive properties of LaL that distinguishes this novel lysozyme from other lysozymes whose mechanistic and structural features have been investigated in greater detail. In recent years, Fastrez and colleagues have selected LaL as a model protein in the study of translational signals (Jespers et al., 1991) and tryptophan starvation (Soumillion & Fastrez, 1992) and their effect on expression efficiency. The incorporation of 7-azatryptophan and characterization of some of the physical properties of the enzyme has also been reported (Soumilliion et al., 1995).

1.4. *E. coli* Lytic Transglycosylases

In order for the peptidoglycan to "grow" with the bacteria, there must exist an intricate and regulated balance between peptidoglycan synthesizing and peptidolglycan hydrolyzing activities such that new subunits can be inserted into the pre-existing sacculus (for reviews on peptidoglycan biosynthesis see Rogers et al., 1980; Bugg & Walsh, 1992; Van Heijenoort, 1994; Matsuhashi, 1994; for reviews on peptidoglycan hydrolases see Rogers et al., 1980; Höltje & Tuomanen, 1991; Shockman & Höltje, 1994). Solubilization of peptidoglycan can be achieved by three classes of peptidoglycan hydrolases; glycosidases cleave the glycosidic bonds between the carbohydrate groups; *N*-acetylmuramyl-*L*-alanine-amidases cleave the amide bond between the *L*-Ala and the lactyl group of MurNAc; and peptidases which cleave within the peptide chains. Each of these endogenous activities have been found in *E. coli* and their sites of action on the peptidoglycan are shown in Fig. 1.5.

A lysozyme-like muramidase is present in *E. coli*. However, like LaL, the enzyme was found to be an intramolecular transglycosylase catalyzing cleavage of peptidoglycan with concomitant production of 1,6-anhydromuramic acid (Höltje et al., 1975). The enzyme has been termed lytic transglycosylase to avoid confusion with the biosynthetic transglycosylases which polymerize the murein precursors during peptidoglycan biosynthesis. It has been proposed that part of the chemical energy of the β-1,4-glycosidic bond that is cleaved is conserved in the newly synthesized 1,6-anhydro bond and that this

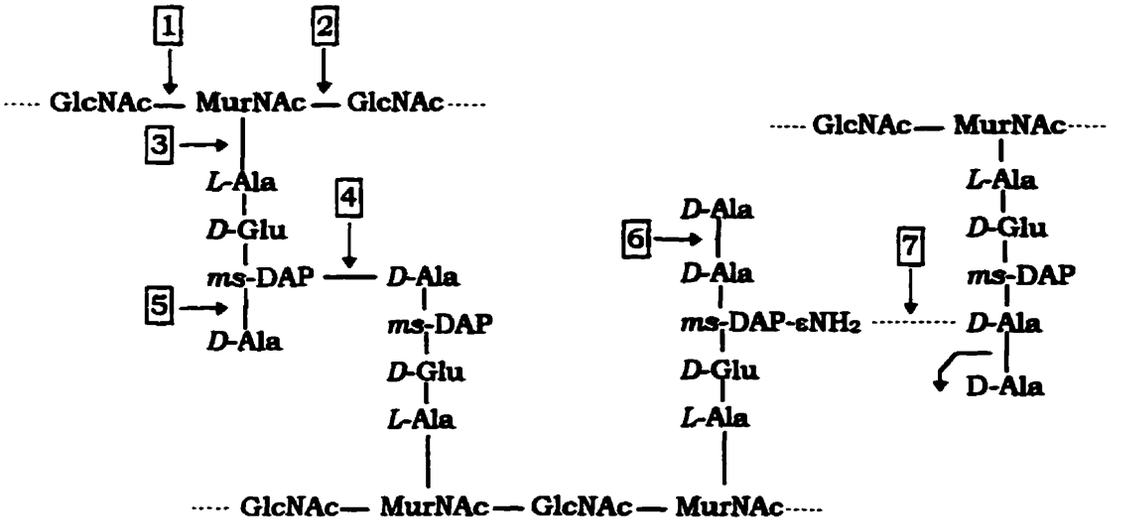


Figure 1.5. Places of action of *E. coli* peptidoglycan hydrolases. (1) β-N-acetylglucosamidase (2) lytic transglycosylase; (3) N-acetylmuramyl-L-alanine-amidase; (4) D,D-endopeptidase; (5) L,D-carboxypeptidase; (6) D,D-carboxypeptidase. Also shown is the site of action of D,D-transpeptidase (7) which is involved in the formation of peptide crosslinks. Adapted from Mirelman (1979) and Höltje & Tuomanen (1991).

energy may be utilized during other reactions involving recycling, rearrangements or repair to the existing peptidoglycan structure (Höltje et al., 1975).

Three distinct lytic transglycosylases have been purified from *E. coli*; Slt70 (Engel et al., 1991), Slt35 (Engel et al., 1992a) and Mlt38 (Ursinus & Höltje, 1994). Both Slt70 (for soluble lytic transglycosylase, 70 kDa) and Slt35 are soluble proteins found in the periplasm while Mlt38 is membrane-bound. The lytic transglycosylases are exomuramidases and are believed to cleave the peptidoglycan from the non-reducing GlcNAc termini of the glycan strands (Romeis et al., 1993). Mlt38 differs from the two soluble enzymes in its ability to degrade $(MurNAc\text{-}GlcNAc)_n$ (Ursinus & Höltje, 1994) while only Slt70 is inhibited by the glucosamine derivative bulgecin whereas Slt35 and Mlt38 are not (Templin et al., 1992; Dijkstra & Thunnissen, 1994). Some sequence homology has been detected between the C-terminal domain of Slt70 and other bacterial and bacteriophage proteins (Koonin & Rudd, 1994). The interest in the lytic transglycosylases has been enhanced by the discovery that *anh*MurNAc containing muropeptides exhibit several biological activities (see section 1.11).

The three-dimensional structure of Slt70 has been recently reported (Thunnissen et al., 1994; Thunnissen et al., 1995a). The 618 residue protein folds into three domains (Fig. 1.6). The N-terminal U-domain and linker-domain are highly helical in nature and form a ring or "doughnut" shaped structure with a large central hole approximately 25-30 Å in diameter. The C-domain which comprises residues 449-618, is believed to contain the active site. This C-domain or C-Slt70 lies on top of the ring leaving the hole accessible. It has been suggested that glycan chains of the peptidoglycan pass through the central hole towards the C-domain and the active site. Furthermore, this organization may account for the exomuramidase activity of Slt70 as only the "loose" ends of the peptidoglycan could access and pass into the hole (Dijkstra & Thunnissen, 1994; Thunnissen et al., 1995b). As will be discussed in the section 1.6.1, the overall structural features of C-Slt70 show a strong resemblance to other known lysozyme structures even though the sequence similarties are weak.

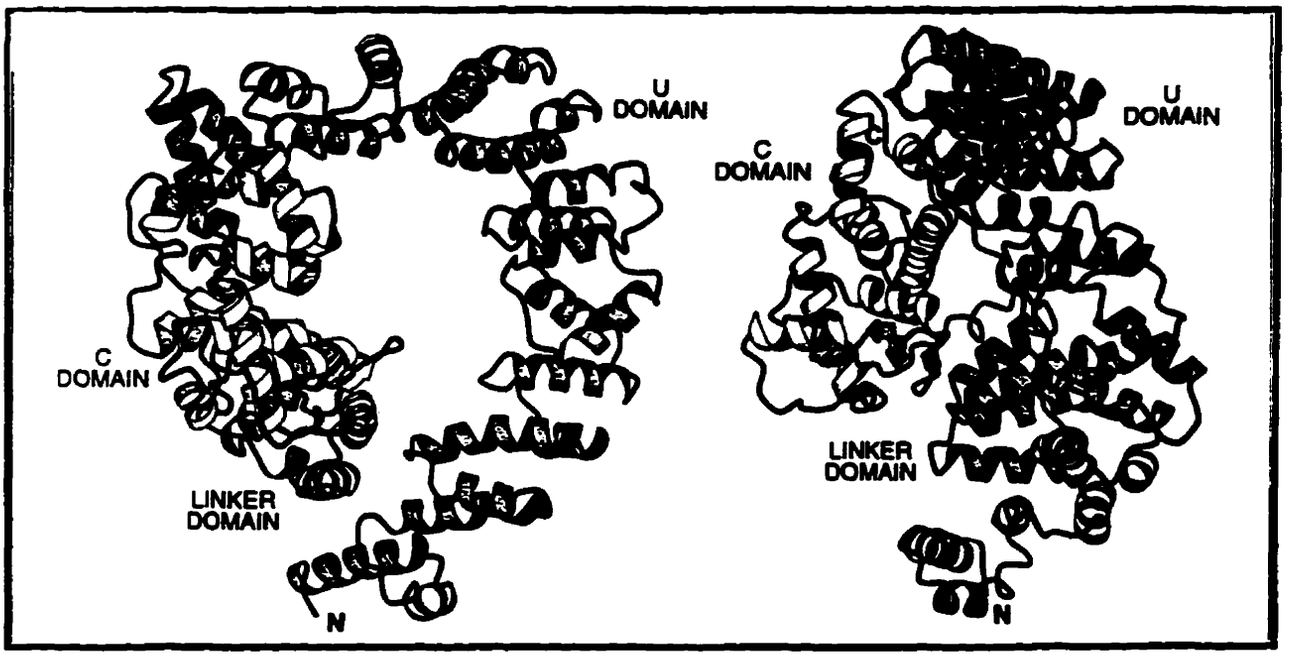


Figure 1.6. Structure of the *E. coli* soluble lytic transglycosylase Slt70. Reproduced from Dijkstra & Thunnissen, 1994.

Left: Ribbon representation of the structure looking down the central hole.

Right: Same as above but viewed from the side of the molecule showing the position of the C-domain situated on top of the helical ring.

1.5. The Phillips Model for Hexasaccharide Binding and the Lysozyme Mechanism

The currently accepted mechanism for the cleavage of substrates by lysozyme was initially proposed by Phillips and colleagues following the determination of the X-ray structures of HEWL cocrystallised with a number of small inhibitors (Phillips, 1966; Blake et al., 1967; Phillips, 1967). The structure of the HEWL/$(GlcNAc)_3$ complex indicated that $(GlcNAc)_3$ occupied a stable position within the cleft of the enzyme. The binding subsites for six monosaccharides, labelled A to F, were recognized which led to the development of the Phillips model for hexasaccharide binding (Fig. 1.7). The binding mode of the first three residues in sites A-C was determined experimentally from the crystallographic studies while the existence of sites D-F was deduced from the modelling of additional GlcNAc residues onto the $(GlcNAc)_3$ complex. The model indicated that the binding of a GlcNAc residue in the normal full-chair conformation in the D site would lead to seriously close contacts between the enzyme and the saccharide. This overcrowding could be relieved by distortion of the D site pyranose ring to a conformation which resembled a half-chair or sofa conformation. Inspection of the model and the knowledge that $(GlcNAc)_6$ is cleaved preferentially between the fourth and fifth residues (discussed in section 1.9.1) suggested that cleavage of the glycosidic bond would occur between the saccharides

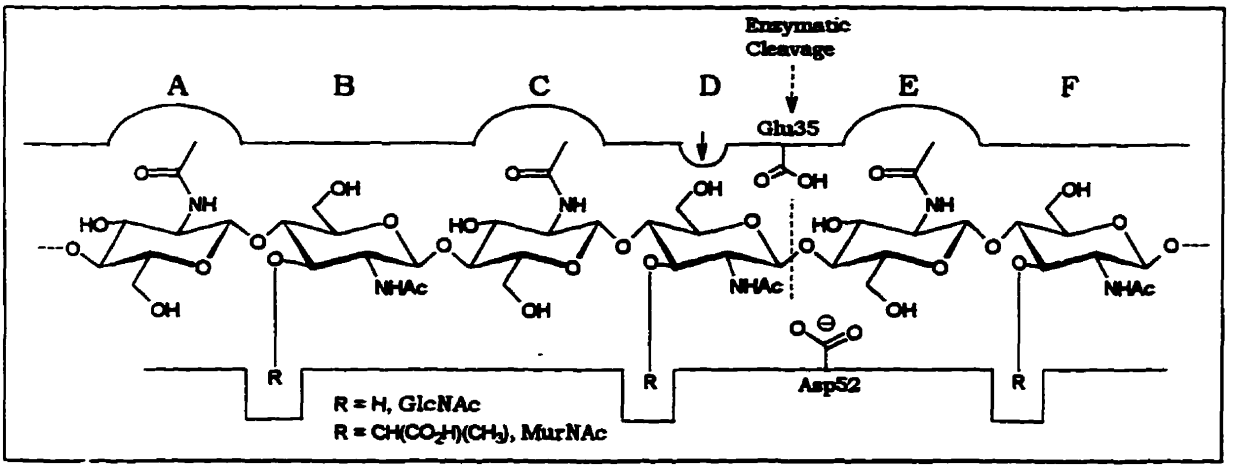


Figure 1.7. Proposed Phillips model for hexasaccharide binding to HEWL. The active site cleft contains six monosaccharide binding sites A-F and cleavage occurs between the residues in the D and E sites. The acetamido groups of residues in sites A, C, and E make specific hydrogen bonding interactions with the enzyme. The model indicated that neither the C nor E sites could accommodate the lactyl side chain of a MurNAc residue but binding of MurNAc can be accomodated in sites B, D and F. Steric interactions (indicated by the solid arrow) existed between the enzyme and for a D site residue which was modelled in a full-chair conformation. Modelling of an oligosaccharide larger than a hexamer so that more than six residues make contact with the enzyme was not possible.

located in sites D and E. Furthermore, the model depicted a position for this labile bond within the cleft that was located very near to the side chains of Glu35 and Asp52. Other relevant features of the Phillips model are described in the legend to Fig. 1.7.

The modelled position of the hexasaccharide within the cleft of HEWL led immediately to a proposal for the catalytic mechanism. The mechanism of lysozyme action as proposed by Phillips is presented in Fig. 1.8 and the following discussion will make reference to the individual intermediates A-E along the pathway.

(A). For the reasons outlined in Fig. 1.7, the substrate binds such that the labile glycosidic bond between the site D MurNAc and site E GlcNAc residues is flanked by Glu35 and Asp52. The environments occupied by these residues suggested that Glu35 is protonated while Asp52 is ionized.

(B). Optimal binding of the substrate requires distortion of the site D pyranose ring towards a half-chair conformation consequently weakening the bond between atoms C1 and O4 of residues D and E. Bond rearrangement is promoted by the lone pair electrons of the site D ring oxygen. Glu35, acting as a general acid, would donate a proton to the glycosidic oxygen.

(C). Bond cleavage results in the formation of an oxocarbonium ion intermediate. The enzyme-bound oxocarbonium ion is stabilized by the half-chair conformation required of ring D and electrostatically by the neighbouring negative charge of Asp52 which forms an ion pair with the oxocarbonium.

(D). The lifetime of the oxocarbonium-Asp52 ion pair is of a duration that allows the aglycone in the E site to diffuse away and a water molecule (or another acceptor molecule, i.e. ROH) to diffuse in and attack C1 of the oxocarbonium ion. Glu35 would assist as a general base.

(E). The participation of Glu35 and the proximity of Asp52 prevents attack of water from the α direction and the reaction is completed with retention of configuration (β-anomer) in the product.

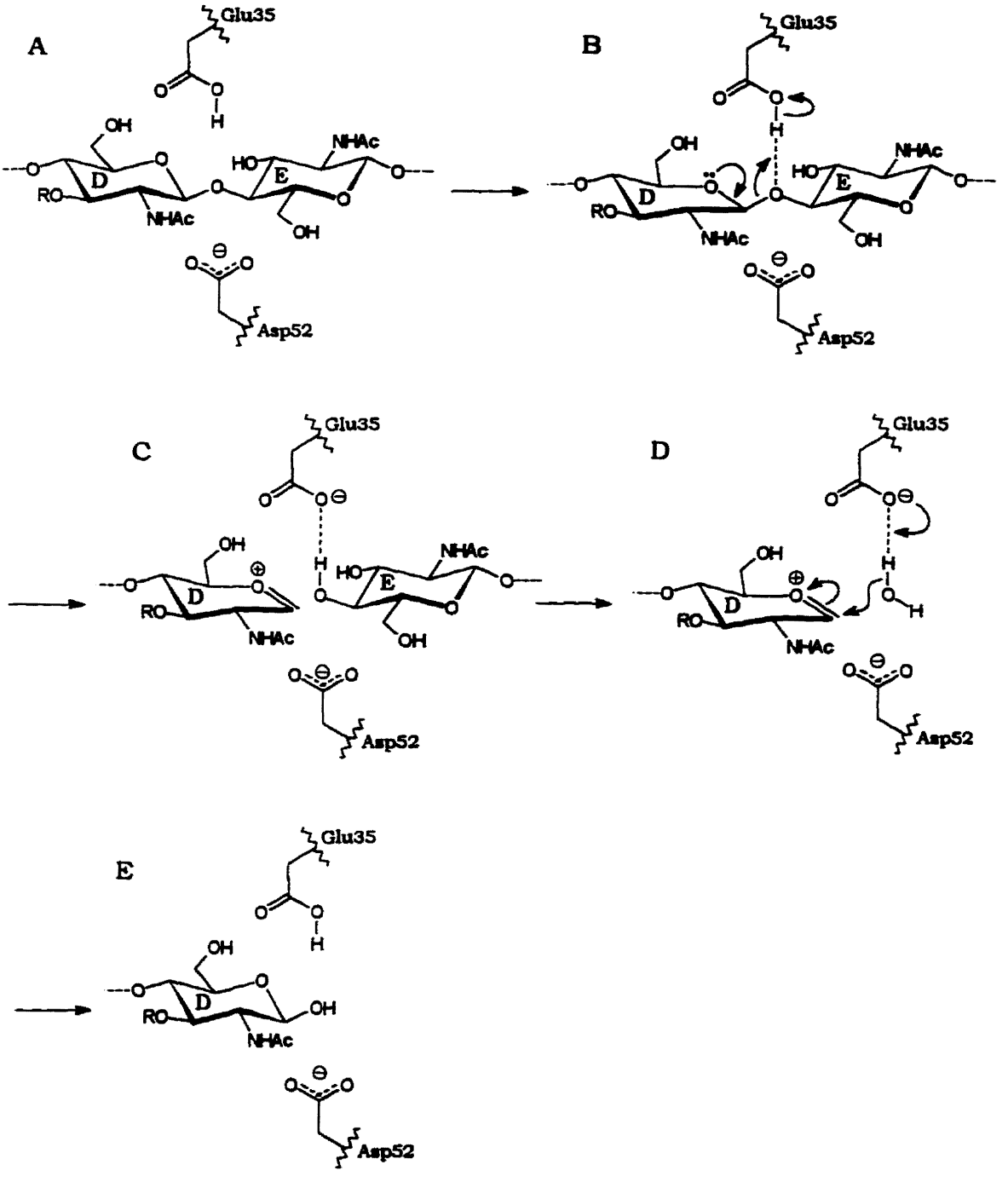


Figure 1.8. The catalytic mechanism for HEWL as proposed by Phillips.

1.5.1. Evidence Supporting the Phillips Mechanism

Recently, Strynadka & James (1991) have reviewed the current standing of the Phillips mechanism. In addition to detailing a comprehensive summary of previously reported investigations, the authors have also provided additional crystallographic evidence that is strongly supportive of the proposed mechanism. Some aspects characteristic to the mechanistic proposal for HEWL have also been suggested as essential mechanistic features of other lysozymes.

All sequences of *c* type lysozymes have a Glu and Asp in analogous positions to Glu35 and Asp52 of HEWL (Nitta & Sugai, 1989). As indicated in Table 1.3, alteration of these residues in HEWL by mutagenesis, chemical modification or affinity labelling is deleterious to enzyme activity. The existence of corresponding acidic residues in other lysozymes has been suggested from similar studies (Table 1.3). In those enzymes for which the evidence has supported the involvement of residues equivalent to both Glu35 and Asp52, modification of the "Glu35" residue is apparently more harmful to activity than modification of the "Asp52" residue. If the analogous "Glu35" residues of the various lysozymes function similarly and are assumed to act as general acids, it appears that this role is more critical to activity than any possible role of the "Asp52" residues. It has been suggested from crystallographic analysis that GEWL (Weaver et al., 1995) and Slt70 (Thunnissen et al., 1994) have no true counterpart to Asp52 of HEWL although some possible candidates have been considered (Table 1.3).

Table 1.3. Catalytic acid residues suggested for several lysozymes.

Enzyme (and Type)	Modification to Indicated Residue and Resulting Residual Activity or Consequence	Reference
HEWL (c)	<u>Glu35</u>	
	-mutation to Gln; 0%	Malcolm et al., 1989
	-chemically converted to Gln; 3-4%†	Kuroki et al., 1986
	-modified with ethylenimine; 3%†	Yamada et al., 1982
	<u>Asp52</u>	
	-mutation to Asn; ≈ 5% residual activity	Malcolm et al., 1989
	-mutation to Ser; <1%	Lumb et al., 1992
	-mutation to Ala; <4%	Matsumura & Kirsch, 1996
	-chemically converted to Asn; 1-3%†	Kuroki et al., 1986
	-modified with ethylenimine; 1.4%†	Yamada et al., 1982
	-specifically labelled by 2',3'-epoxypropyl β-glycoside of $(GlcNAc)_2$ with loss of activity	Eshdat et al., 1973; Moult et al., 1973
	-chemically converted to homoserine; <10%†	Eshdat et al., 1974

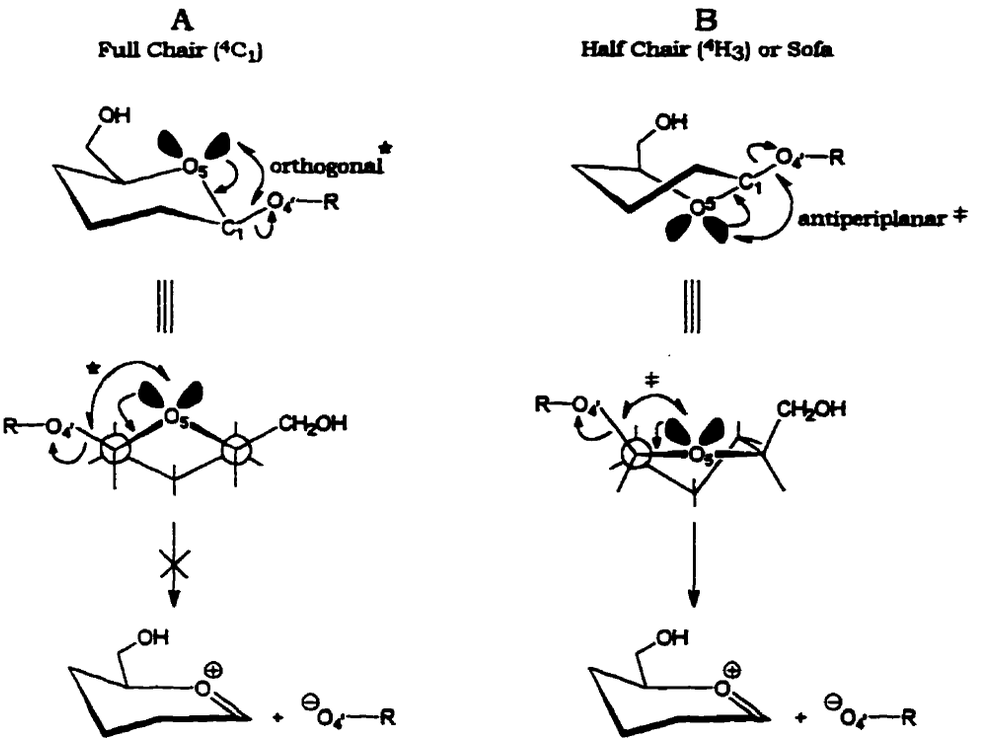
Table 1.3. cont'd.

Enzyme (and Type)	Modification to indicated Residue and Resulting Residual Activity or Consequence	Reference
HuL (*c*)	Asp53: analogous to Asp52 of HEWL	
	-mutation to Glu; <5%	Muraki et al., 1991
	-specifically labelled by 2',3'-epoxypropyl β-glycoside of $(GlcNAc)_2$ with loss of activity	Muraki et al., 1996
T4L (*p*)	Glu11: proposed analogy to Glu35 of HEWL	
	-mutation to i) Asp; 16% and ii) Gln; 0%	Anand et al., 1988
	-mutation to 14 other amino acids abolishes halo forming ability‡	Renell et al., 1991; Kuroki et al., 1995
	Asp20 : proposed analogy to Asp52 of HEWL	
	-mutation to i) Glu; 1% and ii) Asn; 1%	Anand et al., 1988
	-mutation to Cys; 80%	Hardy & Poteete, 1991
	-of 16 mutants, 11 retain while 5 lose halo forming ability‡	Renell et al., 1991; Kuroki et al., 1995
GEWL (*g*)	Glu73 : proposed analogy to Glu35 of HEWL	Weaver et al., 1985, 1995
	Asp? : no apparent counterpart to Asp52 of HEWL; Asp86 and Asp97 have been considered	
CPL-1 (*ch*)	Asp9: proposed analogy to Glu35 of HEWL	Sanz et al., 1992
	-mutation to Glu or Asn; ≈2%	
	-mutation to His, Ala, Lys; <0.1%	
	Glu36: proposed analogy to Asp52 of HEWL	
	-mutation to Asp; 37% or Gln; 67%	
	-mutation to Ala; 5%	
ChL (*ch*)	Asp6 and Glu33 were identified from chemical modification and protection studies	Fouche & Hash, 1978
Slt70	Glu478: proposed analogy to Glu35 of HEWL	Thunnissen et al., 1994
	-mutation to Gln abolishes activity	
	Asp?: no apparent counterpart to Asp52 of HEWL; Glu583 has been considered	
LaL	Glu19: proposed analogy to Glu35 of HEWL	Jespers et al., 1992
	-mutation to Gln abolishes activity	
	Asp34: proposed analogy to Asp52 of HEWL	
	-mutation to i) Asn; 5% and ii) Ala; <1%	

† In the cases of chemical modifications, part of the residual activity has been attributed to unreacted enzyme or regenerated wild type enzyme.
‡ The mutant lysozyme is expressed from phage that have been plated onto a bacterial lawn. The formation of a zone of clearing or halo indicates activity.

The evidence for the involvement of Glu35 of HEWL as a general acid is very strong. The crystal structure of HEWL complexed with MurNAc-GlcNAc-MurNAc (MGM) bound in sites B-D indicated that the side chain of Glu35 is not only ideally positioned to serve as a proton donor but is hydrogen bonded to the anomeric oxygen of the MurNAc in site D (Strynadka & James, 1991). Furthermore, several studies have suggested that the pK_a for Glu35 is 5.9-6.3 in the native enzyme and is raised by 0.4-0.6 when the enzyme binds substrates (Parsons & Raftery, 1972a,b; Banerjee et al., 1974,1975; Spassov et al., 1989). These studies have also indicated that the pK_a for Asp52 is 3.8-5.1 and is not perturbed on substrate binding. Therefore, at the pH of maximal activity for HEWL (pH 5.2; Banerjee et al., 1973) Glu35 is expected to be protonated while Asp52 is ionized. The negative charge of Asp52, a hydrophobic environment promoted largely by Trp108, Val109 and Ala110 and the exposure of the side chain of Glu35 to the negative dipole (C-terminus) of a helix formed by residues Leu25-Ser36 are each considered to contribute to the abnormally high pK_a of Glu35 (Spassov et al., 1989; Strynadka & James, 1991; Inoue et al., 1992). In the case of human lysozyme, pK_a values of 6.8 and 3.8 for Glu35 and Asp53 respectively have been reported (Kuramitsu et al., 1974). This abnormally high basicity for a glutamic acid residue would favour its protonation thereby heightening its function as a general acid. Finally, structural comparisons have indicated that the glutamic acid residues at position 35 in HEWL, 11 in T4L, 73 in GEWL and 478 in Slt70 are structurally superimposable in the active sites of the respective enzymes (Matthews et al., 1981; Weaver et al., 1995; Thunnissen et al., 1995b).

An essential element of the proposed Phillips mechanism is the preferred binding of a distorted D site saccharide. This hypothesis has historically been regarded as a weak point of the Phillips mechanism. There is, however, general agreement that such distortion will weaken the glycosidic bond and favour the formation of an oxocarbonium ion (Post & Karplus, 1986; Kirby, 1987; Frank, 1992). On the basis of stereoelectronic arguments, the glycosidic bond can be broken only if the pyranose ring adopts a conformation such that the glycosidic bond is antiperiplanar to an sp^3 lone pair on the ring oxygen. For a pyranose ring in a full-chair conformation, the lone pair electrons of the ring oxygen (O_5) are orthogonal to the C_1-$O_{4'}$ bond and bond breakage is not geometrically favourable (Scheme 1.1 A). However, distortion of the ring into a half-chair conformation approximates the correct antiperiplanar geometry of the lone pair electrons of the ring oxygen and the C_1-$O_{4'}$ bond (Scheme 1.1 B; the pyranose ring has been depicted as a twist boat for reasons of clarity. The conformation of the site D residue observed in crystal structures is actually one in which C_1, C_2, C_5 and O_5 are approximately



Scheme 1.1.

coplanar; refer to section 1.10). Therefore, oxocarbonium ion formation is facilitated by distortion of the pyranose ring.

The potential energy barrier for ring distortion has been estimated to be approximately 6-12 kcal/mol (Imoto et al., 1972). Phillips envisioned that a role of HEWL is to induce the ring distortion during ground state binding of the substrate in addition to providing the required energy through stabilizing interactions. If indeed the case, then the ground state distorted conformation would approximate the conformation of an oxocarbonium ion intermediate. Since interactions between the enzyme and a distorted site D residue have already been developed in the ground state, then an oxocarbonium ion should also be favourably stabilized by the enzyme. Jencks has stated that part of the intrinsic binding energy used for distortion of the substrate is converted to an acceleration of the rate (Jencks, 1975).

There are several lines of evidence suggesting that enzymatic cleavage of the glycosidic bond by HEWL involves considerable carbonium ion character in the transition state. Secondary isotope effects for deuterium or tritium α-substitution on C_1 of HEWL substrates have indicated a sp^3 to sp^2 transition in the rate-limiting step (Dahlquist et al.,

1969a; Smith et al., 1973). Oxygen-18 leaving group kinetic isotope effects have predicted that C_1-O_4 bond breakage is virtually complete in the transition state for the first irreversible step and that there is little proton transfer (i.e. from Glu35) to the leaving group during the transition state (Rosenberg & Kirsch, 1981). A proposed transition state for the HEWL catalysed reaction is presented in Figure 1.9.

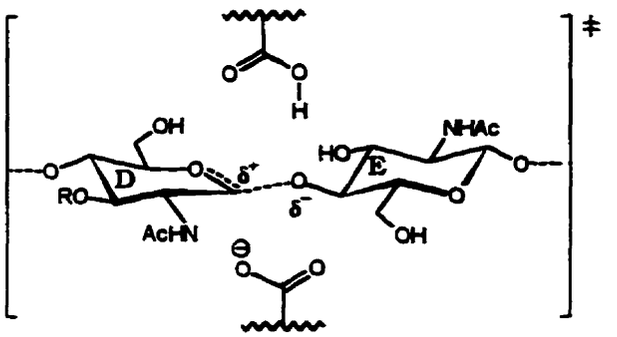


Figure 1.9. Possible transition state for the HEWL catalysed reaction proposed by Rosenberg & Kirsch (1981).

A reaction that proceeds through an oxocarbonium ion intermediate could, in principle, result in either retention or inversion of configuration as the attacking water molecule could approach from either side of the plane. Phillips had predicted that due to the geometry of the active site, the oxocarbonium ion intermediate could only be approached by an acceptor molecule from the β-direction. This prediction was confirmed from methanolyis studies that have indicated that the reaction catalysed by both human lysozyme and HEWL proceeds with retention of configuration to an extent of 99.9 and 99.7% respectively (Dahlquist et al., 1969b). In addition to the hydrogen bond between the anomeric hydroxyl of the site D MurNAc and Glu35 observed in the HEWL/MGM complex described earlier, the complex indicated that there is insufficient room for an α-anomeric oxygen atom on the D site MurNAc due to the proximity of Asp52 (Strynadka & James, 1991). However, in the HEWL/$(GlcNAc)_4$ complex in which Asp52 has been mutated to Ser, an α-configuration was observed for the D site GlcNAc (Hadfield et al., 1994). In the D52S mutant, mutarotation of the ligand and a vacant pocket created by the smaller serine residue are thought to account for the favoured binding of the α-anomer. Efforts to determine the stereospecifity of reactions catalysed by lysozymes belonging to the *g*, *ch* and until recently, *p* type families have been hampered for the reason that the kind of substrates required for these determinations are not readily available. As will be discussed in the following section, T4L hydrolyses its substrate to yield the α-anomer.

1.5.2. Alternatives to the Phillips Mechanism

Over the years, the mechanism of lysozyme action, which is considered a paradigm for other glycosidases has been the subject of extensive experimental and theoretical investigations; nevertheless, some features of the mechanism continue to be questioned (Vernon, 1967; Lowe, 1967; Warshel & Levitt, 1976; Ballardie et al., 1977; Post & Karplus, 1986; Kirby, 1987; Sinnott, 1987, 1990; Frank, 1992; Hardy & Poteete, 1991; Sanz et al., 1992; Weaver et al., 1995; Matsumura & Kirsch, 1996a,b). Some of these points are related to the proposed ground-state distortion of the site D residue and the suggested stabilizing role of the catalytic aspartate. Further discussion of the ring distortion hypothesis with reference to experimental evidence will appear throughout sections 1.7-1.10.

Although the available data from kinetic isotope effects are consistent with the formation of an oxocarbonium ion in the transition state, there has been much debate as to the existence of an oxocarbonium ion intermediate. The lifetime of a nonstabilized glycosyl oxocarbonium ion has been estimated to be on the border line of a molecular vibration or approximately 10^{-15} s whereas such an ion, when solvated, may have a duration as long as 10^{-10} s (Sinnott & Jencks, 1980; Bennet & Sinnot, 1986). In an enzymatic reaction, the oxocarbonium ion intermediate must survive long enough to permit the departure of the aglycon and the diffusion of water (or another glycosyl acceptor in the case of transglycosylation reactions) into the active site to complete the reaction. To fulfil this requirement and based on the estimated 10^{-15} s lifetime for the oxocarbonium ion, it has been suggested that the amount of stabilization provided by the enzyme must be at least 5-7 kcal·mol^{-1} (Young & Jencks, 1977). In the case of HEWL, the stabilization required is theoretically feasible as calculations have predicited that Asp52 would stabilize the oxocarbonium ion by 9 kcal·mol^{-1} and this stabilization has been argued to be the most significant factor contributing to the reaction catalyzed by HEWL (Warshel & Levitt, 1976).

It has been repeatedly suggested however, that the oxocarbonium ion transition state collapses to form a covalent glycosyl-enzyme intermediate. A double displacement mechanism for HEWL is consistent with much of the evidence advocating the Phillips mechanism but does not require a long lived oxocarbonium ion intermediate (Fig. 1.10). This mechanism was initially proposed by Koshland (1953) as a general mechanism for retaining glycosidases and has received supporting evidence from studies on β-glucosidase (Street et al., 1992) and β-galactosidase (Gebler et al., 1992). As is depicted in Fig 1.10, Glu35 serves as a proton donor and an oxocarbonium ion transition state develops that is attacked by Asp52 to form a covalent intermediate. In the second

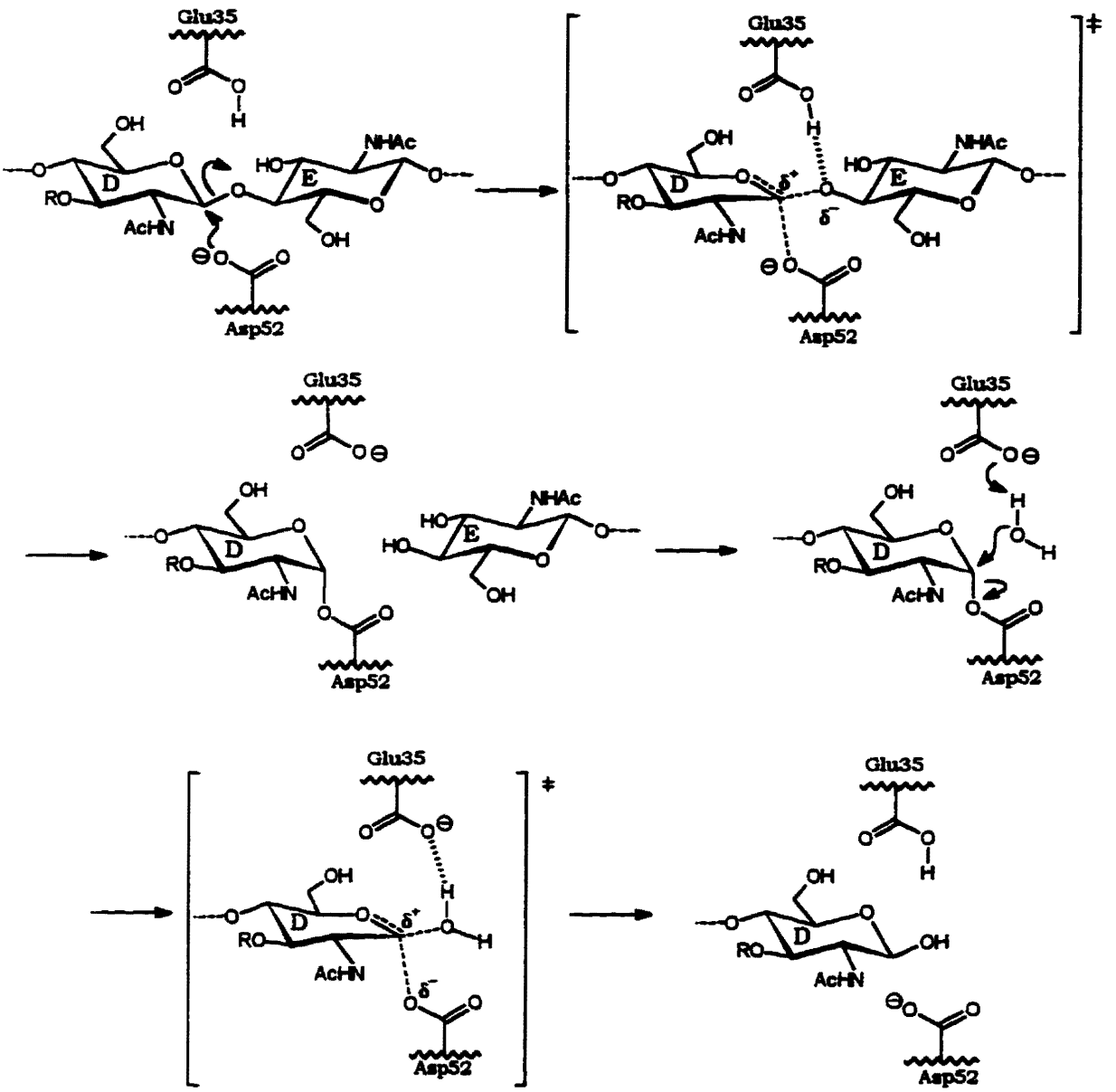


Figure 1.10. Proposed double displacement mechanism for HEWL.

displacement step, the carboxylate of Glu35 promotes the attack of water onto C_1 and Asp52 is the leaving group. As with the Phillips mechanism, the reaction is completed with retention of the original β-configuration.

Although all the essential features of the Phillips mechanism are applicable to the double-displacement mechanism, a glycosyl-enzyme intermediate with native HEWL has yet to be observed. In the case of the reaction of $(GlcNAc)_6$ with the D52S mutant HEWL, electrospray mass spectral analysis has indicated the possible formation of a minor

species corresponding to an adduct between the mutant HEWL and $(GlcNAc)_4$ (Lumb et al., 1992). Although noncovalent complexes of native HEWL and $(GlcNAc)_n$ have been observed previously by mass spectrometry (Ganem et al., 1991), the adduct observed with the D52S mutant was thought to arise from hemiacetal formation between Ser52 and an acyclic aldehyde intermediate of the site D residue that underwent mutarotation (Lumb et al., 1992). The possibility of forming a covalent intermediate with the native enzyme has been explored from inspection of the HEWL/MGM complex (Strynadka & James, 1991). These efforts led to a conclusion that at least in the reported complex, the position and geometry of $O^{\delta 2}$ of Asp52 and its lone pair orbitals are not favourably oriented to form a covalent adduct with the C_1 of the site D residue. Furthermore, any movement of Asp52 would also require disruption of hydrogen-bonding interactions with Asn46 and Asn59 (Strynadka & James, 1991). The hydrogen bond between Asp52 and Asn46 has recently been suggested essential for proper recognition of the site D saccharide (Matsumura & Kirsch, 1996b).

In the case of GEWL and Slt70, a double displacement mechansim is highly unlikely as these enzymes have been suggested to possess no amino acid counterpart in an equivalent position to Asp52 in HEWL that could form a covalent intermediate (Weaver et al., 1995; Thunnissen et al., 1995a). Although acidic residues have been located in the active site of GEWL and Slt70 that are in the vicinity of the position occupied by Asp52 in the active site of HEWL (see examples in Table 1.3), they are too far removed to form an ion-pair with an oxocarbonium ion intermediate. For example, Asp86 of GEWL is 10 Å from a superimposed position to Asp52 in HEWL while Asp97 of GEWL would require a conformational change to approach a position that is still somewhat removed to that of Asp52 (Weaver et al., 1995). However, both GEWL and Slt70 have equivalent residues to Glu35 of HEWL (Glu73 in GEWL; Glu478 in Slt70) and therefore, only a single acidic residue appears necessary for the action of these enzymes.

Stabilization of an oxocarbonium ion intermediate is still possible even without the involvement of an acidic enzyme residue. There is support for intramolecular acetamido group participation from studies on the nonenzymatic hydrolysis of glycosides (Piszkiewicz & Bruice, 1968a,b,c). In the general case of the lysozyme mechanism, an appropriately positioned hydrogen bond acceptor could induce a charge distribution such that the acetamido carbonyl acquires a partial negative charge which could stabilize an oxocarbonium ion intermediate (Fig. 1.11 A). This type of arrangement has been suggested for Slt70 based on studies of a GlcNAc-MurNAc disaccharide modeled into the

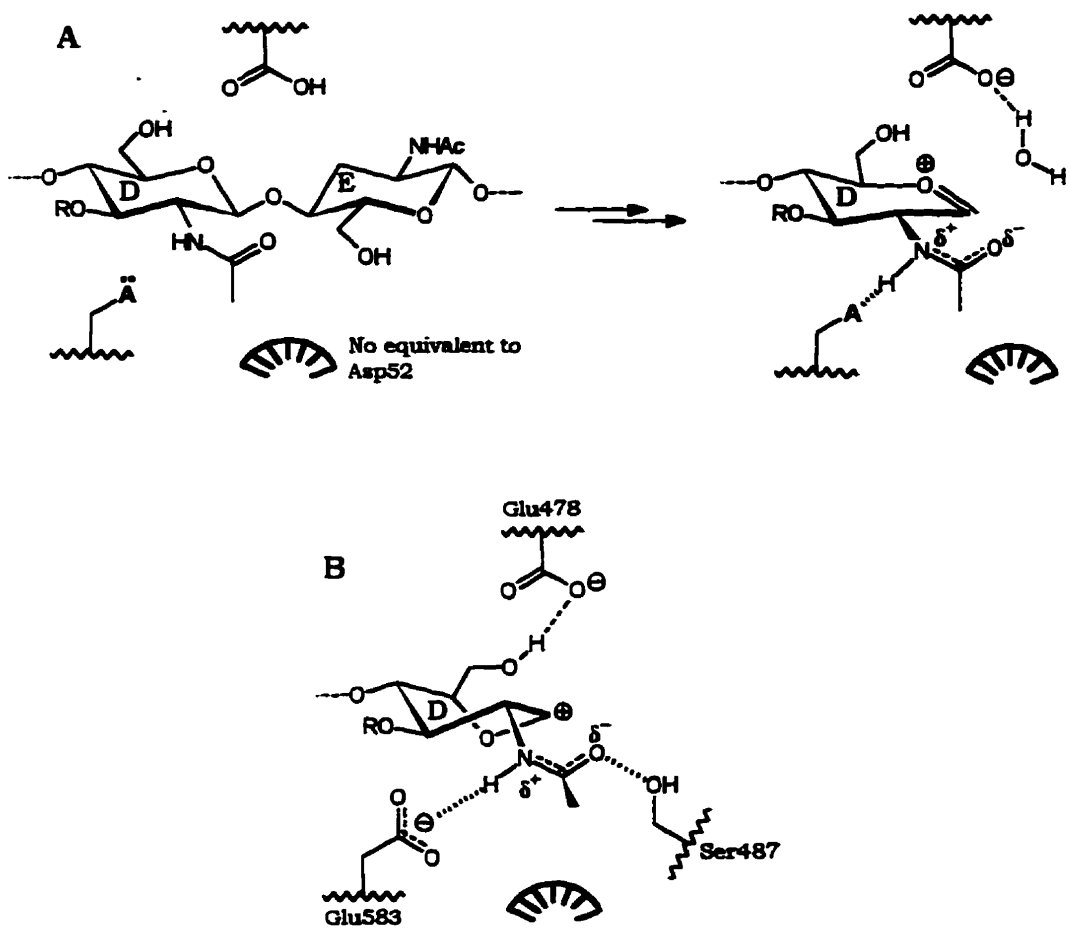


Figure 1.11. Participation of the 2-acetamido carbonyl oxygen in oxocarbonium ion stabilization.

(A) General scheme for lysozymes that lack an equivalent to Asp52 in HEWL. A hydrogen bond interaction between the acetamido amide and an enzyme acceptor **A** could aid in the promotion of partial negative charge on the acetamido carbonyl oxygen.

(B) Possible charge distribution of the acetamido group promoted in the active site of Slt70. Adapted from Thunnissen et al. (1995a).

active site (Thunnissen et al., 1995a). The model indicated that Glu583 and Ser487 could promote a partial negative charge and a conformation of the carbonyl oxygen of the acetamido group that could approach the C_1 atom of MurNAc (Fig. 1.11 B). Such an intramolecular involvement of the acetamido group in addition to the possible formation of a bond between the carbonyl oxygen and C_1 has been suggested as an alternative to the Phillips mechanism for HEWL (Lowe, 1967; Vernon, 1967). Furthermore, electrodynamic calculations on HEWL, HuL and T4L have predicited that in each, a large electrostatic field

exists across the active site cleft that was suggested to play an important role with respect to electrostatic interactions between the enzymes and charged intermediates during catalysis (Dao-Pin et al., 1989).

Until recently, the mechanism of action for T4L was thought to closely mimic that of HEWL since both Glu11 and Asp20 in T4L are found in essentially analogous positions to Glu35 and Asp52 of HEWL when the active sites of these enzymes are structurally superimposed (Matthews et al., 1981b). Glu11 of T4L is still considered to assume the role of general acid but in light of more recent evidence, Asp20 may not merely assume the role of stabilizing counter ion. Mutation of Asp20 to Cys generated an enzyme with 80% activity (at pH 7) of the wild type T4L (Hardy & Poteete, 1991). Studies on the pH dependency of the reaction catalysed by the D20C T4L suggested that the cysteine thiol was protonated and not in the form of a thiolate that would be required for electrostatic stabilization of an oxocarbonium ion. However, the near wild type activity of the mutant was taken as an indication that both Asp20 and Cys20 in the wild type and mutant lysozymes respectively assumed a nucleophilic role in formation of a covalent intermediate in a double-diplacement mechanism (Hardy & Poteete, 1991).

The double diplacement mechanistic proposal for T4L is expected to result in retention of configuration (β-anomer) in the products. However, unlike HEWL, the stereochemistry of the T4L reaction has only very recently been established and was shown to produce the α-anomer in the hydrolysed products of a peptidoglycan substrate (Kuroki et al., 1995). This observation has led to the proposal of a single displacement mechanism for T4L (Fig. 1.12). Glu11 is presumed to donate a proton to the glycosidic oxygen. A tightly bound water molecule (H_2O^{211}) hydrogen bonded by Asp20 and Thr26 in the native structure has a lone pair directed toward C_1 of the site D residue (Kuroki et al., 1993). Acting as a general base, Asp20 promotes attack of H_2O^{211} onto C_1 in a single diplacement reaction inverting the configuration to give the α-anomer as product. Following catalysis, the proton is presumably restored from Asp20 to Glu11 through solvent transfer.

Mutation of Asp20 to Glu, Ala and Cys each produced catalytically active enzymes that hydrolysed the same peptidoglycan substrate used for the wild type T4 lysozyme and in which the α-anomer was also observed in the product (Kuroki et al., 1995). The Ala mutant clearly indicates that a negatively charged residue at position 20 is not essential for the reaction catalysed by T4L. However the observed rates (k_{obs}, determined at pH 5) for the mutants as compared to the wild type T4L (0.03% for E20; 0.005% for A20; and

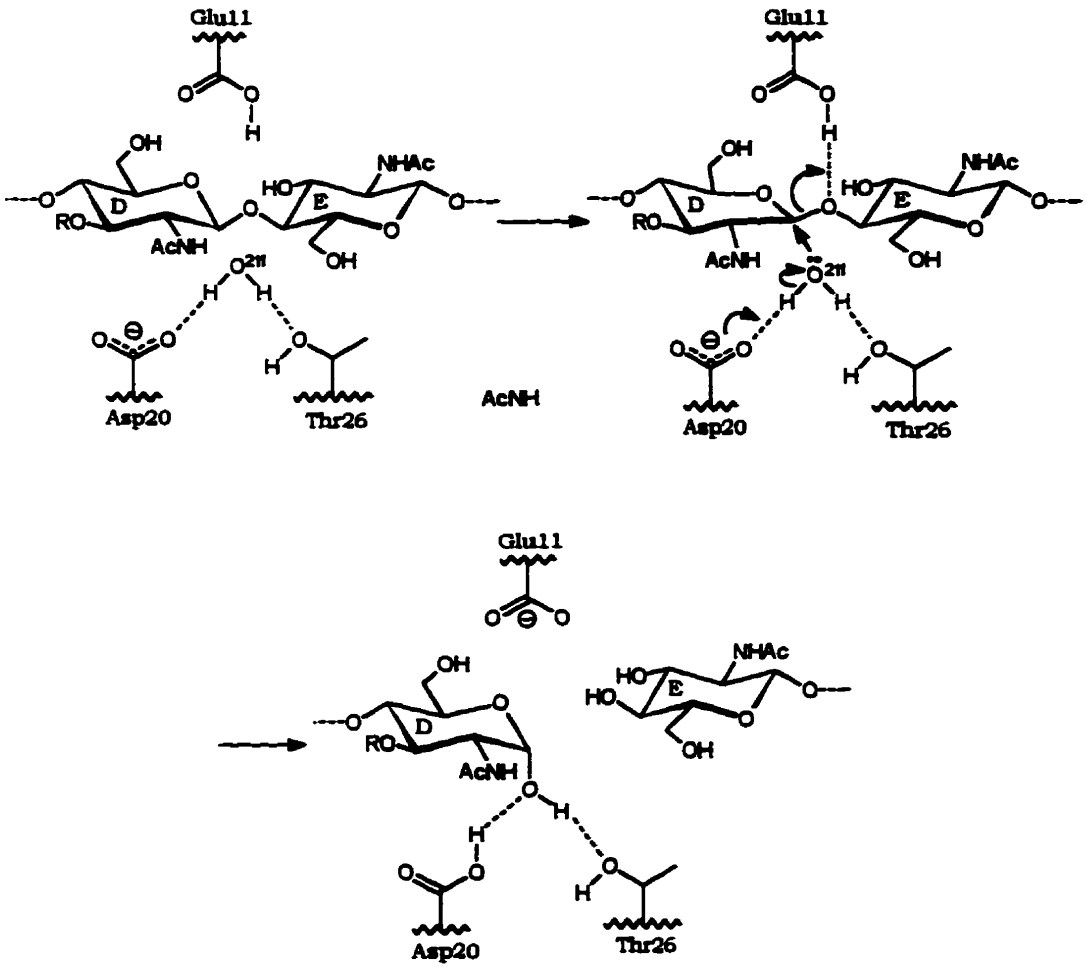


Figure 1.12. Single displacement mechanism proposed for T4L. Adapted from Kuroki et al. (1993, 1995).

0.05% for C20; Kuroki et al., 1995) nonetheless suggest an important role of Asp20 in enhancing the rate of reaction catalysed by the wild type T4L.

Mutation of Thr26 of T4L has also produced interesting results. Replacement of Thr26 with Glu resulted in an enzyme that cleaved the peptidoglycan substrate, yet a covalent adduct between $O^{\varepsilon 1}$ of Glu26 and C_1 of the site D MurNAc residue was generated (Kuroki et al., 1993). The structure of the covalent T26E T4L/peptidoglycan indicated that the site D MurNAc residue was distorted suggesting that ring distortion may play a similar role for T4L as was proposed for HEWL. The water molecule (H_2O^{211}) in the native structure and $O^{\varepsilon 1}$ of Glu26 in the covalent adduct structure were found to be almost exactly superimposable (Kuroki et al., 1993). The mutation of Thr26 to His generated an enzyme that cleaved the peptidoglycan substrate (k_{obs} = 0.17% of the wild type) but with

production of the β-anomer in the product (Kuroki et al., 1995). As with the T26E mutant, the T26H mutant is thought to form a covalent intermediate during a double displacement mechanism. As pointed out by Kirby (1995), the T26H mutant T4L is one, if not the first, example in which genetic engineering has been used to alter the catalytic mechanism of an enzyme.

Each of the mechanisms that have been described thusfar are related in that the general acid protonation of the exocyclic oxygen (i.e. $O_{4'}$) is succeeded with the formation of a cyclic oxocarbonium intermediate or transition state. This has been termed an *exocyclic* cleavage mechanism. Fleet has proposed an alternative general mechanism for glycosidases involving an *endocyclic* cleavage mechanism (Fleet, 1985). In a controversial paper, Post and Karplus have suggested based on their molecular dynamics simulations of HEWL (Post et al., 1986) that the lysozyme mechanism might follow a new pathway involving endocyclic cleavage (Post & Karplus, 1986). In the proposed mechanism (Fig 1.13 A) protonation of the ring oxygen (O_5) rather than the exocyclic oxygen ($O_{4'}$) occurs. The simulation predicted a position for Glu35 that led to a hydrogen bond with the ring oxygen and not the exocyclic oxygen in addition to the favourable binding of an undistorted, full-chair conformation for the site D residue. The full-chair conformation permits stereoelectronic assistance from the lone pair of $O_{4'}$ antiperiplanar to the C_1-O_5 bond in the formation of the acyclic oxocarbonium ion which is stabilized in a similar manner by the carboxylate of Asp52 as in the Phillips model. In a somewhat more elaborate mechanism, the endocyclic pathway has also been used to reconcile the formation of a ring-closed glycosyl-enzyme intermediate (Frank, 1992). Although it has been argued that the endocyclic pathway is consistent with existing experimental evidence on lysozyme hydrolysis (Post & Karplus, 1986; Frank, 1992) the endocyclic pathway has been discredited by others (Kirby, 1987; Sinnott, 1990) in its inapplicability to those glucosidases known to hydrolyse glycosyl fluorides and glycosyl pyridinium salts (Fig. 1.13 B: for specific examples see Legler et al., 1980; Sinnott, 1987; Matsui et al., 1993). In a recent report Sinnott has presented convincing evidence that strongly opposes endocyclic cleavage in the enzyme catalysed pathway (Sinnott, 1993). However in purely chemical models, definitive support for the endocyclic hydrolysis of pyranosides has appeared (Liras & Anslyn, 1994).

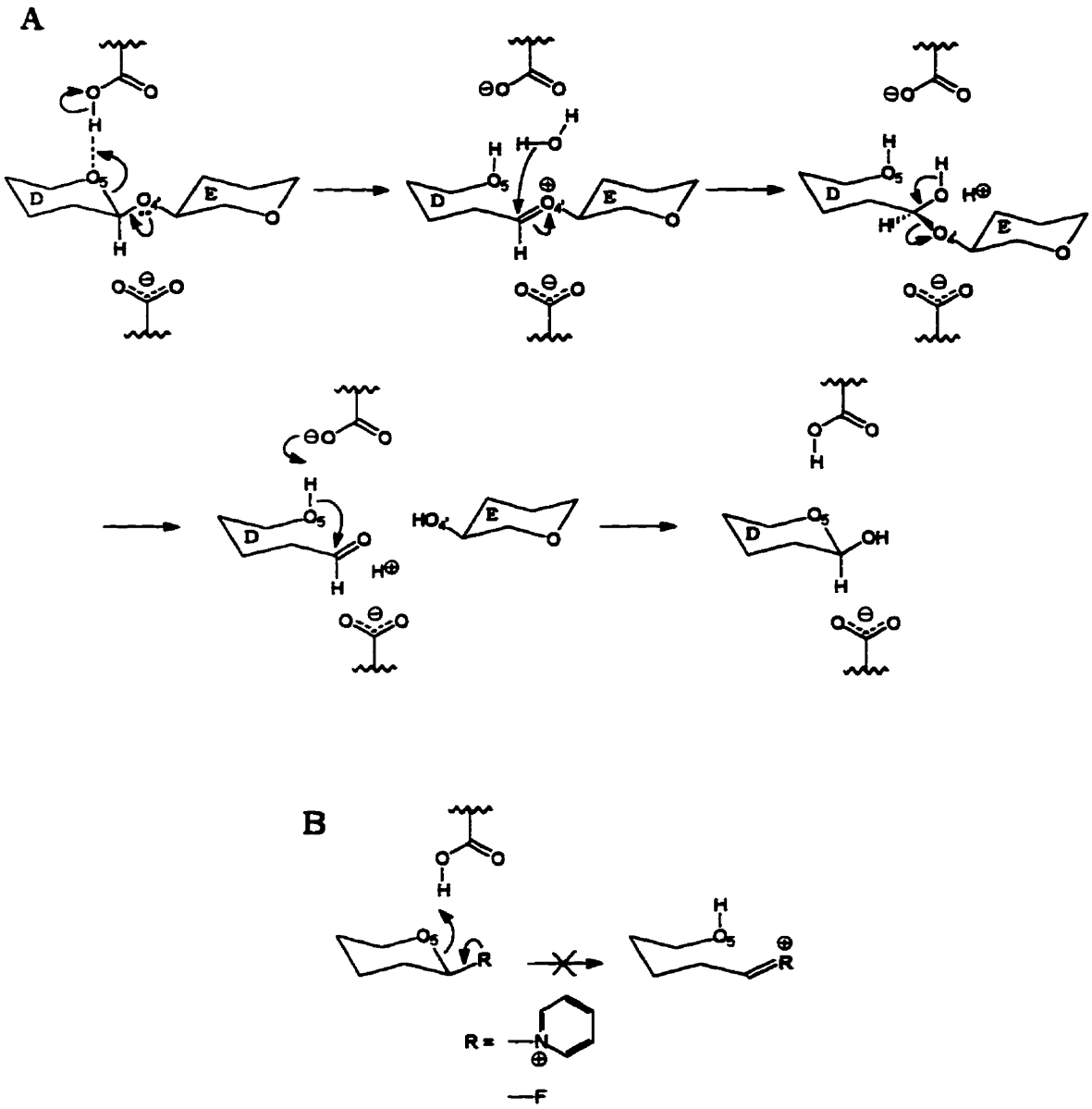


Figure 1.13. (A) Proposed endocyclic cleavage mechanism for HEWL. Adapted from Post & Karplus (1986). (B) Glycosides that are not compatible with endocyclic cleavage.

1.5.3. Mechanism for Lambda Lysozyme and Slt70

It has become apparent that enzymes which cleave peptidoglycan have elements in common with respect to their mechanism of action. Lambda lysozyme is also expected to share in these common mechanistic features. An even greater analogy in mechanism for LaL is expected with the *E. coli* exomuramidase Slt70 as both enzymes have been shown to cleave peptidoglycan resulting in 1,6-anhydromuramic acid production. A possible reaction mechanism for LaL and Slt70, which has received recent support from crystallographic studies on Slt70 (Thunnissen et al., 1994, 1995a), is presented in Figure 1.14.

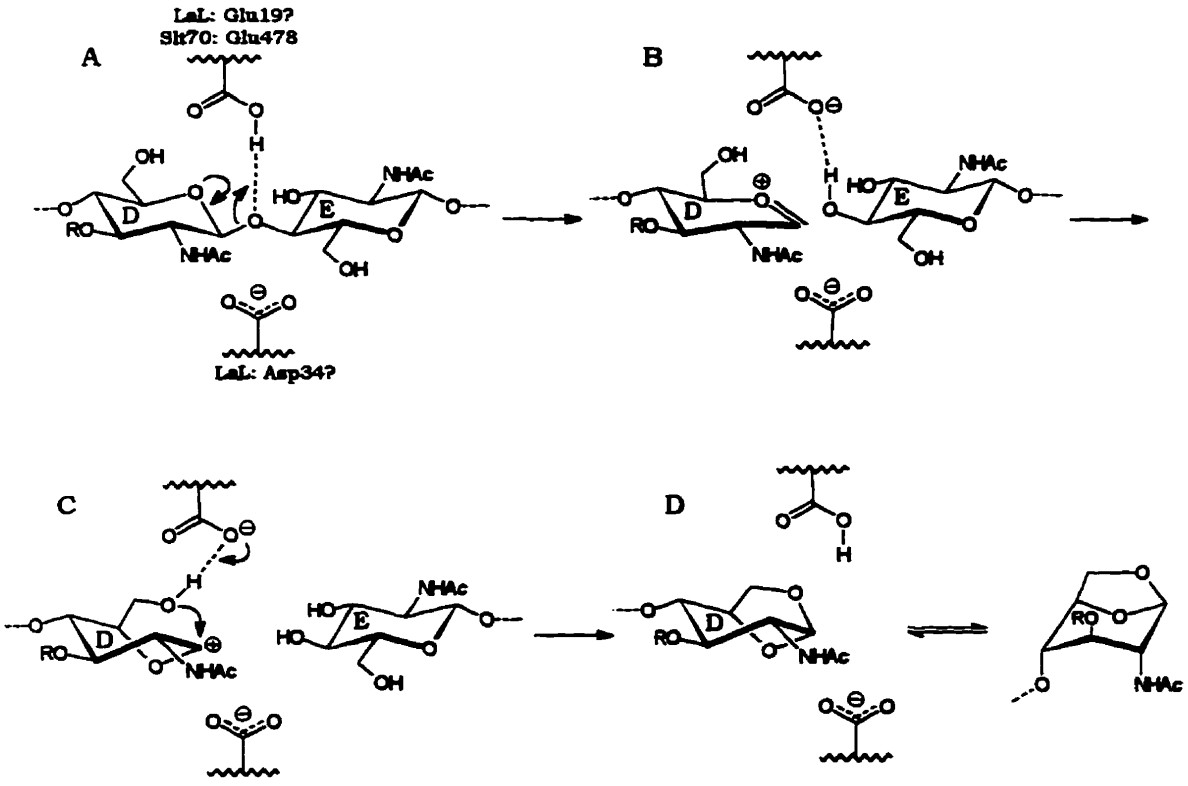


Figure 1.14. Possible mechanism for lambda lysozyme and Slt70. Adapted from Thunnissen et al. (1995a).

The first two steps in the intermediates along the pathway (A and B, Fig. 1.14) are expected to be analogous to the Phillips mechanism involving protonation of the glycosidic oxygen and bond rearrangment to an oxocarbonium ion intermediate. It was previously stated that Glu478 has been implicated as the general acid in Slt70 while this role has been suggested for Glu19 in LaL from sequence comparisons to HEWL and T4L and by

mutation to Gln (Jespers et al., 1992). Subsequent reactions along the pathway must be different than predicted by the Phillips mechanism. Instead of an attack by water, the C6-hydroxyl functionality of the MurNAc is the nucleophile participating in an intramolecular reaction with C_1 of the oxocarbonium ion. This reaction (C, Fig. 1.14) must involve even greater contortions of the pyranose ring in order for the C6-hydroxyl group to optimally approach C_1 as a nucleophile (this can be easily be appreciated with the use of a simple molecular modelling set). It has been argued that such an intramolecular attack would probably require a shorter lifetime of the oxocarbonium ion intermediate as it would no longer be dependent on the diffusion-controlled replacement of the aglycone by water (Thunnissen et al., 1994). For this reason, there may be less of a demand for the enzymes to stabilize an oxocarbonium ion intermediate and could explain the lack of an equivalent to Asp52 of HEWL in Slt70. In light of the newly ascribed role for Asp20 in T4L, it has become questionable as to whether LaL has an equivalent residue to Asp52 in HEWL with a related function. However, Asp34 in LaL has received some attention as a possible candidate from the results of mutations at this position (Jespers et al., 1992; refer to Table 1.3). The determination of the three dimensional structure of LaL will be quintessential in providing greater insight as to the residues critical to the mechanism of this enzyme.

1.6. Structural Homologies between Lysozymes

A brief account on the exisiting sequence similarities between lysozymes belonging to the same family has been presented previously in section 1.1. Within a given family there are clear sequence homologies; however, there are no obvious sequence similarities between one family and another (i.e. inter-family sequence homologies between the *c*, *g* and *p* type lysozymes)(Jollès & Jollès, 1984; Grütter et al., 1985). It is beyond the scope of this discussion to detail the inter-family homologies and incongruities that have been observed to exist both in sequence and in structure. However, some generalizations can be made with reference to these types of detailed studies reported elsewhere[1.1]. The homology of regions which have been considered as potentially conserved are weak and very little (if any) inter-family sequence homologies may exist. More important are the existence of conserved structural elements and motifs.

[1.1] For comparisons between HEWL and T4L, see Rossmann & Argos (1976), Remington & Matthews (1978), Artymiuk et al. (1981), Matthews et al. (1981a,b). For comparisons between GEWL, HEWL and T4L, see Grütter et al., 1983; Weaver et al. (1985a, 1995). For a discussion on the comparison of the lysozyme from *S. erythraeus* to HEWL and T4L, see Harada et al. (1981).

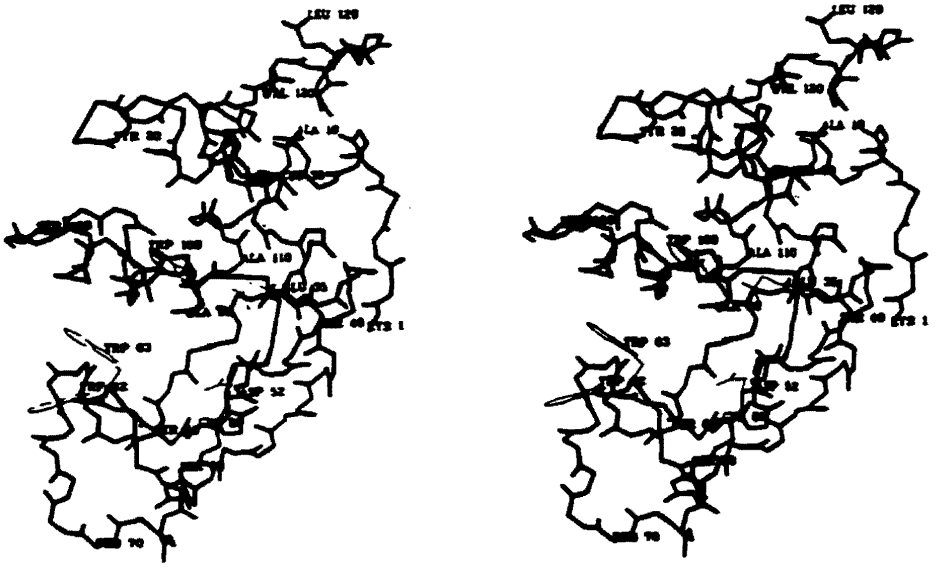
It would appear that the *c* type lysozymes are structurally homologous with each other. The sequence of partridge egg white lysozyme differs from HEWL in only three positions and the overall structure of the two *c* type lysozymes are very similar with an average RMS (root-mean-square) main-chain (including carbonyl oxygens) deviations of 0.29 Å (Turner & Howell, 1995). Although there are several replacements of amino acid residues between HEWL and human lysozyme (53 residues out of 130), there is only a 0.74 Å RMS difference in the α-carbon positions, the active-site geometries between the enzymes are closely similar and the residues essential for catalysis and substrate recognition are highly conserved (Artymiuk & Blake, 1981). It is unfortunate that there are fewer representative examples of structures from the *g* and *p* type lysozymes so that structural homologies within their respective families might be investigated.

The structure of HEWL (129 aa) is presented in Figure 1.15 A. The protein is bilobal with overall dimensions of approximately 45 × 30 × 30 Å (Imoto et al., 1972). HEWL is considered to consist of four structural fragments that include residues: **1**, 1-39; **2**, 40-85; **3**, 86-100 and **4**, 101-129 (Jung et al., 1980). The active site is contained within a cleft that divides HEWL into its two lobes. On the bottom of the cleft (the bottom portion of Fig. 1.15 A) is situated fragment **2**, a region primarily β-sheet in structure. The top lobe contains the N-terminal (**1**) and C-terminal (**4**) fragments which are essentially α-helical in nature. The two lobes are joined by the 15 residue helix formed by fragment **3**. The putative general acid (Glu35) is located within the helix at the end of fragment **1**. This helix extends directly into the β-sheet structure of **2** in which is located the catalytic Asp52.

The T4L molecule (164 aa) has overall dimensions of approximately 50 × 30 × 30 Å and it is clear from Figure 1.15 B that in common to HEWL, T4L is also bilobal and contains a deep crevice in which is located the active site (Matthews & Remington, 1974). Approximately 60% of T4L is comprised of α-helices with the prominent α-helix formed by residues 60-79 joining the upper and lower lobes. The only extended β-structure is formed from residues 18-34 in the lower lobe and reminiscent to HEWL, Asp20 is found within this β-structure fragment while Glu11 is located in the preceding α-helix. Unlike the HEWL fold, the upper and lower lobes of T4L consist exclusively of the amino and carboxy terminal portions of the protein.

The X-ray structure of GEWL (185 aa, Fig. 1.15 C) indicated overall dimensions of approximately 58 × 40 × 40 Å for this lysozyme (Grütter et al., 1979). The proposed

(A) HEWL



(B) T4L

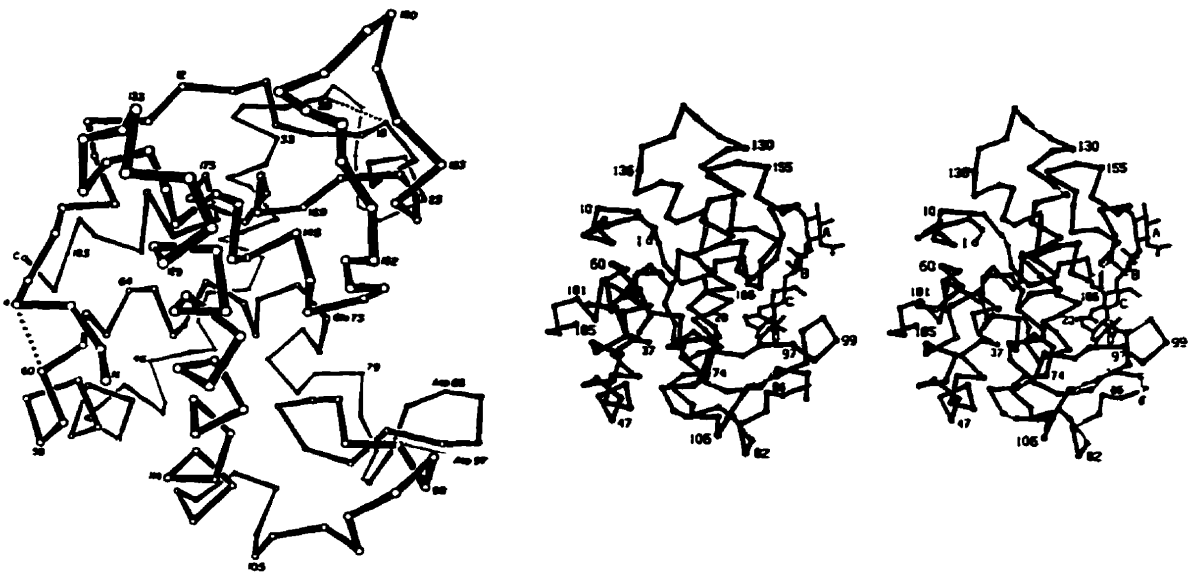


Figure 1.15. X-ray structures of lysozymes.

(A) HEWL. Reproduced from Strynadka & James (1991).

(B) T4L. Reproduced from Matthews & Remington (1974).

(C) GEWL



(D) *S. erythraeus* lysozyme

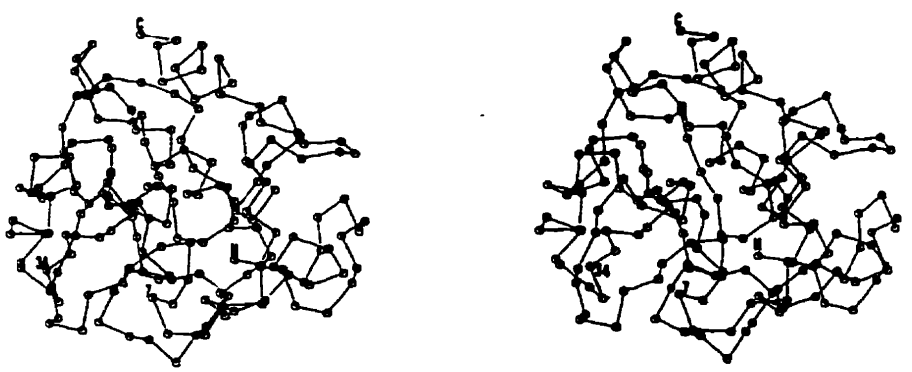


Figure 1.15. X-ray structures of lysozymes.

(C) GEWL. Reproduced from Weaver et al. (1985).

(D) *S. erythraeus* lysozyme. Reproduced from Harada et al. (1981).

general acid (Glu73) is found in a helix, and the only β-structured region of this mostly helical protein is essentially made from residues 83-98. The two acidic residues, Asp86 and Asp97, each at one time considered as a possible equivalent to Asp52 in HEWL, are found in this β-region. Like HEWL and T4L, proteins in which Asp52 and Asp20 respectively are found in β-structures that are preceded by the helix containing the proposed general acid, the β-structure of GEWL is preceded by the helix containing Glu73. The occurrence in GEWL of a lower domain (containing the β-structure) and a predominantly helical upper domain joined by a long α-helix (residues 111-131 in GEWL) is familiar to both HEWL and T4L (Grütter et al., 1985). This structural organization appears to define a possible lysozyme fold.

Relative to HEWL and T4L, the "waist of the GEWL molecule is thickened" by the amino terminal residues that extend across the upper and lower domains (Grütter et al., 1985). With the exception of the respective helices in HEWL, T4L and GEWL which join the two domains in these proteins, only in the case of GEWL is there a region of the protein that extends into each of the two domains.

The *ch* type lysozyme from *Streptomyces erythraeus* (185 aa) has overall dimensions of approximately 45 × 30 × 30 Å (Harada et al., 1981). The structure of the bacterial lysozyme (Fig. 1.15 D) indicated the occurrence of three domains made from residues 1-49, 50-110 and the remaining C-terminal domain. Only approximately 18% of the bacterial lysozyme is in a helical form, a much smaller percentage than that observed for the *c*, *g* and *p* type lysozymes. In addition, the bacterial lysozyme appears to show no homology in its sequence nor in its overall three-dimensional structure with the other lysozymes (Harada et al., 1981). The proposed catalytic residues of the *Streptomyces* enzyme, Asp7 and Glu34, appear in a region of this protein that has been suggested to be similar to the region containing Glu11 and Asp20 of T4L (Harada et al., 1981). The apparent scarcity of structural homology between the bacterial lysozyme and the lysozymes from hen, goose and T4 is suggestive of a unique type of lysozyme fold for the bacterial enzyme and possibly, for other *ch* type lysozymes. Little information is available comparing the structure of the *S. erthryaeus* lysozyme with those of other lysozymes and therefore, no future reference to this lysozyme will be made in the following discussion.

The essence of structural relations bewteen HEWL, GEWL and T4L is shown schematically in Fig 1.16. The symbols illustrated in Fig. 1.16 do not represent distinct structural domains. Instead, the squares represent those structural elements that are common to all three lysozymes, the triangles depict shared features unique to HEWL and



Figure 1.16. Schematic of the structural relationship between HEWL, GEWL and T4L. See text for details. Adapted from Grütter et al. (1983).

GEWL, the circles are representative of structural elements that occur only in GEWL and T4L while the star represents the N-terminal polypeptide segment that is present only in GEWL (Grütter et al., 1983). A stronger appreciation of the overall likeness of these enzymes is obtained from inspection of their superimposed structures (Fig. 1.17). Of the 129 residues of HEWL and 164 in T4L, 78 α-carbons were found to be equivalent within a RMS distance of 4.1 Å (Rossmann & Argos, 1976). The structural alignment of GEWL and T4L indicated 92 equivalent α-carbons with a RMS discrepancy of 3.5 Å, 52 α-carbons with a RMS difference of 2.7 Å and 40 α-carbons that superimoposed within a 2.4 Å RMS difference (Grütter et al., 1985). Of the more striking differences visible in the respective architectures of the proteins (Fig. 1.17) are the larger upper lobes found in GEWL and T4L (formed by the respective C-terminal domains) that have no counterpart in HEWL. HEWL has been shown to exhibit a broad substrate specificity differing from those of GEWL and T4L. Some evidence (presented in Chapter 3) has suggested that the C-terminal lobes of GEWL and T4L are involved in recognition of the peptide-portion of peptidoglycan substrates and that these interactions are necessary for the function of these enzymes.

Oligosaccharide complexes observed from crystallographic investigations on HEWL (Blake et al., 1967; Strynadka & James, 1991), T4L (Anderson et al., 1981; Kuroki et al., 1993) and GEWL (Weaver et al., 1995) have indicated that in addition to the shared elements relating the backbone structures of these proteins, there are common features relating these proteins with respect to their interactions with bound saccharides. Oligosaccharides bind to the cleft that is defined by the upper and lower domains respective of each protein. As in the case of the Phillips model for HEWL, the active site clefts of T4L and GEWL are considered to contain discrete monosaccharide binding sites. The organization of the amino acids lining the cleft dictates the alignment of the substrate within the cleft such that cleavage results between MurNAc and GlcNAc located in sites D and E respectively. The binding of saccharides in site A has been less well defined for the complexes with T4L and GEWL than for HEWL.

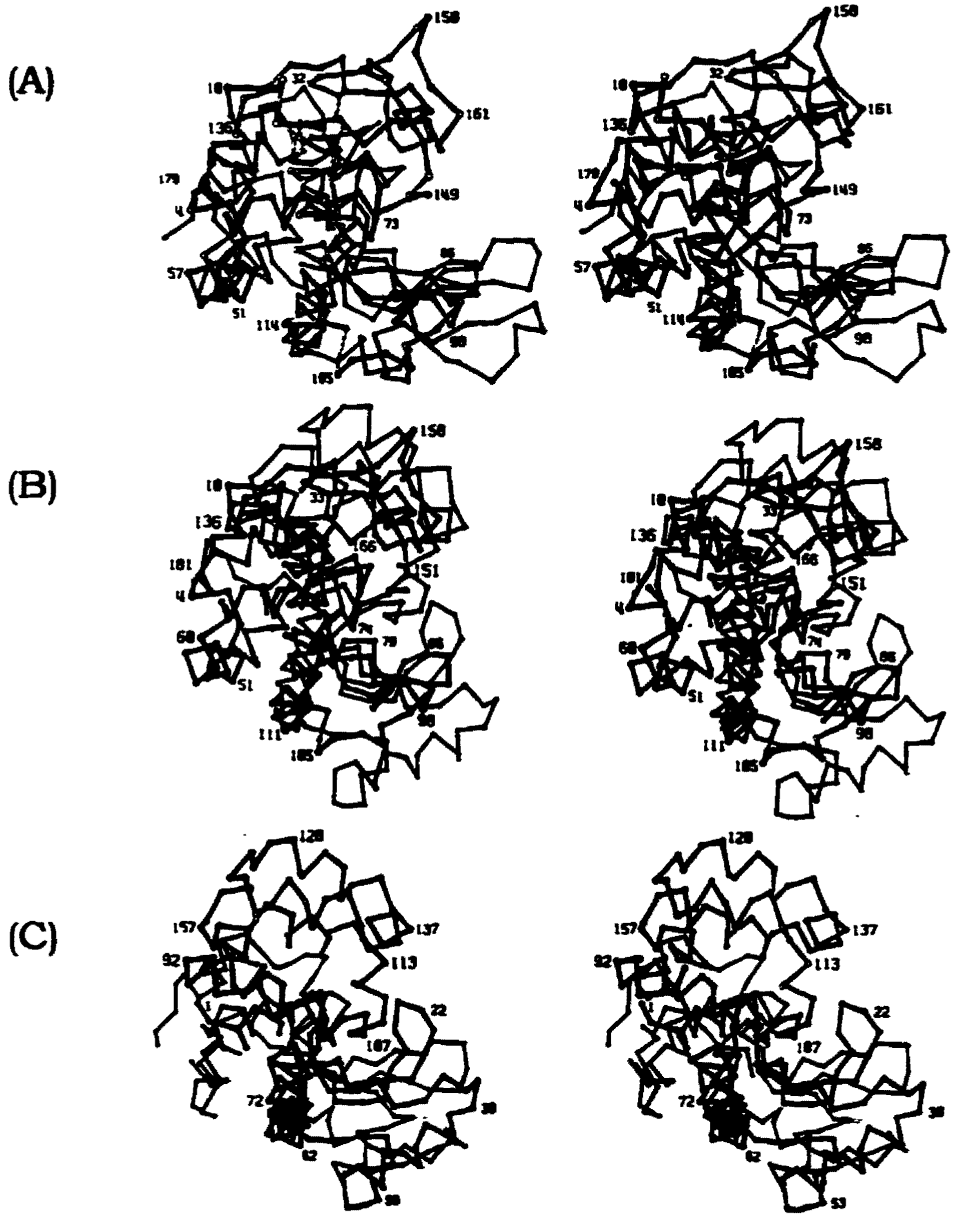


Figure 1.17. Superposition of the structures of HEWL, GEWL, and T4L. Reproduced from Grütter et al. (1983).

(A) GEWL (open bonds) and HEWL (solid bonds). Numbering for GEWL.
(B) GEWL (open bonds) and T4L (solid bonds). Numbering for GEWL.
(C) T4L (open bonds) and HEWL (solid bonds). Numbering for T4L.

The single most distinctive interaction that aligns a bound oligosaccharide with the cleft occurs in subsite C. The complex of GlcNAc with HEWL and T4L indicated exclusive binding of this monosaccharide to site C in the respective proteins (Blake et al., 1967; Anderson et al., 1981). In the two complexes, the most specific interactions resulted from the pair of hydrogen bonds formed between the acetamido group of GlcNAc with a main chain carbonyl and amide of the proteins (Fig. 1.18). The main chain carbonyl (Ala107 of HEWL; Phe104 of T4L) is located in the upper lobe of each protein while the main chain amide (Asn59 of HEWL; Leu32 of T4L) is located in the lower lobe. Therefore, the hydrogen bonding interactions with the site C acetamido group bridges the two lobes on either side of the cleft. An analogous pattern of hydrogen bonding with the site C acetamido group is diplayed by GEWL and involves Tyr147 and Asp97 (Weaver et al., 1995). Other saccharide-protein interactions are presented in Fig. 1.18. Note that in particular, the hydrophobic interaction made between Trp62 of HEWL and the pyranose ring of the site B saccharide (refer to Fig. 1.19 A) has no apparent counterpart in either GEWL or T4L (Weaver et al., 1995).

Throughout this thesis, brief discussions will appear with specific reference to the amino acid residues located in the active site of HEWL. Presented in Fig. 1.19 is the crystallographic data for the binding of the trisaccharide MurNAc-GlcNAc-MurNAc to sites B, C and D of HEWL as reported by Strynadka and James (1991). The figure has been included not only to illustrate the interactions between protein and saccharide, but also to provide the reader with a useful illustration of the specific amino acid or protein-saccharide interactions that will be eluded to in later sections.

1.6.1. Homologies Existing with Lambda Lysozyme

Different strategies have been applied by several groups (see references given in footnote[1.1], p. 36) in the alignment of the structures of HEWL and T4L. These studies have involved either an overall alignment of the proteins or a somewhat more biased approach aligning localized regions having structural relatedness (from crystallographic observations) and the observed alignments vary slightly depending on the method used. Studies comparing the sequence of lambda lysoyzme to that of T4L and other *p* type lysozymes have indicated a very low degree of sequence homology (Rennell & Poteete, 1985; Jespers et al., 1992; Young, 1992). An 18% overall sequence homology between LaL and T4L has been suggested, while a weaker overall homology was found to exist between LaL other c type lysozyme sequences (Jespers et al., 1992).

(A) HEWL

(B) T4L

(C) GEWL

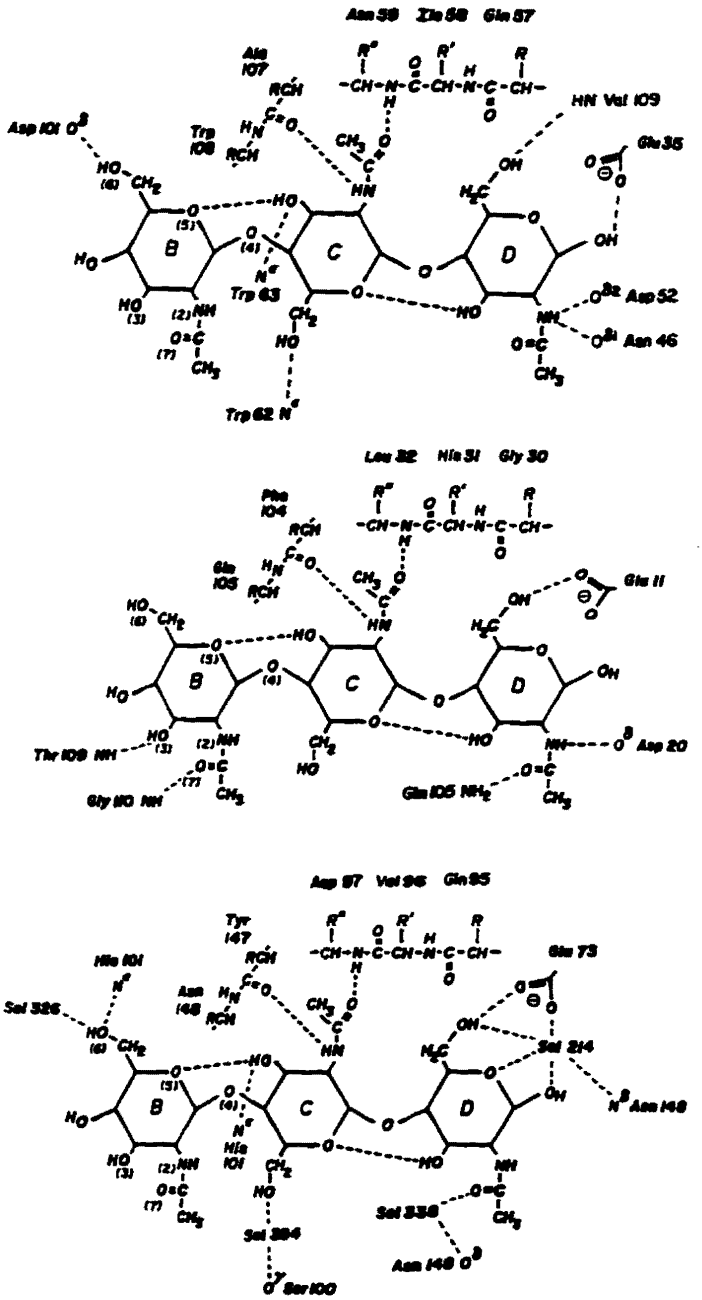
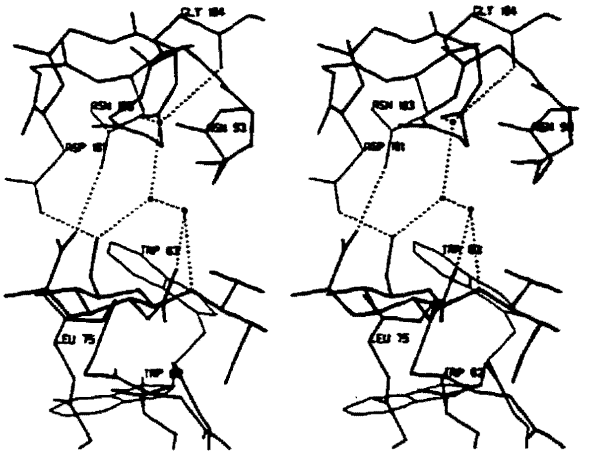
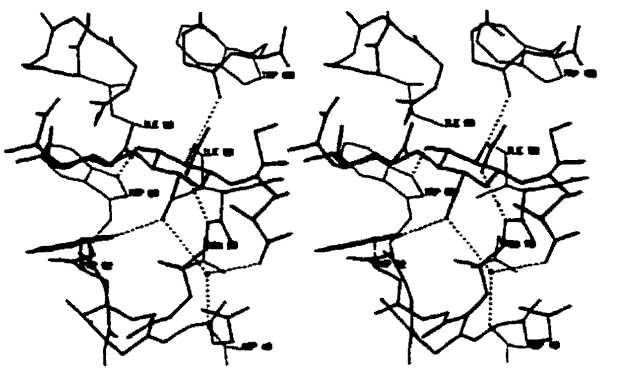


Figure 1.18. Schematic of the inferred saccharide-protein interactions for (A) HEWL, (B) T4L and (C) GEWL. Reproduced from Weaver et al. (1995).

B site (MurNAc)



C site (GlcNAc)



D site (MurNAc)

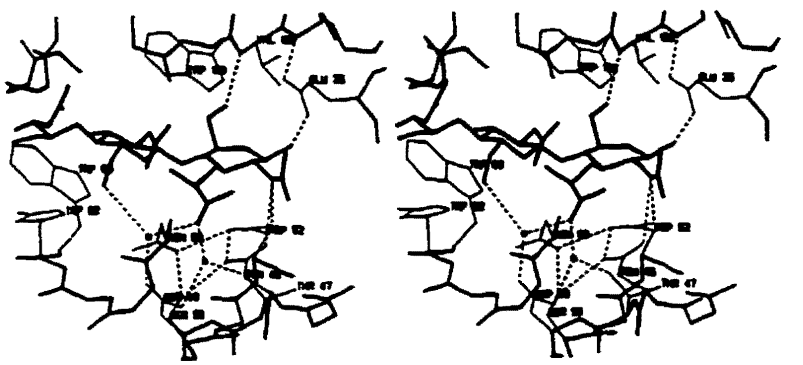


Figure 1.19. Binding of MurNAc-GlcNAc-MurNAc to the B, C, and D sites of HEWL. Reproduced from Strynadka & James (1991).

Using secondary structural predictions however, Fastrez and colleagues have predicted homology between the amino terminal part of LaL and T4L. Residues 9-19 of LaL were proposed to be in a helix with some sequence homology to the helix based on residues 1-11 of T4L (Scheme 1.2). The proposed general acid (Glu11) located at the end of the helix in T4L aligned with Glu19 of LaL. Alignment of fragments containing Glu19 of LaL and Glu35 of HEWL indicated no sequence homology (Jespers et al. 1992).

```
          1                                  11
T4L:      M  N  I  F  E  M  L  R  I  D  E  G
                     :  |  |           |  |
LaL:      K  A  F  L  D  M  L  A  W  S  E  G
          9                             19
```

Scheme 1.2.

Somewhat less confidence was found for the alignment of a residue in LaL equivalent to Asp52 and Asp20 of HEWL and T4L respectively. Secondary structural predictions of a motif of 29 residues in LaL (Gly32-Ser61) suggested that this motif might have possible secondary structure similar to the region of β-structure in HEWL (Jespers et al., 1992). Comparison of the two β-regions of LaL and HEWL led to the suggestion that Asp34 in LaL might be analogous to Asp52 of HEWL even though there is no sequence homology in these regions of the two proteins. A reduction in activity for LaL was observed from mutatation of Asp34 to Asn and Ala (refer to Table 1.3). Interestingly, if the two sequences of T4L and LaL shown in Scheme 1.2 are extended, Asp34 of LaL aligns with Thr26 of T4L. Recall that the mutation of Thr26 to Glu in T4L generated a protein that cleaved peptidoglycan but in which the product remained covalently attached to the glutamic acid.

The recently determined structure of the transglycosylase Slt70 has indicated that the catalytic domain (residues 449-618, or C-Slt70) is very reminiscent of the lysozyme fold with conserved features with SEWL (from swan, a *g* type lysozyme), HEWL and T4L (Fig. 1.20). The structural similarity is very strong despite the absence of any significant sequence homology between C-Slt70 and the lysozymes (Thunnissen et al., 1994; 1995b). Not only is C-Slt70 bilobal (overall dimensions of ≈ 50 × 35 × 30 Å) with a prominent crevice defining the upper and lower lobes, but the proposed general acid of Slt70, Glu478, is located at the end of helix Cα2 (Fig. 1.20) which precedes the only β-structured region of the protein. Like the lysozymes, the acetamido group of a site C GlcNAc residue

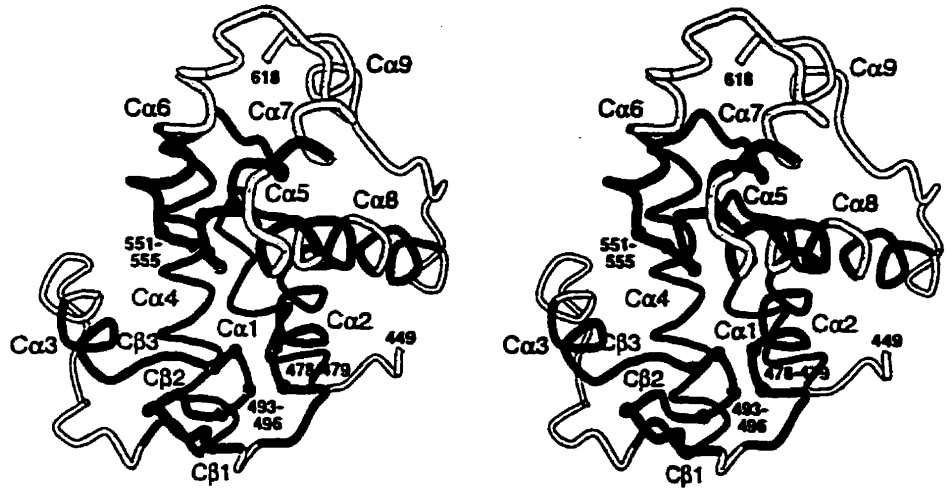


Figure 1.20. Structure of C-Slt70 and conserved features with lysozymes. Regions that are equivalent to regions of HEWL, T4L and SEWL are shown in black, to only one of the three lysozymes are shown in light gray, and to two of the three lysozymes are shown in dark gray. Reproduced from Thunnissen et al. (1995b).

forms ordered hydrogen bonds with the carbonyl of Tyr552 and amide of M498 (Thunnissen et al., 1995b). The C site of Slt70 can only accommodate a GlcNAc residue which would align a peptidoglycan substrate in the cleft such that cleavage would occur between a MurNAc and GlcNAc residue in the D and E sites respectively (Thunnissen et al., 1995b).

It has been reported that the similarity in the amino acid sequence of C-Slt70 with LaL is low (Thunnissen et al., 1995b). However, the structural similarity between C-Slt70 and the lysozymes and the mechanistic similarity between LaL and C-Slt70 (with respect to *anh*MurNAc production) might suggest that LaL could also be structurally related to C-Slt70 and other lysozymes. This confirmation will have to await the elucidation of the three-dimensional structure of lambda lysozyme. Recently, the sequence of the lysis operon of bacteriophage P2, a bacteriophage of *E. coli*, has been reported (Ziermann et al., 1994). The product of the P2 gene *K*, which is required for phage induced lysis and is believed to be an endolysin, was found to have 57-69% sequence homology with LaL, the highest homology known for LaL reported to date. Although as of yet not characterized, it will be of great interest to observe whether the endolysin activity of the product of the P2 *K* gene will be related to LaL with respect to *anh*MurNAc production.

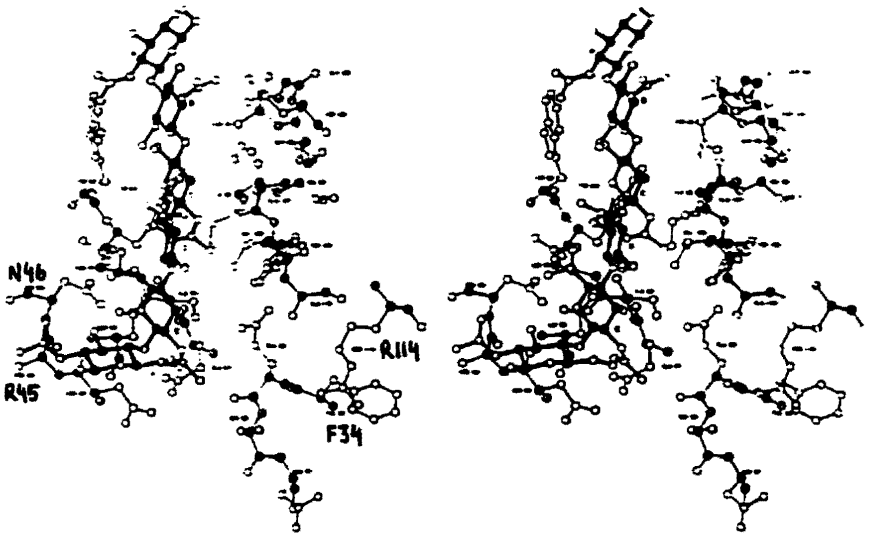
1.7. Theoretical Predictions of Hexasaccharide Binding

There currently exists a dearth of direct structural evidence for any lysozyme concerning the binding of saccharides beyond and including site D. The sites D, E and F are also collectively known as the "lower" part of the active site. The lower site interactions with HEWL have been inferred primarily from model building studies and kinetic observations from transglycosylation reactions (summarized by Imoto et al., 1972). Unfortunately, the known crystalline forms of HEWL preclude the diffusion of large substrates into the active site due to occlusion by neighboring enzyme molecules. A recent attempt to cocrystallise $(GlcNAc)_6$ with the D52S mutant HEWL was unsuccessful due to catalytic hydrolysis of the substrate and only the complex of $(GlcNAc)_4$ filling the A-D sites was observed (Hadfield et al., 1994). An attempted cocrystallization of $(GlcNAc)_6$ with human lysozyme resulted once again in cleavage of the substrate and generated two protein molecules, MOL1 and MOL2 (Song et al., 1994). In MOL2, only $(GlcNAc)_4$ was found in the A-D sites. However in MOL1, $(GlcNAc)_4$ was bound to sites A-D but $(GlcNAc)_2$ was also observed and was bound close to the E and F sites proposed in the Phillips model (Song et al., 1994). The binding of GlcNAc-MurNAc to the lower active site of turkey egg white lysozyme has also been reported (Sarma & Bott, 1977). Therefore, no structural data is available for any lysozyme in which a saccharide is bound "intact" to the entire cleft.

Studies to elucidate the role of the lower active site for this type of complete hexasaccharide binding have emerged. On the basis of conformational energy calculations, it was predicted that HEWL exhibits two distinct lower site binding modes for both $(GlcNAc)_6$ (Pincus & Scheraga, 1979) and $(GlcNAc\text{-}MurNAc)_3$ (Pincus & Scheraga, 1981), a left-sided and a right-sided mode.

In the case of $(GlcNAc)_6$, the structure of lowest energy predicted that the fifth and sixth residues of $(GlcNAc)_6$ bound to E and F sites that involved a β-sheet region of the enzyme including residues Arg45, Asn46 and Thr47 on the left side of the active site (the left-sided structure, Fig. 1.21 A). Calculations on the left-sided structure predicted that the GlcNAc residue in the D site adopts the full-chair conformation and occupies a position that is somewhat removed from the catalytic residues Glu35 and Asp52 and lies close to the surface of the active site cleft. A second low energy structure found (the right-sided structure) is one in which the D site GlcNAc is bound more deeply in the cleft near the catalytic residues but maintains the full chair conformation (Fig. 1.21 B). In the right-sided structure, the fifth and sixth GlcNAc residues bound to E and F sites involving residues Phe34, Asn113 and Arg114 on the right side of the cleft. In both the left- and

(A) Left-Sided: $(GlcNAc)_6$, Full-Chair D Site Residue



(B) Right-Sided: $(GlcNAc)_6$, Full-Chair D Site Residue

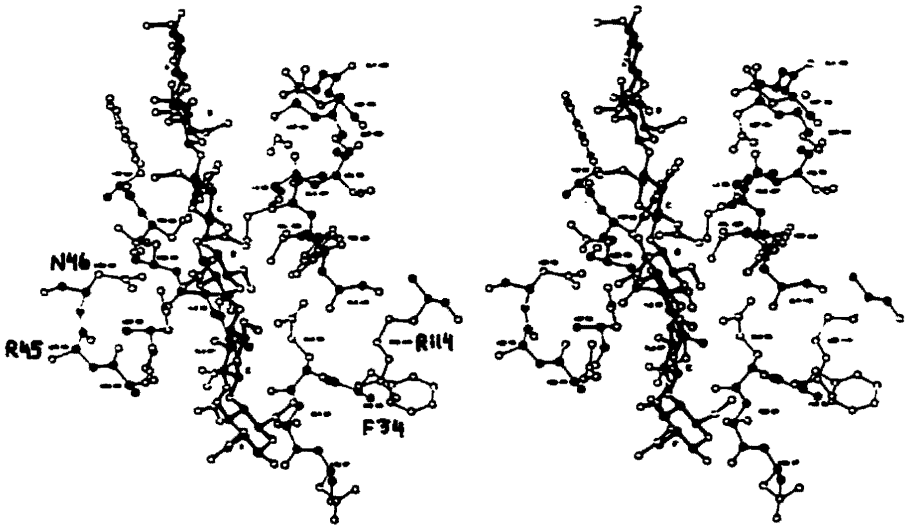
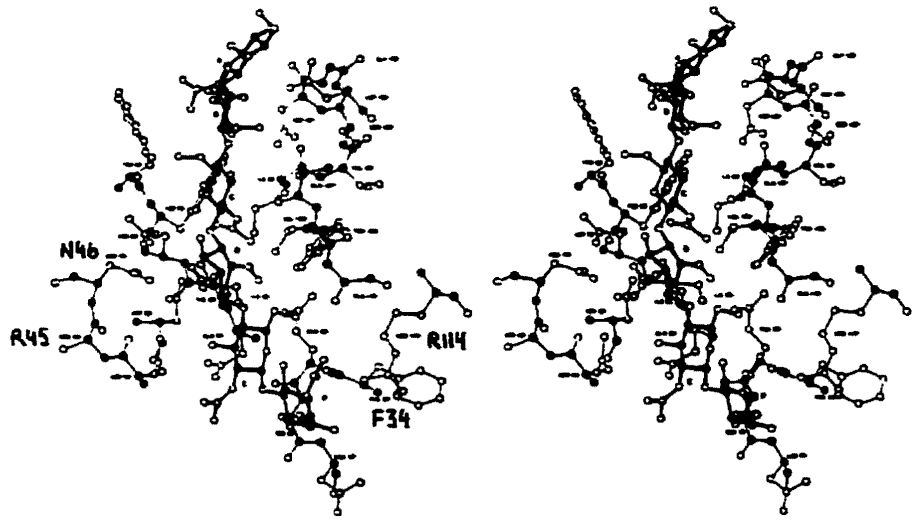


Figure 1.21 (A) and (B). Predicted structures describing left- and right-sided binding of hexasaccharides to HEWL. Shown are the energy minimized structures illustrating the binding mode of HEWL with the indicated hexasaccharide and the site D residue conformation. Reproduced from Pincus & Scheraga (1979).

(C) Right-Sided: $(GlcNAc)_6$, Distorted D Site Residue



(D) Right-Sided: $(GlcNAc\text{-}MurNAc)_3$, Full-Chair D Site Residue

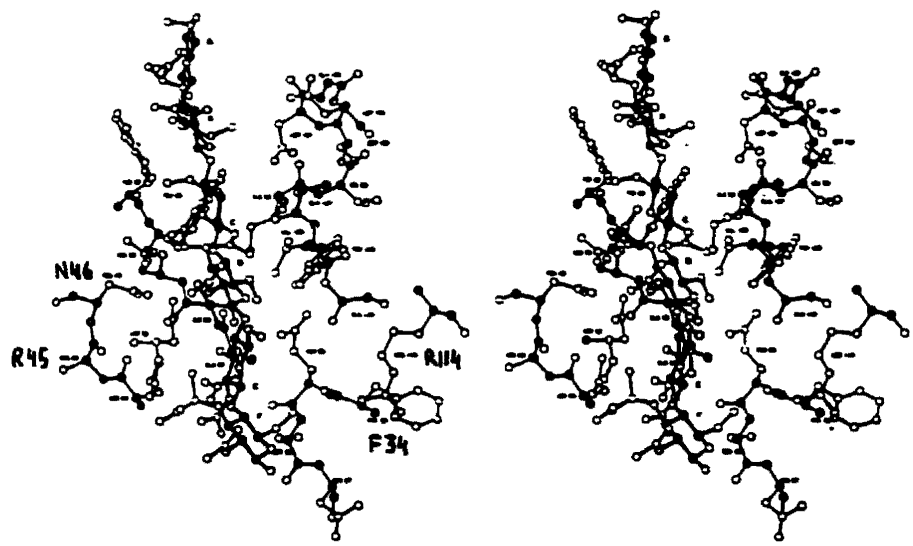


Figure 1.21 (C) and (D). Reproduced from Pincus & Scheraga (1979, 1981).

right-sided structures, the first three GlcNAc residues occupied positions in the A-C sites very similar to that described by the Phillips model. However, the positions of the E and F residues described by the Phillips model were found only for the right-sided structure.

The calculations predicted an overall lower energy for the left-sided structure than for the right-sided structure by approximately 5 kcal·mol^{-1}. However, the conformational energy of the right-sided structure could be lowered even further by approximately 15 kcal·mol^{-1} if the D site residue was distorted to a half-chair conformation. The right-sided-distorted structure (Fig. 1.21 C) made the best contacts for the E and F site residues and was calculated to have the lowest conformational energy of all structures explored if the strain energy for distortion was ignored. However, it was suggested that if the strain energy was taken into account, the overall conformational energy of the right-sided-distorted structure would be higher than that for the left-sided structure (Pincus & Sheraga, 1979).

Although the calculated right-sided structure agreed well with the one proposed by the Phillips model, the left-sided structure was predicted to predominate (Pincus & Scheraga, 1979). It was proposed that initial binding of $(GlcNAc)_6$ to HEWL involved the left side of the cleft. Futhermore, left-sided binding was thought to stabilize a productive but catalytically inactive complex while interactions between $(GlcNAc)_6$ in the right-sided structure stabilized a reactive complex (Pincus & Sheraga, 1979). Experimental evidence has demonstrated that the right-sided mode is not involved in the initial binding step of $(GlcNAc)_6$ to the active site whereas the left-sided mode is. The reaction of $(GlcNAc)_6$ with HEWL and ring neck pheasant lysozyme (RNPL; a c type lysozyme) indicated that although no significant difference in K_m values were observed, k_{cat} for HEWL was three times greater than for RNPL (Smith-Gill et al., 1984). RNPL is identical in sequence to HEWL around the left side of the active site but differs on the right side where Lys and His replace the Asn and Arg of HEWL at postitions 113 and 114 respectively. The authors concluded that the left-sided mode is involved in the formation of the equilibrium complex but the right-sided mode interacts with the substrate in the transition state during catalysis (Smith-Gill et al., 1984). Similar conclusions have been suggested from mutant lysozymes. Mutations of left-sided residues in HEWL were found to affect the stability of the enzyme-substrate complex but not of the transition state while mutations of residues found in the right-side of cleft did not affect equlibrium binding but did affect the energy of the transition state (Inoue et al., 1992b). There is also evidence supporting participation of the right-sided residue Arg115 in human lysozyme (analogous to Arg114 in HEWL) during catalysis from studies on mutants at this position (Muraki et al., 1989).

Involvement of the lower active site during equilibrium binding has also been probed using a monoclonal anti-HEWL antibody which was directed against residues found specifically in the the left side of the lower active site (Smith-Gill et al., 1984). The antibody was competitively displaced from HEWL by $(GlcNAc)_5$ and $(GlcNAc)_6$ illustrating that these substrates bound to the lower cleft in the left-sided mode. $(GlcNAc)_{n\leq4}$ were weak inhibitors of antibody binding and the inhibition was believed to occur in a noncompetitive fashion consistent with the binding of the smaller saccharides only to the A-D sites (or exclusively to site C for GlcNAc, to sites B-C for $(GlcNAc)_2$ etc.) and not to the E and F sites of the lower active site.

In the case of copolymeric GlcNAc and MurNAc hexasaccharides, conformational energy calculations on the binding of these saccharides to HEWL have also predicted both the left-sided and right-sided binding modes (Pincus & Sheraga, 1981). For the saccharide $(GlcNAc\text{-}MurNAc)_2\text{-}(GlcNAc)_2$, left-sided binding was favoured over right-sided binding as was predicted for $(GlcNAc)_6$. The overall conformational energy of the $(GlcNAc\text{-}MurNAc)_2\text{-}(GlcNAc)_2$ left-sided structure was lower than the right-sided structure by approximately 6 $kcal{\cdot}mol^{-1}$. However, the opposite was predicted for the binding mode for $(GlcNAc\text{-}MurNAc)_3$. For this hexasaccharide, the right-sided structure was predicted to be favoured and the overall energy of the right-sided conformation was 10 $kcal{\cdot}mol^{-1}$ less than for the left-sided structure. The right-sided binding conformation for $(GlcNAc\text{-}MurNAc)_3$ is shown in Fig. 1.21 D. Because the site D MurNAc residue of the right-sided structure is positioned deeply in the cleft between the catalytic Glu and Asp residues, the right sided-structure was associated with being the catalytically active conformation (Pincus & Sheraga, 1981).

The reversal of specificity of binding (i.e. $(GlcNAc)_6$ and $(GlcNAc\text{-}MurNAc)_2\text{-}(GlcNAc)_2$ were predicted to bind preferentially to the left-side while $(GlcNAc\text{-}MurNAc)_3$ was predicted to bind preferentially to the right-side of the lower acitve site) is due to unfavourable contacts of the lactic acid side chain with the left-side F site. The lactic acid side chain can be well accomodated in the right-side F site (Pincus & Sheraga, 1981). In addition, energy calculations on several other MurNAc containing saccharides predicted that binding of MurNAc in sites other than the B, D or F was highly unfavourable.

The role of the lower active site in lysozyme classes other than the *c* type has not been explored experimentally and a lower active site may not exist in these lysozymes. Comparison of the main-chain topologies suggested that neither GEWL nor T4L has binding sites analogous to subsites E and F of HEWL (Weaver et al., 1995) although previous efforts from model building did describe an analogous E site (but not an F site)

for T4L (Anderson et al., 1981). The structure of MurNAc-GlcNAc-MurNAc has been modelled into the B, C and D sites of Slt70 but no inference was made regarding enzyme contacts with saccharides that might extend beyond the D site (Thunissen et al., 1995a,b).

1.8. Analysis of the Contributions to Binding of Saccharides to the Subsites of HEWL

A comparison of the binding properties of structurally diverse oligosaccharides to HEWL has permitted estimates to be made regarding the contributions to binding by specific residues. This type of analysis has long been made with the assumption that the various saccharide complexes are superimposable and that monosaccharide residues contribute additively to the unitary free energy of association (Chipman & Sharon, 1969). Table 1.4 summarizes some of the relevant data for the association[1-2] of several saccharides to HEWL. The assumed binding subsites for the saccharides are made with the considerations given in the legend for Table 1.4, with the emphasis that a MurNAc residue cannot occupy sites A, C or E. With this argument, the assignment of the positions of saccharides **9-16** is probably correct. However, more than one binding mode might exist for some of the other saccharides (eg. **4**, **7**). Both productive and nonproductive complexes are possible for chitotetraose in which the reducing sugar resides either in the C or D site (discussed in section 1.9.2). The association constant for $(GlcNAc)_4$ binding over sites A-D ($\approx 2\times10^3$ M^{-1}) was estimated based on the premise that the additional sugar in the tetrasaccharide **12**, which occupies the D site, reduces the association of **12** with HEWL by a factor of 130 relative to the binding of **9**. In analogy, a similar reduction in K_A was proposed for $(GlcNAc)_4$ when bound across the A-D sites (Secemski et al., 1972).

This premise does not appear to be valid. It is seen on comparison of the K_A values for saccharides **11** and **9** that a GlcNAc residue in the D site can actually stabilize the complex with HEWL by 0.6 kcal·mol^{-1}. A XylNAc residue (which lacks the C5 hydroxymethyl group of GlcNAc) in subsite D stabilizes the complex only slightly more by

---

[1-2] In the late 1960's and continued through the 1970's, many studies on the binding of saccharides to HEWL were reported (an excellent summary of these reports are given in Table XXXI from Imoto et al., 1972). The data were compiled from various approaches using NMR, dialysis, inhibition and circular dichroism however UV spectroscopy and fluorescence were two methods predominantly used and considered to be the most reliable. Consequently, values from different sources are found to vary depending on the methodology utilized and attempts will be made in this thesis to cite values determined under similar experimental conditions so that comparisons can be made with some confidence.

Table 1.4. **Comparison of the association of various saccharides to HEWL. Adopted from Schindler et al. (1977a) with additional data taken from other references cited.**

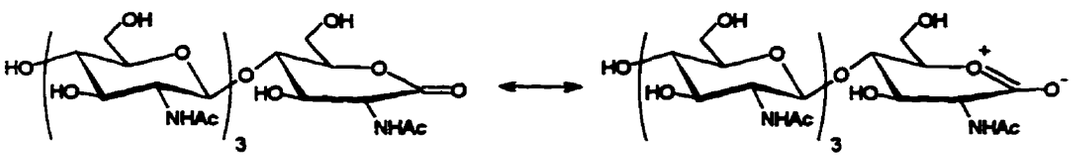
	Assumed Subsite[a]				ΔG_n[b]	K_d (M)[c]	K_A ($^1/_{Kd}$) (M^{-1})	CFGB[d]	ΔG_n Gd[k]
	A	B	C	D					
1		**GlcNAc**			-4.5	2.9×10^{-2}	34		
2		**GlcNAc**-GlcNAc			-7.2	3.0×10^{-4}	3.3×10^{3}		
3		**GlcNAc**-GlcNAc-GlcNAc			-9.3	8.8×10^{-6}	1.1×10^{5}		-9.0
						5.0×10^{-6}[f]	2.0×10^{5}[f]		
4	**GlcNAc**-GlcNAc-GlcNAc-GlcNAc				-9.6	5.3×10^{-6}	1.9×10^{5}		-9.5
						9.1×10^{-6}[f]	1.1×10^{5}[f]		
		GlcNAc-GlcNAc-GlcNAc-GlcNAc					$\approx2\times10^{3}$[f]		
5		**GlcNAc**-GlcNAc-GlcNAc-***XylNAc***			-9.7	4.5×10^{-6}	2.2×10^{5}	-0.4	-9.5
					-11.6[g]	1.8×10^{-7}[g]	5.5×10^{6}[g]		
6		**MurNAc**-GlcNAc			-7.9[e]	9.3×10^{-5}	1.1×10^{4}		
7		GlcNAc-**MurNAc**			-4.1[e]	5.7×10^{-2}	1.8×10^{1}		
			GlcNAc-MurNAc						
8		**GlcNAc**-GlcNAc-***MurNAc***			-6.7	7.0×10^{-4}	1.4×10^{3}	+0.5	
9		**GlcNAc**-MurNAc-GlcNAc			-9.9	3.2×10^{-6}	3.1×10^{5}		-9.5
						3.6×10^{-6}[e]	2.8×10^{5}[e]		
10		**GlcNAc**-MurNAc-GlcNAc-***XylNAc***			-10.8	7.0×10^{-7}	1.4×10^{6}	-0.9	-10.5
11		**GlcNAc**-MurNAc-GlcNAc-***GlcNAc***			-10.5	1.2×10^{-6}	8.7×10^{5}	-0.6	-9.9
12		**GlcNAc**-MurNAc-GlcNAc-***MurNAc***			-7.0	4.2×10^{-4}	2.4×10^{3}	+2.9	-6.7
						4.8×10^{-4}[e]	2.1×10^{3}[e]		
13		**MurNAc**-GlcNAc-MurNAc				2.8×10^{-3}[i]	3.6×10^{2}[i]		
14		**GlcNAc**-GlcNAc-GlcNAc-(***Glcδ-lactone***)			-12.1	7.8×10^{-8}	1.3×10^{7}	-2.8	
						3.0×10^{-7}[f]	3.3×10^{6}[f]		
15		**GlcNAc**-MurNAc-GlcNAc-$\Delta^{2,3}$-***GlcNAc***			-9.6[h]	5.3×10^{-6}[h]	1.9×10^{5}[h]	+0.3	
16		**GlcNAc**-MurNAc-glucose			-6.8[j]		1.8×10^{3}[j]		

[a] The assumed binding region based on considerations of i) known binding modes from crystallographic studies available at the time of this work (namely Imoto et al. (1972) and Ford et al. (1974)) ii) exclusion of MurNAc from sites A, C and E iii) that a residue outside of the cleft cannot affect ΔG_n. [b] Unitary free energy of binding in kcal·mol^{-1}. [c] Calculated from $\Delta G_n = RT \ln K_d$ - 2.4 kcal·mol^{-1} (Schindler et al., 1977a). [d] Contribution From Group in Bold type expressed as difference between ΔG_n for indicated compound and that for compound lacking the residue in bold; since ΔG_n measurements are dependent on experimental conditions only values from this work (pH 6.6, 25 °C) are compared. All values taken from (Schindler et al., 1977a) except for Chipman et al. (1967)[e], Secemski et al. (1972)[f], van Eikeren & Chipman (1972)[g], Schindler & Sharon (1976)[h], Patt et al. (1978)[i] and Chipman & Sharon (1969)[j]. [k] Determined in the presence of 20 mM $GdCl_3$.

The preparation of the unique saccharides containing muramic acid (**8**, **11**, **13**) or with novel derivatives (**8**, **10**, **14**, **15**, **16**) was achieved by transglycosylation reactions using the appropriate acceptor with either $(GlcNAc)_n$ or $(GlcNAc\text{-}MurNAc)_2$ as the substrate (Pollock & Sharon, 1970; Schindler et al., 1977a).

0.9 kcal·mol^{-1} for **10** or by only 0.4 kcal·mol^{-1} for **5**. The $\Delta^{2,3}$-2-acetamido-2,3-dideoxy-*D*-glucose residue ($\Delta^{2,3}$-GlcNAc in **15**) has little effect on the complex when bound in the D site (the inhibition of HEWL is 35 times greater by **15** than by **12** (Schindler & Sharon, 1976)). These observations can be interpreted to indicate that, at least for a tetrasaccharide bound in sites A-D with a terminal GlcNAc residue, the hydroxymethyl group *does not* lead to severe steric interactions with the protein in site D. The "strain" of 2.9 kcal·mol^{-1} observed for **12** by the introduction of a MurNAc residue into site D must therefore be largely due to the lactate ether of the muramic acid. Indeed, the structure of the HEWL/MGM complex indicated that the lactyl group of the site D muramic acid lies in close proximity to the side chains of Asn46 and Asn59 and is thereby forced to make close contacts with the GlcNAc residue in site C (Strynadka & James, 1991).

The tetrasaccharide TACL (Scheme 1.3) is presently the saccharide of highest affinity to bind to HEWL and exhibits a dissociation constant of approximately 80 nM (**14**, Table 1.4). The crystal structure of the HEWL/TACL complex indicated that TACL bound to sites A-D but the crystal data could not precisely distinguish the structure of the terminal δ-lactone in the D site (Ford et al., 1974). Best fits were obtained by modelling the lactone in the sofa and the $B_{3,O}$ boat conformations (refer to Fig. 1.25, p. 73) although the sofa conformation was favoured by the authors as this conformation more appropriately satisfied the planar nature of the lactone (Ford et al., 1974). Because of this planar nature and a distribution of positive charge (Scheme 1.3) that is expected to be similar to the carbonium ion for lysozyme catalysis, TACL was proposed to behave as a transition state analogue (Secemski et al., 1972; Ford et al., 1974). The terminal lactone of TACL is bound in the D site only 30-40 times or by 2.2 kcal·mol^{-1} (Table 1.4) more tightly than a GlcNAc residue (Secemski et al., 1972; Schindler et al., 1977a). This enhancement in affinity was seen as insignificant compared to the acceleration ascribed to



TACL
(tetra-N-acetylchitotetraose δ-lactone)

Scheme 1.3.

lysozyme catalysis and the status of TACL as a transition state analogue was questioned by Chipman and colleagues (Schindler et al., 1977a). These authors have also suggested that the hydrogen bond that forms between the lactone carbonyl and Asp52 in the crystal (Ford et al., 1974) may account for part of the affinity of TACL[1,3] .

If the residues at the reducing termini of tetrasaccharides **8**, **10**, **11** or **12** (Table 1.4) do indeed bind to the D site in positions similar to that proposed in the Phillips model, then C1 and O1 (anomeric oxygen) of these residues would lie between the side chains of Asp52 and Glu35. Table 1.4 lists the effect of the lanthanide, gadolinium (Gd(III)), on the affinity of HEWL for several saccharides. Lanthanide ions have been shown to inhibit both the activity of HEWL and the binding of chitosaccharides to HEWL, producing in the latter, dissociation constants of 0.2-0.8 mM (Secemski & Lienhard, 1974; Perkins et al., 1981a). Crystallographic studies have shown that there are two Gd(III) ions bound directly between the two catalytic carboxyls (that are separated by 3.6 Å) and are thought to occupy the same positions in solution (Kurachi et al., 1975; Perkins et al., 1979).

The fact that the presence of Gd(III) does not markedly decrease the energies of association of **8**, **10**, **11** or **12** (Table 1.4) led to the suggestion that the terminal groups of the tetrasaccharides in question are bound in the D site in positions that are different from that proposed by the Phillips model or observed for the complex with TACL (Schindler et al., 1977a). These authors propose that it is likely that these groups are not bound deeply in the cleft but rather lie near its surface. This suggestion is in agreement with the calculated lowest energy conformation of the HEWL/$(GlcNAc)_6$ complex (the left-sided structure; Fig. 1.21 A) which also predicted that the position of the site D residue was somewhat removed from the cleft region.

The concept initiated by Beddell et al. (1970) that if a residue in the D site is moved significantly out of the cleft relative to its position described in the Phillips model, then the residue in the C site must also change appears valid. For example, the calculated contribution of a given residue in the D site to ΔG_u is variable (i.e. compare the contribution of MurNAc for **8** and **12** or of XylNAc for **8** and **10**) which implies that the

---

[1,3] TACL is also a substrate for HEWL being hydrolysed to the free acid. Frank (1992) has noted on examination of the data presented by Secemski et al. (1972) that HEWL catalyzes ring-opening of TACL approximately 13 times that of the rate of spontaneous hydrolysis. This observation was also seen as being suggestive of endocyclic cleavage (Frank, 1992).

complexes in question are not superimposable particularly when the saccharide extends into the D site. Therefore the contribution to binding of a given residue is not only a function of the contacts in the D site, but will also include the energy associated with conformational changes elsewhere in the complex that are caused by the introduction of the given residue (Schindler et al., 1977a).

Other notable data can be extracted from Table 1.4 and from a summary of the estimated contributions of the enzyme-saccharide interactions at each subsite (Table 1.5). A comparison of the association constants of **3** with **2** or of **9** with **6** (Table 1.4) indicates that the contribution of a GlcNAc residue in the A site enhances binding approximately 30-fold or by 2 kcal·mol^{-1} (Table 1.5). The association of a saccharide with a MurNAc residue in site B is greater than the association for the corresponding saccharide with a GlcNAc located at this position. Differences in the binding energies between compounds **6** & **2**, **9** & **3**, **10** & **5**, and **11** & **4** amount to a stablilization of 0.6-1.1 kcal·mol^{-1} (Table 1.4). This level of stabilization is reasonable considering that in the crystal structure of the HEWL/MGM complex the lactyl carboxylate of the MurNAc residue in site B makes a hydrogen bond to Asn130 (Strynadka & James, 1991). Lastly, the binding importance of

Table 1.5. Estimated binding contributions of the subsites in HEWL. Adapted from Chipman & Sharon (1969) and Imoto et al. (1972).

		ΔG_u (kcal·mol^{-1})		
	Site	GlcNAc	MurNAc	Contacts[†]
Group 1	A	-1.8 – -2.3		7
	B	-2.7 – -2.9	-3.4	11
	C	-4.6 – -5.7		30
	D	-0.6	+2.9	35
Group 2	D	≈ +6.2		
	E	-4		45
	F	-1.4 – 1.9	≈ -1	13
	E and F	-1.7		
	D, E and F	≈ +1.2		

[†] Number of contacts <4 Å between enzyme and saccharide. The Group 1 values are estimated from the comparisons of ΔG_u for saccharides which are assumed to bind only to sites A-D. See text for details concerning the Group 2 values.

the acetamido group contributes 3.1 kcal·mol^{-1} to binding at the C site as illustrated by comparison of the binding affinities of saccharides **16** and **9** (Table 1.4).

It is difficult to accurately estimate the contributions to binding of saccharides in the D-E sites. From comparison of the binding energies for (GlcNAc-MurNAc)$_2$ (binding to sites A-D) and for (GlcNAc-MurNAc)$_3$ (binding to sites A-F) the contribution to binding of GlcNAc-MurNAc to sites E-F was estimated to be -1.7 kcal·mol^{-1} (Chipman & Sharon, 1969; Table 1.5). A similar comparison of the binding energies for (GlcNAc-MurNAc)$_3$ and for GlcNAc-MurNAc-GlcNAc (binding to sites A-C) suggested that the sum of the interactions of MurNAc-GlcNAc-MurNAc to sites D-F was approximately +1.2 kcal·mol^{-1} (Chipman & Sharon, 1969; Table 1.5). Therefore, the binding of MurNAc to site D can be estimated to be +2.9 kcal·mol^{-1} from the difference in the D-F and E-F binding energies (i.e. (+1.2) - (-1.7) = +2.9). In this case, the calculation of the E-F binding energy has included the binding energy for (GlcNAc-MurNAc)$_2$ in which the terminal MurNAc residue is bound in the D site.

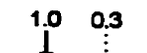
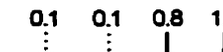
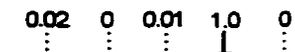
From studies on acceptor reactivities in transglycosylation reactions in which the acceptor is assumed to bind in the E and F sites, the individual binding energies for GlcNAc and MurNAc in sites E and F were estimated to be -4 and -1 kcal·mol^{-1} respectively (Table 1.5). In this case, when the sum of the latter values for binding in sites E and F (i.e. -5 kcal·mol^{-1}) is subtracted from sum of binding to sites D-F (+1.2 kcal·mol^{-1}), the difference indicates that the cost of binding to site D may be as high as +6.2 kcal·mol^{-1} (Table 1.5). From these observations, it was suggested that when MurNAc is bound in the D site as part of a saccharide which binds to only the A-D sites (i.e. a terminal MurNAc residue), it is not "strained" to the same extent as would be a MurNAc residue bound in site D in a saccharide that fills the cleft completely in sites A-F (Chipman & Sharon, 1969).

1.9. Lysozyme/Chitooligosaccharide Interactions

1.9.1. Cleavage Patterns and Transglycosylation Reactions

The reactions of chitooligosaccharides have been intensively investigated with HEWL and have revealed that the processes involved in the binding and hydrolysis of these compounds are complex. As illustrated in Table 1.6 it is seen that although the relative rates of cleavage differ drastically from the trimer to hexamer, there is no appreciable difference in their association constants (K_A), the free energies of association (ΔG) nor in their K_m values (Rupley & Gates, 1967).

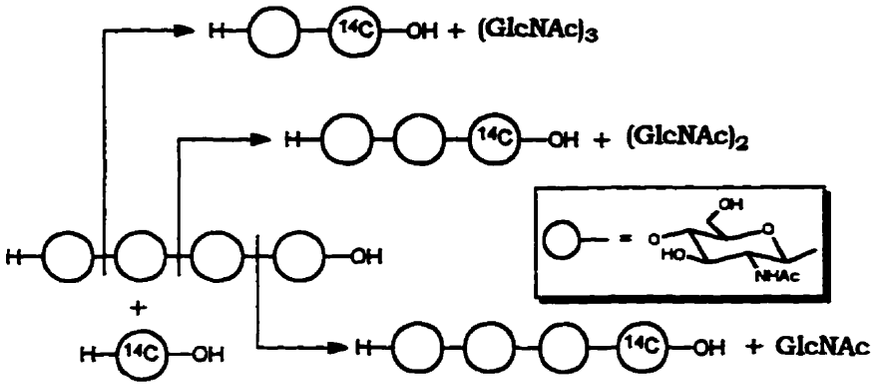
Table 1.6. Binding and cleavage data of chitooligosaccharides with HEWL.

	Association Data [a]		Kinetic Behavior [b]			
	K_A (M^{-1})	$-\Delta G$ (kcal/mol)	k_{cat}[c] (s^{-1})	Relative Rate	Order	Cleavage Pattern
GlcNAc	17-25	4.1-4.3				
$(GlcNAc)_2$	5.7×10^3	7.5	2.5×10^{-8}	0.003	1	
$(GlcNAc)_3$	1.3×10^5	9.3	8.3×10^{-6}	1	1	
$(GlcNAc)_4$	1.06×10^5	9.2	6.6×10^{-5}	8	1	
$(GlcNAc)_5$	1.07×10^5	9.2	0.033	4 000	0	
$(GlcNAc)_6$	1.62×10^5	9.5	0.15	30 000	0	

Taken from Dahlquist et al. (1966)[a], Rupley & Gates (1967)[b] and Imoto et al. (1972)[c]. Kinetic behavior is given for conditions of a 1:1 substrate:enzyme complex saturation (substrate concentration = 10^{-4} M). Cleavage patterns determined from transfer reactions to [^{14}C]-(GlcNAc) showing major (↓) and minor (⋮) cleavage sites and the relative cleavage rates within each substrate is indicated above the arrow.

The data given in Table 1.6 were determined by identification of products formed from the substrates by glycosyl transfer reactions (i.e. transglycosylation) in the presence of high concentrations of [^{14}C]-GlcNAc. The transfer reaction was informative and allowed the identification of the cleavage sites. The use of this strategy is illustrated in Scheme 1.4 demonstrating the cleavage pattern for $(GlcNAc)_4$. For each of the cleaved bonds, unique transfer products are obtained which can be separated and quantitated to give the cleavage pattern and the relative rates of cleavage (Table 1.6). It was later demonstrated

that GlcNAc is approximately 2000 times more effective than water as the acceptor during the catalytic reaction (Rupley et al., 1968). Studies of the hydrolysis reactions were less informative since cleavage of symmetrically located bonds would generate identical products but gave results consistent with those for the transfer reaction (Rupley, 1967; Rupley & Gates, 1967; Imoto et al., 1972).



Scheme 1.4.

The glycosyl acceptor is not limited to a monosaccharide as indicated in Scheme 1.4. Under appropriate conditions, HEWL will catalyze the formation of insoluble chitin products on incubation with $(GlcNAc)_4$ by way of transglycosylation (Kravchenko, 1967). For example, $(GlcNAc)_4$ can be cleaved generating a $(GlcNAc)_3$ donor which reacts with a second $(GlcNAc)_4$ acceptor to yield $(GlcNAc)_7$. The latter can be cleaved (for example between residue 4 and 5) producing a tetrasaccharide donor which reacts with a $(GlcNAc)_n$ acceptor yielding $(GlcNAc)_{n+4}$. Higher oligomers are propagated similarly. In addition, higher oligomers of $(GlcNAc\text{-}MurNAc)_n$ are formed when the tetrasaccharide $(GlcNAc\text{-}MurNAc)_2$ is reacted with HEWL (Sharon, 1967). The transglycosylation activity of HEWL has found application in the synthesis of substrates for this enzyme in which GlcNAc-PNP (Usui et al., 1988) or GlcNAc-4MU (Byers & Leaback, 1988) serve as the glycosyl acceptors. In all cases mentioned thus far, transglycosylation proceeds with formation of a $\beta(1\rightarrow4)$ glycosidic bond between the glycosyl donor and acceptor. Recently, an interesting HEWL transglycosylation product was reported. When $(GlcNAc)_2$ was incubated with α-PNP-maltopentaoside, GlcNAc was transferred to the non-reducing unit of the pentasaccharide (the acceptor) with formation of a β-linkage between C1 of GlcNAc and C3 of the acceptor (Hidenori et al., 1990).

The kinetic behavior listed in Table 1.6 is a familiar set of experimental data often cited today. Interpretation of the results (Sharon, 1967; Rupley & Gates, 1967) has suggested an elaborate mechanism of cleavage for the different chitooligosaccharides. A brief account of these interpretations is offered below.

The relative rates and cleavage pattern listed in Table 1.6 were described for the condition in which one molecule of $(GlcNAc)_n$ binds to one molecule of HEWL, i.e. *1:1 complex saturation* ($(GlcNAc)_{3-6}$ essentially saturate the enzyme at 10^{-4} M and $(GlcNAc)_2$ does at 10^{-2} M; Rupley, 1967). Under these conditions, only in the case of chitohexaose is a unique bond essentially cleaved. With reference to the Phillips model, $(GlcNAc)_6$ is just able to completely fill sites A-F and is therefore cleaved between residues 4 and 5 (i.e. D and E subsites). For any other bond to be cleaved, the hexasaccharide must bind such that all the enzyme sites (or at least one) are not occupied. Such "incomplete" complexes would require loss of binding interactions that will disfavour the formation of any other than the "completely filled" complex (Rupley & Gates, 1967) and are considered as nonproductive complexes if the positions in sites D and E (in which lies the reactive bond) are not occupied simultaneously (productive and nonproductive complexes are addressed in section 1.9.2). The pentasaccharide can form two types of complete complexes occupying sites A-E or site B-F. As such, cleavage of these complexes between sites D and E occur at essentially equal rates between the 3rd and 4th or 4th and 5th residues respectively (Table 1.6). The rates of $(GlcNAc)_5$ and $(GlcNAc)_6$ cleavage approach zeroth order kinetics as saturation is reached; i.e. below 10^{-4} M, the reaction is dependent on concentration (first order), increases with increasing substrate concentration until saturating concentrations are reached and the rate becomes independent of substrate concentration (zeroth order). However, the *cleavage patterns* for the pentamer and hexamer show dependence on concentration. The relative bond cleavage rates at concentrations of 10^{-2} M for $(GlcNAc)_6$ are 0.3, 0.2, 0.2, 1.0, and 0 and for $(GlcNAc)_5$ are 1.9, 0.3, 2.2, and 1.0 (values are read from the non-reducing end as listed in Table 1.6). As will be discussed below, this observation is believed to result from a reaction of a complex containing two molecules of substrate. However, the data obtained at substrate concentrations of 10^{-4} M should represent reaction of a 1:1 saturated complex (Rupley & Gates, 1967).

The trimer and tetramer demonstrated complex cleavage patterns that were cleaved at unequal rates (Table 1.6). In addition, the relative cleavage rates for 10^{-4} M $(GlcNAc)_{2-4}$ are extremely slow (as compared to the pentamer and hexamer; Table 1.6) and show dependency on substrate (first order) at all concentrations (as high as 0.5 M; at the higher concentrations, the rate approaches that of $(GlcNAc)_5$). It was suggested that cleavage of a 1:1 complex (even if bridging sites D and E) of $(GlcNAc)_{2-4}$ may be negligible; however, the

binding of a second chitosaccharide unit may provide additional and required interactions by filling the active site in such a way that one of the substrate molecules is cleaved rapidly (Rupley & Gates, 1967). Therefore, the observed rate would depend on the concentration of the ternary complex (i.e. 2 saccharides bound per active site) and will be first order in substrate at concentrations that will result in ternary complex formation; i.e. above the concentration that leads to 1:1 complex saturation. The rate was too slow to be measured below saturating concentrations but is also expected to be first order in substrate concentration.

1.9.2. Productive and Nonproductive Complexes

The concept of productive and nonproductive binding of substrates to HEWL was initiated to explain the apparently identical association constants but different rates of hydrolysis observed for $(GlcNAc)_n$ ($n \geq 3$). In an extensive and detailed study of the equilibrium, thermodynamics, and kinetics (including pre-steady state and steady-state) of the reaction of HEWL with GlcNAc oligosaccharides, Rupley and co-workers (Banerjee & Rupley, 1973a, b; Holler et al., 1975a, b; Banerjee et al., 1975) have described the binding modes of chitohexaose involved in its reaction with HEWL. A model depicting a summary of their observations of the enzyme forms important for cleavage of chitohexaose is shown in Fig. 1.22.

Nonproductive complexes of $(GlcNAc)_6$ are viewed as having formed through binding of saccharide to the ABC sites of the enzyme with the reducing unit occupying site C while the non-reducing units of saccharides longer than the three units sufficient to fill sites ABC extend into the solvent "beyond" site A. The crystal complex of D52S HEWL with $(GlcNAc)_5$ indeed showed occupancy of a sugar above site A (termed the A-1 site) that made no contacts to its own molecule in the crystal but did interact with a symmetry related molecule (Hadfield et al., 1994). The formation of nonproductive complexes occurs in a process which has been called the α process (Holler et al., 1975b). The α process is a two-step process in which the complexes ENP_1^{α} and ENP_2^{α} (Fig. 1.22) are formed in a rapidly established pre-equilibrium (process α_1) and in a complex isomerization (process α_2) that occur on a time scale of about 5 ms (Holler et al., 1975b). The two-steps of the α process are observed for each of the oligosaccharides $(GlcNAc)_3$ through $(GlcNAc)_6$, and the same set of kinetic and thermodynamic parameters describes these steps for each of these saccharides.

The structure of the stable nonproductive complex ENP_2^{α} is defined by the structure determined from crystallographic analysis of the chitotriose complex and the

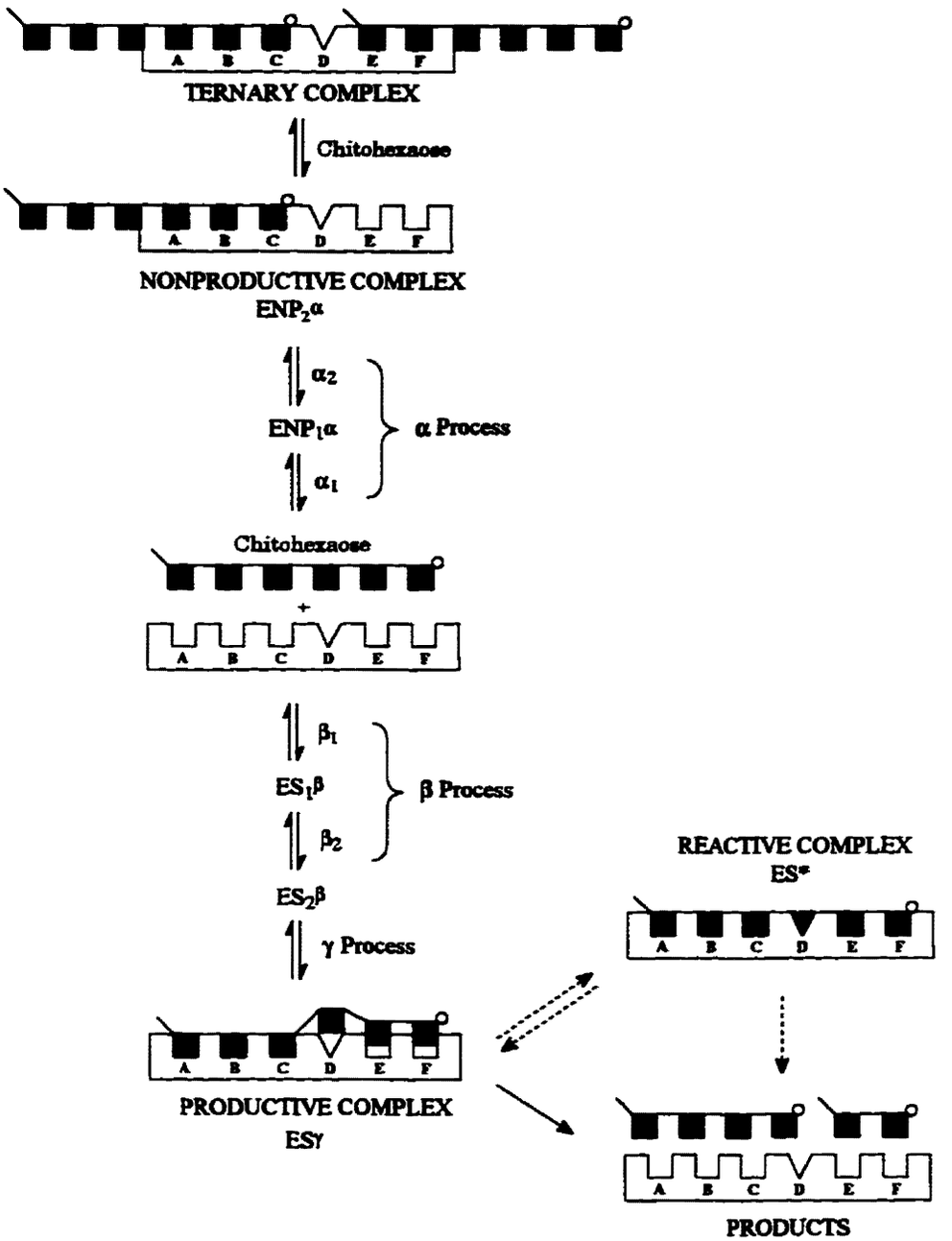
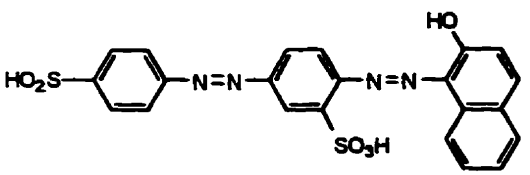


Figure 1.22. Schematic model representing the enzyme forms important for the lysozyme catalyzed cleavage of chitohexaose. The nonproductive complexes (ENP) are formed in the α processes with the reducing end unit (☚) at site C of the enzyme. Productive complexes (ES) formed in the β and γ processes involve binding of the non-reducing unit (☛) at site A. Further details of the model are given in the text. Adopted from Holler et al. (1975a, b) and Banerjee et al. (1975).

specific interactions of bound trimer with the ABC sites described by the Phillips model. The principal enzyme-substrate interactions found in ENP_2^{α} are established in the rapid bimolecular step (α_1) that occurs on formation of ENP_1^{α}. Large changes in thermodynamic parameters (ΔH and ΔS) were found to occur for the interconversion of ENP_1^{α} and ENP_2^{α} (step α_2, Fig. 1.22) and were thought (Banerjee et al., 1975) to reflect changes in enzyme structure that serve to cause a narrowing of the cleft which lead to the optimal interactions of the cleft region with the bound substrate (the specifics of these conformational changes will be described in section 1.9.3). Of relevance, approximately one-quarter of the total fluorescence change for the α process develops on isomerization of ENP_1^{α} to ENP_2^{α}, which reflects the movements of Trp62 and Trp108 in making hydrophobic contacts with substrate in sites B and C respectively.

The existence of the nonproductive ternary complex illustrated in Fig. 1.22, although as of yet not observed directly, is advocated from several lines of evidence. It was found that the dye, Biebrich Scarlet, inhibits HEWL competitively and forms a 1:1



Biebrich Scarlet

complex with the enzyme (Rossi et al., 1969). The absorption spectra of the dye differs for the free dye, the dye/HEWL complex and the dye complexed with nonproductive complexes of HEWL. More elaborate equilibrium studies suggested that the dye interacts only with either the E or F subsite since chitohexaose and not chitotriose displaces the dye (Holler et al., 1975a). Part of the equilibrium data (only conducted for $(GlcNAc)_{n\geq 3}$) was rationalized if the saccharides can form ternary complexes involving sites A-C and E-F simultaneously (Holler et al., 1975a). Similar investigations, but with more detailed kinetic analysis involving the GlcNAc monomer through trimer, again revealed that there are two sites for the simultaneous binding of these saccharides (Ikeda & Hamaguchi, 1976). In contrast to the results of Holler et al. (1975a), the binding of the saccharides $(GlcNAc)_{n=1-3}$ in sites A through D did displace the dye but in a noncompetitive fashion (Ikeda & Hamaguchi, 1976). As was discussed, the simultaneous binding of more than one molecule was also used to justify the first-order rate character (Table 1.6) of trimer and tetramer hydrolysis (Rupley, 1967; Rupley & Gates, 1967). More recently, inhibition of the binding of an anti-HEWL monoclonal antibody directed to the lower active site was

inhibited by $(GlcNAc)_4$, yet the pattern of inhibition was, in part, explained in terms of the co-binding of two molecules of $(GlcNAc)_4$ (Smith-Gill et al., 1984).

The nonproductive complexes are so defined since the substrate does not extend across the reactive D-E cleavage site. Stable *productive* complexes are also formed which can then be transformed into reactive complexes (Holler et al., 1975a). The productive complexes (see ES, Fig. 1.22) for the hexasaccharide are viewed with substrate bound in the ABC sites (as in the nonproductive complex) but with the non-reducing unit now occupying site A. The three saccharide units at the reducing end extend into and *partially* fill the DEF subsites, as indicated by the partial filling of these sites in Fig. 1.22. In this sense, $(GlcNAc)_6$ is "lined up" with the A-F subsites. Productive complexes are observed *only* for $(GlcNAc)_{4-6}$ (hence the immeasurably slow relative rate of dimer and trimer cleavage; Table 1.6). In the case of $(GlcNAc)_4$ and $(GlcNAc)_5$, the reducing end of these sugars would protrude only into the D and DE region of the cleft respectively.

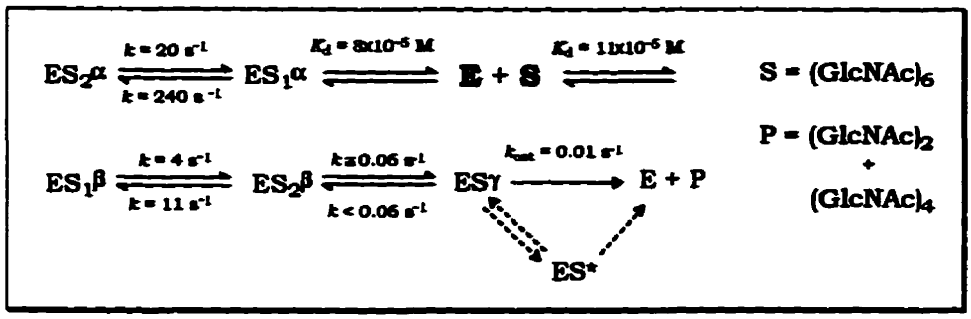
The first productive complexes in the reaction path are ES_1^{β} and ES_2^{β} (Fig. 1.22) which are formed during what is called the β process (Holler et al., 1975b). The β process is also a two step process including a rapid bimolecular reaction (process β_1) and complex isomerization (process β_2). A third productive complex, $ES\gamma$, forms through another isomerization of ES_2^{β} during the γ process. The γ process occurs in a 1-10 s time period and again is only observed for tetra-, penta- and hexasaccharide. Formation of $ES\gamma$ does not contribute to the stability of the complex. It was suggested that the contacts between enzyme and saccharide in the DEF region are still not fully developed in $ES\gamma$. The properties of the $ES\gamma$ complex suggested that during the rate-determining step there must still be movement of part of the substrate or enzyme to allow development of more complete interaction between enzyme and substrate. These changes which would include the ring distortion, were thought to come (but unobserved at the time experimentally) in a possible reaction to give a reactive complex (ES*) (as indicated by the broken arrows in Fig. 1.22) which then breaks down into the products. Alternatively, the substrate (most notably in site D) might move fully into the cleft during the same process in which the glycosidic bond is broken during catalysis (Banerjee et al., 1975).

The structures of the productive complexes (ES_1^{β}, ES_2^{β} and $ES\gamma$) were believed to be similar in structure with each other and have been compared to the structure of the complex of HEWL with GlcNAc-β(1→4)-xylose, which has been studied in the crystal (Beddell et al., 1970) and in solution (Banerjee et al., 1974). The GlcNAc unit of this disaccharide binds in a similar manner, but with some rearrangement to the Phillips

model, in site C. However the xylose unit in site D would be bound in an atypical orientation such that any additional units attached at the reducing end of this disaccharide would be unable to make the proper contacts in the EF sites as described in the Phillips model. A similar structure describes the complex with GlcNAc-β(1→4)-glucose (Beddell et al., 1970). Hence, the argument that there is not full development of interactions in the DEF sites of ESγ was described. It was discussed previously that theoretical energy calculations have predicted three low-energy structures for the productive complex of HEWL with $(GlcNAc)_6$ (i.e. the left-sided, right-sided and right-sided-distorted structures; refer to section 1.7). It is interesting to speculate whether the three low-energy calculated structures may be representative of any of the productive complexes (perhaps $ES_2\beta$, ESγ and ES*) inferred to experimentally and that the isomerization processes β_2 and γ may describe conformational changes that occur between and result in one of the three predicted structures.

There are some particular points of interest that can be concluded regarding the properties of the nonproductive and productive complexes. Perturbation of the pK_a of Glu35 occurs during the isomerization steps involving conformational changes in the complexes (i.e. the α_2 and β_2 processes). Relevant pK_a values determined for this group are: 6.1 for E, ENP_1^{α}, ES_1^{β}; 6.4 for ENP_2^{α}; and 6.7 for ES_2^{β} (Banerjee et al., 1975). Interactions with sites A-C are the same in both the productive and nonproductive complexes. Stable productive complexes may result when a saccharide binds in the active site (even across sites D-E) but without full intrusion toward a productive reactive complex. A recent crystal structure of GEWL with $(GlcNAc)_3$ showed the trisaccharide binding in sites B-D without distortion (Weaver et al., 1995). It was suggested that the $(GlcNAc)_3$/GEWL complex could represent a stable productive complex in which the saccharide did not penetrate fully towards a reactive complex. Therefore, it is thought that the energy for transition from a stable to a reactive complex is provided by interactions of sugar units within subsites E and F (Weaver et al., 1995; Holler et al., 1975b).

Further, nonproductive binding is about 3 times better than productive binding for $(GlcNAc)_6$ at 25 °C (Holler et al., 1975a). The reaction with hexamer indicated that the α process is 10-20 times faster than the β process, which is about 100 times faster than the γ process. The cleavage reaction is about 10-fold slower than the γ process and has nearly twice the energy of activation. An example of the parameters pertaining to the minimal



Scheme 1.5.

pathway for the HEWL catalyzed hydrolysis of $(GlcNAc)_6$ at pH 6.3 and 25 °C is shown in Scheme 1.5 (Holler et al., 1975b). At the pH optimum of chitosaccharide hydrolysis (pH 5.2) and 40 °C, k_{cat} for hydrolysis of $(GlcNAc)_6$ is 0.14 s^{-1} (Banerjee et al., 1973).

The reactions shown in Scheme 1.5 and Fig. 1.22 were corroborated by similar investigations of the individual steps in the reaction of HEWL with $(GlcNAc)_3$ and $(GlcNAc)_6$ at subzero temperatures (Fink et al., 1980). This work again delineated the occurrence of a two step α process (for trimer and hexamer) and the three reactions (β1, β2 and γ process) for hexamer only. In contrast to the case at 25 °C, productive binding for hexamer is favoured with decreasing temperature. As well, the data showed tighter binding of both sugars with decreasing temperature.

A rational explanation to the data given in Table 1.6 can now be suggested. Recall that the experiments of Rupley & Gates (1967) indicated that although $(GlcNAc)_{3-6}$ bound to lysozyme with the same affinity, there were substantial differences in their cleavage rates. It is now evident that the experimentally determined association constants must in some way reflect the collective steps that develop during the processes of productive and nonproductive binding. With the assumption that ES^{γ} (or ES*) represents the reactive lysozyme-substrate complex, Holler et al. (1975b) have suggested that the observed velocity of the catalytic reaction would be expected to be directly proportional to its concentration. The additional interactions which chitohexaose can make with the enzyme, as compared to tetramer or pentamer, are expected to lead to a higher concentration of ES^{γ}, and therefore, to a higher rate of hydrolysis.

1.9.3. Conformational Changes

Conformational changes in enyme structure accompany sugar binding. These changes have been described for the HEWL molecule from early crystallographic studies (Blake et al., 1967; Imoto et al., 1972) but the discussion herein will draw briefly from more recent structural investigations. Strynadka & James (1991) report that in general, the changes in the molecular conformation of HEWL upon complexation with MurNAc-GlcNAc-MurNAc in the crystal are subtle (the r.m.s. deviation of all protein atoms is 0.79 Å and for the main chain atoms is 0.28 Å). The changes involve a narrowing of the substrate cleft by approximately 1.0 Å overall as the residues and secondary structural units that line the sugar binding site move in towards the bound saccharide. Similar minor conformational adjustments of the overall protein molecule are also described for GEWL/$(GlcNAc)_3$ complex (Weaver et al., 1995). The observed cyrstallographic changes compare well to theoretical predictions which suggested that the dominant motions of HEWL involved a slight opening and closing of the active site cleft *via* concerted movements of the flanking structural regions (Levitt et al., 1985).

In particular, three specific regions of HEWL exhibit the greatest conformational differences between the native and complexed molecule (Strynadka & James, 1991). These regions are: R1, Asn59-Cys64; R2, Val99-Thr118; and R3, Arg68-Cys76 (the general locations of R1,R2 and R3 are shown in Fig. 1.23). The regions R1 and R2, contained in the lower and upper lobes of HEWL respectively, contribute directly to substrate binding. Residues Asp52, Trp62, and Trp63 (located in R1) and residues Asp101, Asn103 and Val109 (located in R2) all shift in position to accomodate the incoming substrate. Stacking of the non-polar face of the site B sugar against the side chain of Trp62 results in a shift of the indole ring towards the sugar by up to 1.4 Å (Hadfield et al., 1994). Movement of the catalytic Asp52 (by approximately 0.8 Å, corresponding to small rotations within the side chain) is observed while the conformation of the catalytic Glu35 is identical in the native and complexed enzyme (Strynadka & James, 1991).

The third region comprising residues 69-76 (R3, Fig. 1.23) forms a loop near the surface of HEWL. Athough this region is removed from the active site and does not make direct contact with the substrate, the region shifts in position and becomes more ordered with bound saccharide (Strynadka & James, 1991). A dramatic change on complexation involving the peptide bond between Arg73 and Asn74 has been observed in the crystal complexes (Strynadka & James, 1991; Hadfield et al., 1994). This peptide bond flips 180° and was thought to be correlated with the movement of Trp62 since both the side chains and main chain of residues 73 and 74 and of tryptophans 62 and 63 are in close contact (< 5 Å). The corresponding surface loop of the lysozyme from partridge egg-white (comprising

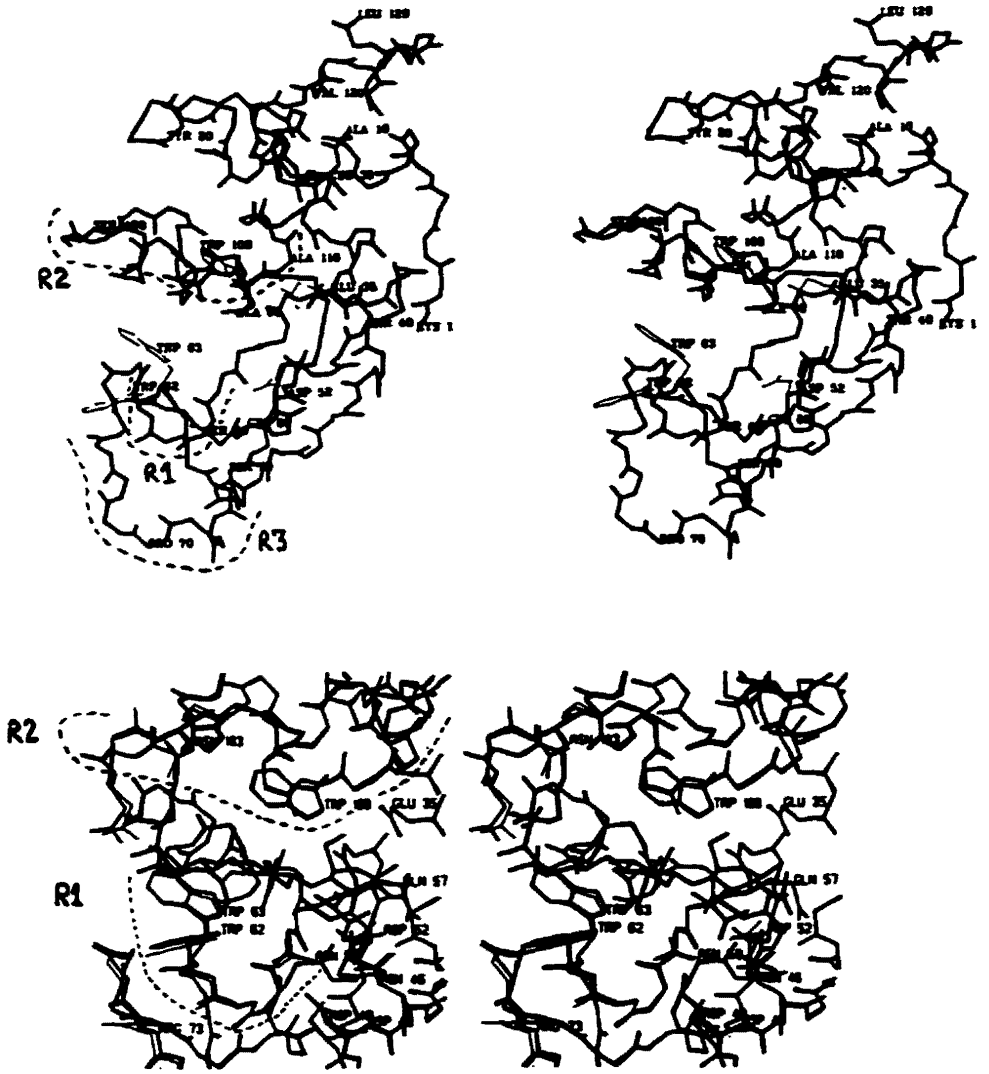


Figure 1.23. Regions of HEWL that experience conformational changes on substrate binding. An attempt has been made at approximating the regions R1, R2 and R3 (refer to text for details) with the dashed lines. The bottom structure is an overlap of native (thin lines) and sugar bound HEWL (thick lines). The MurNAc-GlcNAc-MurNAc is bound to sites B-D. Reproduced from Strynadka & James (1991).

residues 70-75 of this enzyme) was also observed to shift and become more ordered on binding $(GlcNAc)_3$ (Turner & Howell, 1995). Movement of the loop towards the active site was suggested to result in the anchoring of Trp62 (Turner & Howell, 1995).

Two-dimensional proton NMR solution studies (Lumb & Dobson, 1992; Lumb et al., 1994) of the complexes of chitosaccharides with HEWL gave results which were

consistent with the conformational changes observed in the diffraction studies. Furthermore, the majority and largest chemical shift changes of enzyme resonances induced on binding of GlcNAc were in the vicinity of site C in the active site cleft in accordance with the location of the GlcNAc binding site in the crystal. Similar changes in chemical shift were induced by the binding of $(GlcNAc)_2$ and $(GlcNAc)_3$ to those caused by formation of the monomer complex, and these changes in chemical shift on complex formation increase from the monomer to the trimer (there were, however, some resonances perturbed by both dimer and trimer that were not affected by GlcNAc binding). For every significant change observed for GlcNAc binding, a parallel change (i.e. upfield or downfield) was caused by $(GlcNAc)_2$ and $(GlcNAc)_3$. The results were interpreted to indicate that the induced conformational changes are similar for the three inhibitors, arise predominantly from interactions with site C and are governed primarily by interactions with the N-acetyl group. As well, small changes in chemical shift were observed in regions of the protein remote from the active site indicating that effects of sugar binding are transmitted throughout the enzyme.

Unfortunately, the details on the changes in T4L structure have not been described in any detail except for the specific interactions of the molecule with bound saccharides (Anderson et al., 1981; Kuroki et al., 1993). There is experimental evidence which suggests that the two lobes (domains) that comprise T4L (between which is formed the active-site cleft) undergo a hinge-bending motion. The support for hinge-bending came from an interesting study on the M6I mutant T4 lysozyme (Faber & Matthews, 1990). Not only does the mutant crystallize in the standard form with a structure very similar to the wild type protein, but it also crystallizes such that the lattice comprises four independent molecules of different structure. Including the wild type, each of the five structures has a different hinge-bending angle between the two domains. It is possible that the different conformers in the unit cell result from crystal packing forces. However, a recent molecular dynamics simulation on *wt* T4L provided support to the hinge-bending motion and two hinge loci were located (Arnold et al., 1994). It was originally suggested that the hinge-bending was important to allow access of substrates to the active site of the enzyme (Matthews & Remington, 1974). This type of conformational change of the entire protein for T4L appears to be of a more global nature than for HEWL in which movements of the protein predominate in the cleft region; however, hinge-bending has also been advocated to occur in HEWL (McCammon et al., 1976; Sternberg et al., 1979).

In the T4L/$(GlcNAc)_3$ complex in which an additional GlcNAc was modelled into site D, the carboxyl oxygen of the proton donor Glu11 is 5 Å from the glycosidic oxygen of

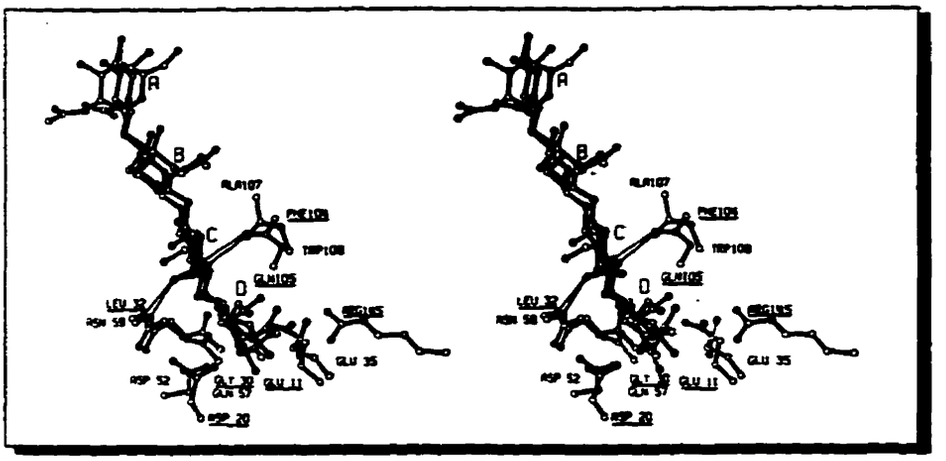


Figure 1.24. Comparison of elements in the active site (including a bound $(GlcNAc)_4$) of T4L (solid bonds, residues underlined) and HEWL (open bonds). Reproduced from Matthews et al. (1981b).

the D sugar to which it would donate its proton (Anderson et al., 1981). As well, Glu11 is linked by a salt bridge to Arg145, in contrast to the nonpolar environment of Glu35 in HEWL. This places Glu11 close to, but not in exactly the same relative position as Glu35 of HEWL (Fig. 1.24). A slight conformational change upon substrate binding was suggested to be possible which could result in removal of the salt bridge and movement of Glu11 into the required position to act as the general acid (Matthews et al., 1981a,b).

1.10. Considerations of Distortion

The original Phillips model detailing the binding of $(GlcNAc)_6$ to HEWL suggested that the C5-hydroxymethyl group of the site D residue was the likely source of ground-state distortion in the substrate. In the model, the C6 and O6 atoms of the site D GlcNAc would make seriously close contacts with the main chain carbonyl of Gln57 and the side chains of Trp108 and Glu35 of the protein. In addition, the hydroxymethyl atoms would also encounter steric clashes with the acetamido group of the GlcNAc residue in the C site. This overcrowding could be relieved by distortion of the chair conformation of the site D GlcNAc to a sofa or half-chair conformation in which the C5-hydroxymethyl substituent assumes a more axial position (Blake et al., 1967; Imoto et al., 1972; Phillips, 1974).

It is now believed that the primary source of distortion results from the close approach of the C5-hydroxymethyl group of the D site residue with the acetamido group of the site C site GlcNAc. In the HEWL/MGM crystal structure, the atoms of the site C acetamido groups were found to have very low thermal factors indicating that the acetamido group was well ordered and tightly bound in the C site. In a full-chair conformation with an equatorial orientation for the site D MurNAc C5-hydroxymethyl group (Fig. 1.25), the atoms of the hydroxymethyl group would make seriously close contacts with the highly ordered site C acetamido group (see also Fig. 1.26 A). In order to relieve these steric contacts, the hydroxymethyl group of the site D MurNAc is forced into the quasi-axial position observed in the crystal (Strynadka & James, 1991). The site D MurNAc residue most closely approached a conformation in which the ring atoms C1, C2, C4, C5 and O5 are approximately co-planar. This conformation can best be described by the sofa or half-chair ($^4H_3$) configurations (Fig. 1.25). The quasi-axial orientation of the C5-hydroxymethyl group in the sofa and half-chair conformations is compared to the equatorial orientation for this substituent in the full chair conformation in Fig 1.25. Although not considered to be a prominant source contributing to distortion, the lactyl group of the site D MurNAc packs tightly against the side chains of Asn46 and Asn59 thereby being forced to make close contacts with the ring oxygen and C6 of the GlcNAc residue in site C (Strynadka & James, 1991; see also Fig. 1.26. A).

There are now four independent crystallographic investigations in which the D site residue has been observed with a distorted (sofa or half-chair) conformation: the study of TACL with HEWL (Ford et al., 1974); the mutant D52S HEWL/$(GlcNAc)_4$ complex (Hadfield et al., 1994); the study of the HEWL/MGM complex (Strynadka & James, 1991); and the covalent complex of T26E T4L with the cell wall peptidyl-disaccharide (Kuroki et al., 1993). The structures of the distorted MurNAc residues from the latter two investigations are presented in Fig. 1.26. Conversely, there have also been crystallographic studies in which the D site residue was found to be in an undistorted conformation. The originally investigated HEWL/MGM complex (Kelly et al., 1979), the crystal structure of $(GlcNAc)_4$ with human lysozyme (Song et al., 1994) and the complex of $(GlcNAc)_3$ with GEWL (Weaver et al., 1995) all have indicated a full-chair conformation for the D site GlcNAc residue.

A key but frequently overlooked feature of the Phillips model is that the binding positions of the site D-F residues were modeled based on the "static" structure determined for the complex of $(GlcNAc)_3$ with HEWL (Blake et al., 1967; Phillips, 1974). It appears that the model building assumed that no further conformational changes in either the enzyme or of the bound trisaccharide would occur upon addition of the remaining residues of

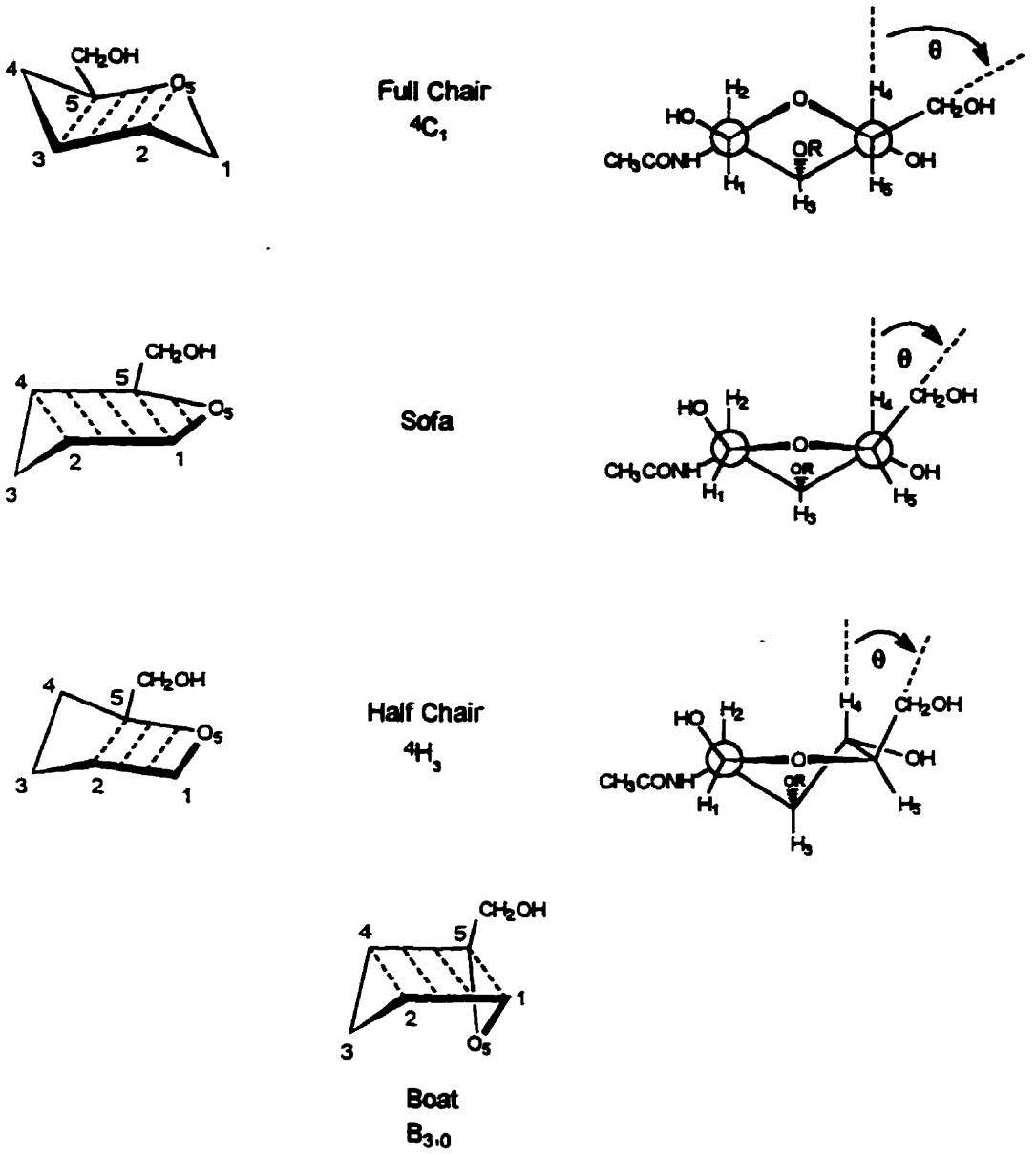


Figure 1.25. Possible conformations of a pyranose ring. Adapted from Ford et al. (1974), Phillips (1974) and Strynadka & James (1991).

Broken lines indicate reference planes containing ring atoms that are coplanar. Conformations are designated C for full-chair, H for half-chair, B for boat and sofa as indicated. Ring atoms which lie above the reference plane are written as superscript and preceed the letter while ring atoms which lie below the plane are written as subscripts and follow the letter.

The substituents for the Newman projections on the right are for a MurNAc residue (R = $CH(CO_2H)(CH_3)$). The change in position of the C5-hydroxymethyl group from an equatorial position in the full chair conformation to a quasi-axial orientation in the half-chair and sofa conformations is indicated.

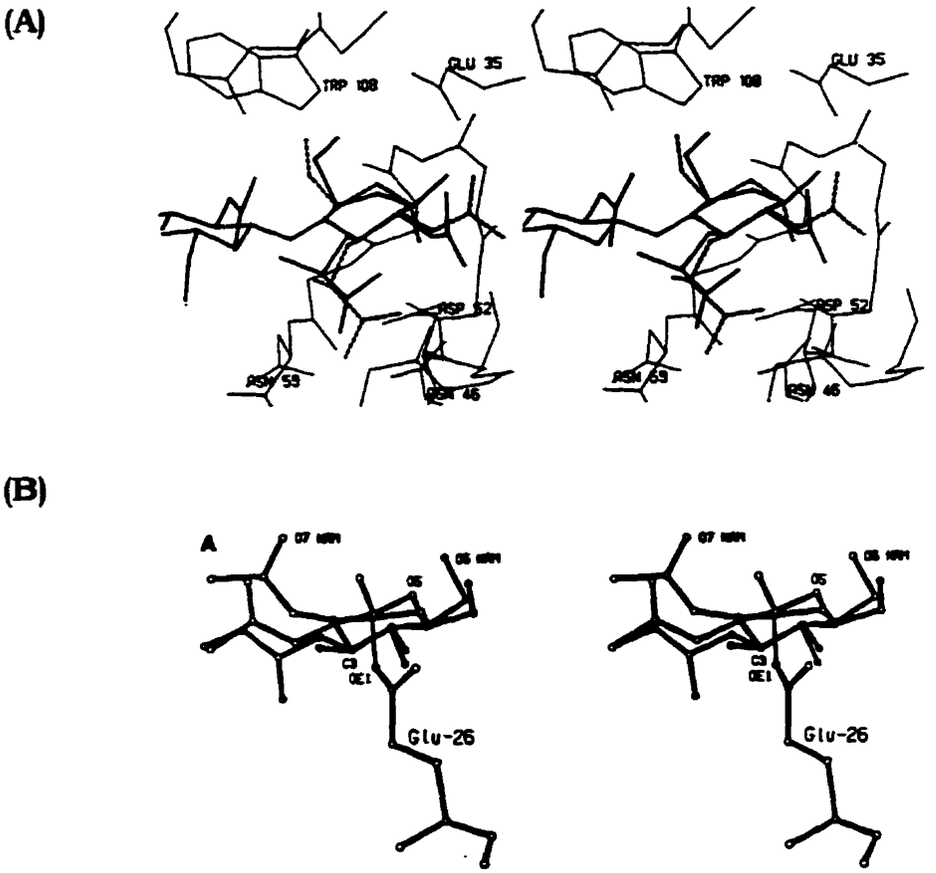


Figure 1.26. Distorted site D MurNAc residues observed in crystal structures of HEWL and T4L.

(A) Superposition of the MurNAc residue bound in the D site of the HEWL/MGM complex (thick lines) on the conformation of MurNAc determined from crystallographic studies (broken lines). Reproduced from Strynadka & James (1991).

(B) Superposition of the MurNAc residue bound in the D site (solid bonds) on the GlcNAc residue from site C (open bonds) of the T26E T4L peptidyl-disaccharide complex. The MurNAc is covalently linked to Glu26. Reproduced from Kuroki et al. (1993).

$(GlcNAc)_6$ in filling the D-F sites. It is reasonable to assume that if other conformational changes result upon further occupation of the lower sites, then it is possible that the site D residue can be accomodated with favourable interactions but without its distortion. Conversely, conformational changes may arise upon occupation of the lower sites in which the energy required to distort the D site residue is provided by newly created interactions with the enzyme that have resulted from these conformational changes.

Another important concept is evident from the crystallographic studies. If emphasis is placed on these results then as noted above, there are equal arguments from structural data supporting and discrediting the distortion of the D site residue. Although the determined binding pattern of residues to sites A-C seems to be consistent not only for repeated studies with HEWL but also for lysozymes from other sources (human, goose, T4) there is an apparent difference for the D site residue as well as to what is suggested for the lower active site positions. However, some of the enzyme-saccharide structures observed may simply be dependent on how the complexes were prepared (i.e. the crystallographic conditions and the technique for the introduction of the saccharide to the protein crystal) and the ensuing interpretation based on the quality of the data and the extent of refinement of the crystal structure. A report by Phillips and colleagues exemplifies the importance of refinement strategies and approaches to prepare protein crystals with bound ligands (Cheetham et al., 1992).

When Phillips and coworkers initially reported the 2.5 Å structure of HEWL complexed with the trisaccharide MGM bound to the B-D sites, the interpretation of the structure led to a conclusion that did not involve a distorted MurNAc residue in the D site (Kelly et al., 1979). However, more recent studies performed at higher resolution (1.5 Å) and with more rigorous refinement of the same HEWL/MGM complex have demonstrated that Phillips' original interpretation appeared to be incorrect and that the D site MurNAc residue is indeed distorted towards a sofa or half-chair conformation (Strynadka & James, 1991). Hadfield and colleagues have also reported a high resolution (1.9 Å) structure of D52S HEWL complexed with $(GlcNAc)_4$ and reported distortion of the site D GlcNAc. Conversely, the complex of GEWL/$(GlcNAc)_3$ refined to 1.6 Å resolution did not show distortion of the site D GlcNAc although the geometry of the D site sugar ring was not well defined due to high thermal factors for the ring atoms (Weaver & Matthews, 1995).

It must also be asked whether the crystal structures themselves are true representations of the solution complexes. The solution binding of MGM to HEWL by NMR analysis is repeatedly cited as a study offering evidence to refute the importance of ring distortion. The coupling constant between H1 and H2 of the terminal MurNAc residue of MGM indicated that the dihedral angle (θ) between H1 and H2 does not change upon binding to HEWL (Patt et al., 1978). For a saccharide in the full-chair conformation having an axial anomeric OH (i.e. α-anomer), $\theta \approx 60°$ and $J_{H1,H2} \approx 2\text{-}3$ Hz. As the ring becomes distorted towards a sofa conformation, θ decreases towards 0° and $J_{H1,H2}$ increases. The authors measured a coupling constant of ≈ 2.6 Hz for both the free and

enzyme bound saccharide and concluded therefore that binding of MGM to HEWL does not induce distortion.

These results are somewhat questionable. The crystal structure of the HEWL/MGM complex indicated that the α-anomer of the MurNAc residue in the D site cannot be accommodated due to potentially extreme contacts with Asp52 (Strynadka & James, 1991). Furthermore, the structure indicated the probability of a hydrogen bond between the anomeric hydroxyl and Glu35 which would favour the binding of the β-anomer. Binding of the β-anomer is also supported by the known stereospecificity of HEWL in retaining the β-anomer upon cleavage of substrates (Dahlquist et al., 1969). The study of Patt et al. (1978) also failed to take into consideration that both GlcNAc (Johnson, 1966) and GlcNAc-MurNAc (Kuroki et al., 1995) mutorotate to equilibrium mixtures of approximately 60% α and 40% β anomer. Finally, the authors do not offer an explanation as to the contradiction in results obtained several years earlier wherein changes in the coupling constant between H1 and H2 of the reducing residue of the tetrasaccharide GMGM were interpreted as evidence for distortion (Patt et al., 1974).

There are two crucial considerations involved with the structural studies. Firstly, it is particularly interesting to note that the two X-ray structures with cell wall substrates in which MurNAc binds to the D site have only been observed in the distorted form (i.e. the HEWL/MGM structure (Strynadka & James, 1991) and the T26E T4L/peptidyl-disaccharide structure (Kuroki et al., 1993)). This conclusion accepts that the original interpretation of the HEWL/MGM structure (Kelly et al., 1979) is incorrect. Only in the case of homopolymers of GlcNAc were both undistorted and distorted conformations reported. Could this suggest that binding of substrates that more appropriately resemble the natural substrate demand distortion in the D site as a consequence of other interactions with the peptide and lactyl group? Secondly, the variety of saccharides examined which have bound in the crystal structures have extended only into the D site. No studies have emerged in which the conformation of the D residue was observed for a substrate that extended over the D and E sites, although there have been unsuccessful attempts using $(GlcNAc)_6$ with HEWL (Hadfield et al., 1994) and human lysozyme (Song et al., 1994). These studies have been hampered by the difficulty of the diffusion of the substrate into the crystal, by the obscuring of the lower active site by symmetry related molecules and by hydrolysis of the substrate. Such studies would give more insight into whether binding of saccharide to enzyme occurs with or without distortion of the D site saccharide.

The various theoretical energy minimization procedures[1-4] of hexasaccharide binding agree in concluding that low energy structures are possible in which there is no distortion of the D site residue in the ground state of a lysozyme-substrate complex. In addition, the unexpected finding that both GlcNAc and XylNAc have favourable binding energies of similar magnitude for the D site (refer to section 1.8 and Table 1.4) suggest no distortion of these residues in the ES complex. Favourable interactions of a XylNAc residue are easily rationalized since it lacks both the lactyl and C5-hydroxymethyl groups of a muramic acid residue. It is possible to position a GlcNAc residue in a full-chair conformation without unfavourable contacts by a slight rotation of the glycosidic bond between the site C and D residues in the Phillips model. This rotation is allowed for GlcNAc but not for MurNAc because of the presence of the lactyl group (Strynadka & James, 1991). Since both the C5-hydroxymethyl and lactyl groups are present in MurNAc, MurNAc is expected, and does, exhibit the greatest unfavourable contribution to binding (+2.9 kcal·mol^{-1}, Table 1.4 & 1.5). This may suggest, as noted above for the structures of HEWL/MGM and T26E T4L/peptidyl-disaccharide, that MurNAc binds in the energetically more demanding distorted state.

There is unambiguous evidence however that interactions between HEWL and the the C5-hydroxymethyl group of the substrate saccharide residue bound in the D site are important for catalysis. It appears that whether or not there is distortion in the ground state ES complex, interactions of the hydroxymethyl group in the complex at the transition state and not the ES complex are an important contributor to catalysis with intrinsic binding energy with the hydroxymethyl group being expressed at the transition state.

In consideration of the reaction coordinate of an enzymic reaction (Fig. 1.27), it was postulated that the standard free energy of activation for the reaction, $\Delta G^{\ddagger}_{T}$, could be written as the sum of an adverse energy term, $\Delta G^{\ddagger}$, which is involved in the chemical process of bond breaking and forming and an energetically favorable term, ΔG_s, due to the realization of the enzyme/substrate binding energy (Fersht, 1974).

$$\Delta G^{\ddagger}_{T} = \Delta G^{\ddagger} + \Delta G_s$$

[1-4] The theoretical investigations reported by Pincus & Scheraga (1979, 1981) were described earlier in section 1.7. Warshel and Levitt (1976) have also predicted from theoretical calculations that the full-chair conformation of the D site is favoured over the sofa conformation.

The binding energy of the transition state may be more favorable than that of the initial ES complex because (i) an energetically favourable interaction is created in ES‡ that does not occur in the ES complex, or (ii) there is an energetically unfavourable interaction in the initial ES state that does not exist in ES‡ (or a combination of these). Further, case (i) can lead to a lowering of $\Delta G^{\ddagger}_{T}$ and an increase in k_{cat}/K_m through a lowering of the free energy level of ES‡; (ii) would lead to an increase of the free energy level of ES, with complementary increases in both k_{cat} and K_m but would not affect $\Delta G^{\ddagger}_{T}$ (Fersht, 1974). Therefore, substrate distortion in the initial state of the ES complex would lead to higher values for k_{cat} and K_m and to more effective catalysis at high substrate concentrations but would not necessarily lead to a higher value of k_{cat}/K_m as described by Fersht (1974).

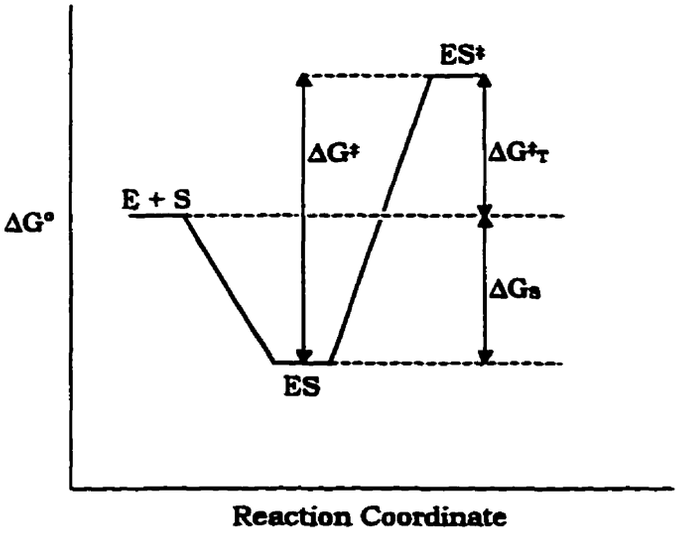
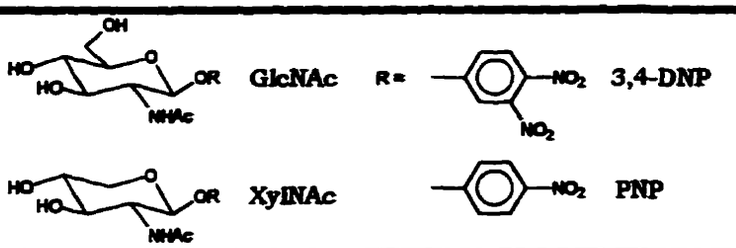


Figure 1.27. Change in standard free energy for an enzymatically catalyzed reaction. Adapted from Fersht (1974).

In order to examine the structural requirements for catalysis, Dearie and coworkers synthesized a series of substrates containing XylNAc (Capon & Dearie, 1974; Ballardie et al., 1977). These authors showed that whereas HEWL catalyzed release of PNP or 3,4-DNP from the chitosaccharide substrates, the XylNAc derivatives were not detectably hydrolysed. The XylNAc derivatives did show slow spontaneous hydrolysis and the measurements of these rates enabled calculation of the upper limits to k_{cat}/K_m for the lysozyme-catalyzed reaction with these substrates (Table 1.7). The difference in the value of k_{cat}/K_m for the hydrolysis of $(GlcNAc)_3$-3,4-DNP and $(GlcNAc)_2$(XylNAc)-3,4-DNP set a lower limit on the difference in the standard free energy of activation (i.e. $\Delta G^{\ddagger}_{T}$) of 4.3

Table 1.7. Values of k_{cat}/K_m ($M^{-1}s^{-1}$) for the HEWL catalyzed hydrolysis of XylNAc and GlcNAc nitrophenyl-glycosides (Ballardie et al., 1977).



$(GlcNAc)_2(XylNAc)$-PNP	0	$(GlcNAc)_3$-PNP	0.28
$(GlcNAc)_3(XylNAc)$-PNP	0.002[a]	$(GlcNAc)_4$-PNP	0.90
$(GlcNAc)_2(XylNAc)$-3,4-DNP	0.006[a]	$(GlcNAc)_3$-3,4-DNP	7.9

[a] upper limit fixed by reproducibility of measurements of the rate of spontaneous hydrolysis.

kcal·mol^{-1}. That is there is a lowering of the free energy of the enzymic transition state of 4.3 kcal·mol^{-1} for the GlcNAc substrate than the XylNAc substrate from interaction of the enzyme with the hydroxymethyl group of the GlcNAc residue.

Since there is a change in $\Delta G^{\ddagger}_{T}$, case (i) described above appears to be the operative one in which favourable interactions are created upon formation of the transition state and that ground state distortion in the ES complex does not occur. This is not compatible with the original Phillips mechanism. An interpretation of these observed results with the XylNAc and GlcNAc substrates and that which is predicted by the Phillips mechanism is suggested in the reaction coordinates presented in Figure 1.28.

The free energy levels for $(GlcNAc)_2(XylNAc)$-3,4-DNP (S_X) and $(GlcNAc)_3$-3,4-DNP (S_G) are shown in Fig. 1.28 A. As mentioned earlier, the finding that both GlcNAc and XylNAc have favourable binding energies of similar magnitude for site D suggest that the complexes, ES_X and ES_G, are at a comparable energy level and that there is no strain involving interactions with the hydroxymethyl group in ES_G. It is also possible that ES_G is strained when compared to ES_X and is therefore, at higher energy (Fig. 1.28 A). As shown, there may exist only a limited number of interactions (↔) with the C5-hydroxymethyl group in ES_G and none with ES_X. Upon formation of the transition states, $ES^{\ddagger}_G$ is 4.3 kcal·mol^{-1} lower in energy than $ES^{\ddagger}_X$. This is because there is now favourable interaction with the C5-hydroxymethyl group in $ES^{\ddagger}_G$ that cannot occur in $ES^{\ddagger}_X$. $\Delta G^{\ddagger}_G$ is therefore lower than $\Delta G^{\ddagger}_X$ as was determined experimentally.

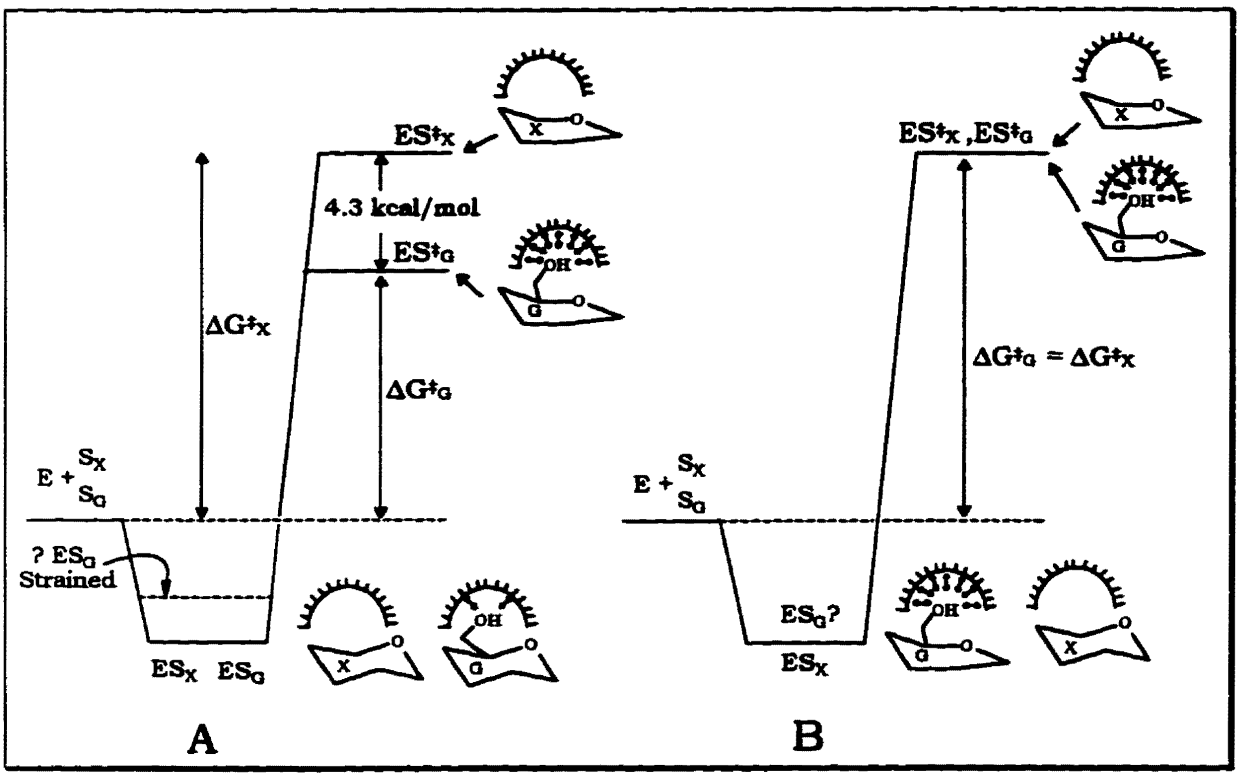


Figure 1.28. Possible reaction coordinates describing the HEWL catalysed reaction with XylNAc or GlcNAc containing substrates.
(A) Reaction coordinate based on experimental data described in the text.
(B) Reaction coordinate predicted by the Phillips model.
S_X and S_G are substrates with a XylNAc or GlcNAc residue bound at the D site. The semicircle with the double headed arrows indicates interactions of the enzyme with the C5-hydroxymethyl group.

In the Phillips model (Fig. 1.28 B), ground state binding of GlcNAc (i.e. ES_G) does not occur without distortion. For the reasons discussed previously, ES_X and ES_G are shown at comparable energy levels but the actual level of ES_G is not certain (the energy level of ES_G in the Phillips model is expected to be different than that in A). In accord with the Phillips model, XylNAc (i.e. ES_X) is not expected to require ring distortion for favourable binding since it is lacking the C5-hydroxymethyl group. The energy cost for distorting the GlcNAc ring must be provided in part, by the favourable interactions with the C5-hydroxymethyl group in ES_G. However upon formation of $ES^{\ddagger}_G$, no new interactions can occur with the C5-hydroxymethyl group and the energy levels of $ES^{\ddagger}_G$ and $ES^{\ddagger}_X$ would be as shown (Fig. 1.28 B) and no net difference in $\Delta G^{\ddagger}_G$ and $\Delta G^{\ddagger}_X$ would result. Therefore, *a priori* substrate distortion in the initial state of the ES complex is not

compatible with the differences in catalysis observed for the XylNAc and GlcNAc containing substrates reported by Dearie and colleagues.

The use of the term "strain" with respect to the lysozyme mechanism can be confusing. For example, in his review on glycosyl group transfer, Sinnott uses the term as follows in discussion of the Phillips model: the model "envisaged introduction of strain in the ES complex which was relieved at the transition state" (Sinnott, 1987). Does strain refer to the crowding of the undistorted saccharide in the D site with the enzyme or does it pertain to the strain conferred to the sugar ring on distortion. This is often an ambiguity in the literature. Warshel has argued that proteins cannot impose sufficient strain to change the ground state geometry of the substrate; however, ground state binding of the substate can lead to the enzyme being deformed (Warshel, 1981). Furthermore, it is the release of the "deformation energy" in the transition state that can lead to a reduction in $\Delta G^{\ddagger}_{T}$ (Warshel, 1981). It is clear from the original descriptions of the model (Blake et al., 1967; Imoto et al., 1972; Phillips, 1974) that a distorted sugar makes *favourable* contacts within site D that are unrealizable in an undistorted state. It should also be stressed and made clear that the model depicts an optimal placement of the distorted residue positioned deeply in the cleft so that the labile glycosidic bond is situated in a position that is compatible for the roles ascribed to the two catalytic acid residues (i.e. a catalytic positioning of residue D). In the lowest calculated energy conformation (the left-side structure) of the HEWL/$(GlcNAc)_6$ complex (Pincus & Sheraga, 1979) and in the crystal structure of GEWL/$(GlcNAc)_3$ (Weaver et al., 1995), the site D GlcNAc residue is removed from a deep position in the active site cleft and is not distorted.

In conclusion, the following is offered. It is evident that there may be more than one mode for a saccharide to bind to the active site cleft. The various processes detailed for the productive binding of $(GlcNAc)_6$ to HEWL (i.e. the β and γ processes; section 1.9.2) suggest that changes to the structure of the complex develop amid the events of initial binding of the substrate and the formation of a reactive complex. It is possible that the initial (or equilibrium) binding mode of the substrate with HEWL involves lower site positions of the enzyme that are somewhat removed from the cleft in addition to an undistorted site D residue. The calculated conformation involving the left side of the lower active site (i.e. the left-sided structure; section 1.7 and Fig. 1.21 A) may be representative of the initial binding mode. Conformational changes to this initial complex involving some (or all) of the saccharides in their respective subsites, but especially those in the D-F sites, could result in the movement of the substrate to the right side of the

active site and of the site D residue towards the catalytic position. Additional conformational changes to the complex could contribute to the distortion of the site D saccharide and this structure may be the one described in the Phillips model for hexasaccharide binding. The energy required for distortion of the ring may be provided by newly created interactions between the enzyme and substrate that have arisen during movement of the substrate deeper into the cleft. The distorted ring conformation is now very close to that required for the putative carbonium ion transition state and the enzyme interactions with the substrate are now able to stabilize the transition state during catalysis. The enzyme-induced stabilization of the carbonium ion allows diffusion of the aglycone from the active site and the entry of a new acceptor molecule into the cleft to react with the site D residue to complete the reaction. This description is in accord with transition state theory and the role of an enzyme as a catalyst in which part of the catalytic power of an enzyme is derived from its ability to bind the transition state more strongly.

Recently, an X-ray crystallography study was reported in which the different stages in the reaction pathway of haloalkane dehalogenase were trapped by varying the pH and temperature of the crystal (Verschueren et al., 1993). Usually, the lifetimes of reaction intermediates in enzymes is too short for crystallographic observation. However, this work is promising in that perhaps application of similar methodology may provide direct observation of the reaction intermediates which occur with lysozyme.

1.11. Biological Properties of Peptidoglycan and Muramyl Peptides

The bacterial peptidoglycan and muropeptides derived from peptidoglycan display a diverse spectrum of biological activity on mammalian organisms. Complete Freund's adjuvant (mycobacterial cell walls in a water/oil immulsion containing the antigen in the aqueous phase) is commonly used to increase the titer of antibodies against the antigen used. It was demonstrated that N-acetylmuramly-*L*-alanyl-*D*-isoglutamine (adjuvant peptide; muramyldipeptide; MDP; Fig. 1.29) could replace the mycobacterial cell walls in Freund's adjuvant and is the smallest common structural unit of the peptidoglycan possessing immunostimulatory properties (Elonz et al., 1974). Much interest in the chemistry and biological activity of MDP and related muramyl peptide analogues (eg. variations in the peptide and carbohydrate moieties, disaccharide compounds, 6-O-lipid derivatives) has arisen as these compounds have proven to improve the efficiency of vaccines, stimulate nonspecific resistance to bacterial infections and to act as possible

immunosuppressants and antitumour agents (for reviews see Dukor et al., 1979; Lederer, 1980; Lefrancier, 1981; Kotani et al., 1986; Bahr & Chedid, 1986; Lefrancier & Lederer, 1987; Baschang, 1989). MDP and peptidolglycan can also induce undesirable responses reproducing most of the major symptoms associated with bacterial infections such as fever (pyrogenic response), arthritis, meningitis, sleepiness, inflammation, macrophage and lymphocyte activation and cytokine secretion (Dziarski, 1986).

The exact biochemical mechanism of how the peptidoglycan products elicit their responses on hosts cells remains unknown. Some evidence has indicated that the processes are receptor-mediated involving a signal transduction mechanism of cell activation. Polymeric (or high molecular weight) peptidoglycan was shown to bind to the surfaces of mouse lymphocyte and macrophage cells (Dziarski, 1991a,b). The binding of peptidoglycan to the surfaces of these cells in both mouse and human cell lines was later shown to involve cell-bound albumin (Dziarski, 1994).

Among the peptidoglycan derived muropeptides, those containing *anh*MurNAc have also demonstrated biological activities. It was demonstrated that in sleep-deprived animals, substances accumulate in the spinal fluid, urine and brain and that these substances are active somnogens inducing slow-wave sleep (Krueger et al., 1980). Purification and structural characterization of the sleep-promoting factors from human urine had identified the anyhydro-disaccharide-tetrapeptide (G(Anh)MTetra) and the corresponding anhydro-muramyl-tripeptide (G(Anh)MTri) (Fig. 1.29) as the most potent somnogens (Martin et al., 1984; Krueger et al., 1984). Muropeptides that did not possess the 1,6-anhydro linkage were much less active somnogens. Furthermore, anhydro-muropeptides obtained directly from bacterial peptidoglycan and that were amidated at the α- and ε-carboxylic acid positions of the Glu and DAP moieties respectively were not somnogenic. These observations led to the suggestion that amidation/deamidation reactions in the brain could serve as a means to regulate the sleep-promoting activities (Krueger, 1984).

The current interest in the physiological effects of anhydro-muropeptides has prompted the development of methods to obtain these compounds in large amounts. Recently, G(Anh)MTetra and G(Anh)MTri as well as other anhydro-muropeptides have been isolated from the digestion of *E. coli* peptidoglycan with immobilized forms of the *E. coli* soluble lytic transglycosylase and D,D-endopeptidase (Engel et al., 1992b). No reports on the utilization of lambda lysozyme in the preparation of anhydro-muropeptides have appeared. The availability of this methodology has permitted further investigations into the activities of anhydro-muropeptides. Recent studies have indicated that G(Anh)MTetra

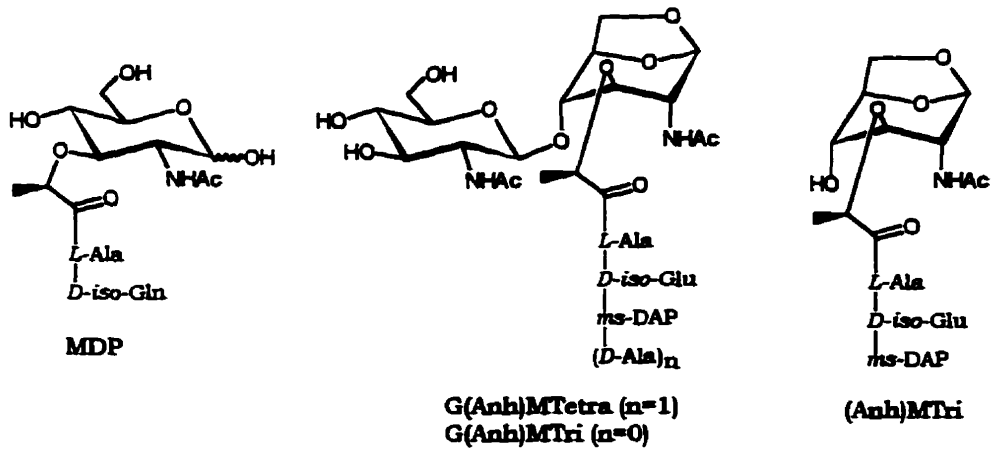


Figure. 1.29. Structures of some biologically active muramyl peptides.

induces interleukin (Dokter et al., 1994a) and granulocyte colony-stimulating factor (Dokter et al., 1994b) expression in human monocytes and that anhydro-muropeptides suppressed appetite and weight gain in rats (Biberstine & Rosenthal, 1995). In some cases, anhydro-muropetides released by some human pathogens (eg. *Bordetella pertussis*, *Neisseria gonorrhoeae*) have demonstrated cytotoxicity to epithelial cells in higher organisms including humans (Melly et al., 1984; Cookson et al., 1989).

The presence of peptidoglycan components in animals has been attributed to the catabolic products originating from the intestinal flora (Sen & Karnovsky, 1984). Evidence has surfaced suggesting that large molecular weight peptidoglycan is stored in the macrophages of human spleen (Hoijer et al., 1995). Others have also suggested that muramyl peptides might be stored in various tissues in mammals, perhaps as a modified inactive form, and then released into the sera where they exert their somnogenic, pyrogenic and immunomodulatory effects (Krueger et al., 1984). It is possible that mammalian lysozymes may participate in the regulation of the levels of the muramyl peptides and of higher molecular weight peptidoglycan present in the circulation. High levels of N-acetyl-muramyl-*L*-alanine amidase activity is found in many mammalian sera in addition to the saliva, milk, and spinal fluid of humans. The physiological role of the amidase was suggested to maintain the ratio between the biologically active muropeptides and their inactive, cleaved saccharide and peptide components in which the amidase would antagonize the production of biologically active molecules by lysozyme (Vanderwinkel et al., 1995). Given the widespread activities of peptidoglycan related

molecules in mammals, a novel concept has been proposed in which eukaryotic cells themselves are responsible for the endogenous synthesis of peptidoglycan, that the genes for peptidoglycan metabolism have been inherited from endosymbiotic bacteria and that endogenously synthesized muropeptides play a vital role in cell division and the biological clock (Roten & Karamata, 1992).

The biological significance of anhydro-muropeptides has also been established in bacteria. *E. coli* recycles components from approximately 40-50% of the murein sacculus each generation. As the peptidoglycan is degraded in the periplasm, the products are transported into the cytoplasm are subsequently reutilized for peptidoglycan biosynthesis (Goodell, 1985; Park, 1993). Recent evidence has suggested that a specific permease and cytosolic N-acetylmuramyl-*L*-alanine amidase exist in *E. coli* for the transport of (Anh)MTri (Fig. 1.29) into the cell and its cleavage into *anh*MurNAc and tripeptide respectively (Jacobs et al., 1994). The cytosolic accumulation of (Anh)MTri in *E. coli* has also been suggested to serve as a signal for β-lactamase induction (Jacobs et al., 1994).

1.12. General Objectives

While it is evident that lysozymes and in particular HEWL and T4L, have been characterized in great detail both in structure and mechanism, lambda lysozyme has not received similar attention. Several specific objectives were pursued in this study in order to further our understanding of the biochemical properties of LaL.

In order to initiate these studies, our primary objective was to establish an efficient expression system which would facilitate the the large scale isolation of LaL. In Chapter 2, our goals were to incorporate the λ *R* gene into a plasmid for the inducible expression of LaL in *E. coli* and to develop effective methodology for the purification of the enzyme in the quantities ultimately required to initiate crystallographic studies.

In Chapter 3, the introduction will highlight that lysozymes have evolved with remarkable differences in both their substrate specificities and abilities to interact with saccharide ligands. The ligand and substrate specificities of LaL are not understood. We have attempted to gain knowledge of the components of the peptidoglycan, namely the saccharide and peptide constituents, that will bind to LaL and that are critical for the activity of LaL. These questions have been addressed employing a variety of biophysical techniques and with the rational synthesis of possible substrates for LaL. An essential prerequisite for these studies was to establish a reliable and accurate approach to measure the activity of LaL.

We were interested in exploring the potential applications of incorporating a methionine analogue, trifluoromethionine into proteins and have chosen LaL as a model system for these purposes. Our objective was to develop methodolgy for the replacement of the three methionines in LaL (at positions 1, 14 and 107) with trifluoromethionine and to study the labelled protein by ^{19}F NMR. Presented in Chapter 4 is our commitment to introduce the experimental approaches necessary to achieve this goal and to address the consequences of trifluormethionine incorporation suggested from our characterization of lambda lysozyme labelled with this analogue.

CHAPTER 2

Overexpression, Purification and Characterization of Lambda Lysozyme

2.1. INTRODUCTION

Lambda lysozyme is produced in nature during the lytic cycle of phage λ and therefore, appropriately induced *E. coli* lysogens could serve as the source of the enzyme for purification purposes. Indeed, LaL has been purified to apparent homogeneity from λ lysogens (Black & Hogness, 1969a; Bienkowska-Szewczyk & Taylor, 1980). However, these strategies involved extremely tedious methodology and resulted with the recovery of less than 1 mg of protein per litre of lysogen. The biophysical and structural (including crystallographic) studies that we had envisioned with LaL required a source for the enzyme that would allow the isolation of LaL in yields higher than that achievable from a lysogen. With the advent of modern molecular biology techniques, the ability to isolate genes with their subsequent introduction into efficient expression systems in both prokaryotes and eukaryotes has greatly reduced the requirement to purify proteins from their native sources. Our goal was foremost to incorporate the λ *R* gene into an appropriate plasmid for its overexpression in *E. coli*, as well as for potential genetic manipulations to produce mutant lysozymes.

Our studies first required the overexpression of LaL (to any level) to allow for the development of a purification strategy that might improve upon the more elaborate protocols previously reported by others (Black & Hogness, 1969a; Bienkowska-Szewczyk & Taylor, 1980). Attempts to clone the *R* gene from λ DNA were initiated in the laboratory of Dr. E. Daub at the University of Guelph prior to the onset of this work. As we were primarily interested in developing a system that would allow for the recovery of the greatest amount of LaL, several expression systems were developed since these initial efforts and examined for their capacity to produce LaL. As we became able to improve upon the overexpression of LaL, the purification strategy was continually modified to accommodate the increased amounts of protein. Purification protocols were not only revised to obtain maximal protein recovery, but also to become facile and as efficient as possible. The contributions of Robert Ruman, Martha Cox, Denise Miedema and especially those of Dr. E. Daub in the preparation of expression plasmids is appreciatively noted.

During the course of this study and following the development of our current purification methodology, others have also detailed the overexpression and purification of LaL with yields of 11.5 mg/L (Jespers et al., 1991). Our current expression system (*E. coli*

BL21(λDE3)/pLR102) and purification methodology readily permits one individual to isolate 150 mg of purified protein over the course of 3-4 days at a level of 25-30 mg per litre of culture.

2.2. EXPERIMENTAL

2.2.1. General Experimental

Preparation of all growth media and buffers used for protein purification and DNA manipulations are described in Appendix A. Optical spectroscopy was performed on a Varian Cary 3 UV-Vis Spectrometer. Culture growth was performed in either a Gyrotory® G76 Water Bath Shaker (New Brunswick Scientific) for small cultures (< 200 mL) or in a Series 25 Controlled Environment Incubator Shaker (New Brunswick Scientific) for larger cultures. Cells were grown aerobically by agitation of test tubes or flasks at 180 rpm. Low speed preparative centrifugation was performed at 2-4 °C in a Beckman J2-21 instrument using appropriate rotors (JA-10, JA-14 or JA-20) dependent on the volume demand of the sample. Small volume centrifugation was performed with a BioFuge A table top microcentrifuge or an IEC Clinical Centrifuge.

2.2.2. Materials

All materials used were of the highest grade commercially available and used without further purification. Reagents used in the preparation of growth media and buffers are given in Appendix A. Formaldehyde solution (37-40%, AnalaR), phosphorous pentoxide (P_2O_5), potassium carbonate (K_2CO_3), potassium chloride (KCl, AnalaR grade) sodium acetate (CH_3CO_2Na), sodium carbonate (Na_2CO_3), sodium hydrogen carbonate ($NaHCO_3$) and sodium thiosulphate ($Na_2S_2O_3$) were obtained from BDH (Toronto, ON) and were of Assurance grade except where noted. Ultra pure grades of ammonium sulfate ($(NH_4)_2SO_4$) and guanidine hydrochloride (Gdn·HCl) were purchased from ICN Biomedicals (Montreal, PQ). Albumin (bovine serum), ampicillin, bromophenol blue, Coomassie Brilliant Blue R-250 and Blue G-250, glutaraldehyde solution (50%, Grade I), 2-mercaptoethanol, silver nitrate ($AgNO_3$) and streptomycin sulfate were obtained from Sigma (St. Louis, MO). Ammonium persulfate and N,N,N',N'-tetra-methylethylenediamine (TEMED) were obtained from Bio-Rad (Mississauga, ON). Ethidium bromide, phenylmethylsulfonyl fluoride (PMSF), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and ultra-pure urea were products from Boehringer Manheim (Montreal, PQ). Electrophoretic grade acrylamide, agarose, bisacrylamide, glycine and sodium

dodecylsulfate (SDS) were obtained from either ICN Biomedicals or Bio-Rad while electrophoretic grade Tris-HCl and Tris-base were obtained from both Bio-Rad and Boehringer Manheim. Glycerol was obtained from Fisher Scientific (Don Mills, ON). Dithiothreitol (DTT) and isopropyl-β-D-thiogalactopyranoside (IPTG) were from Diagnostic Chemicals (Charlottetown, PEI). Syringe filter units (0.22 μm and 0.45 μM, PVDF) were products of Chromatographic Specialties (Brockville, ON.). Schleicher and Schuell (Keene, NH) provided Elutip-D columns and NA45 membranes while the Prep-A-Gene® Miniprep Kit is a product of Bio-Rad. Restriction endonucleases, T4 DNA ligase, *Bam*HI linkers, λ DNA, RNase A and reagents for performing PCR were purchased from either New England Biolabs (Mississauga, ON.) or Boehringer Mannheim. DNA standards were also obtained from Bio-Rad. Spectra/Por® 3 and Spectra/Por® 7 dialysis tubing (MWCO 3500) were obtained from Canadawide Scientific (Ottawa, ON). Other details concerning materials and equipment are given when necessary in the text.

2.2.3. Plasmids and Bacterial Strains

The bacterial strain *E. coli* BL21(λDE3) and plasmid pET-22b were purchased from Novagen while plasmid pTTQ18 was obtained from Amersham. Dr. Reggie Lo from the University of Guelph provided the synthetic primers used for the polymerase chain reaction (PCR) and the generous gifts of replicative form (RF) M13 mp18 and mp19 and *E. coli* strain TG-1. Plasmid pCM101 and *E. coli* OR1265 were brought by Dr. E. Daub from the laboratory of Dr. H. Murialdo (Dept. of Medical Genetics, University of Toronto, ON).

All *E. coli* strains were maintained in 50% glycerol suspensions (free of antibiotic) and stored at -80 °C. Strains maintained on buffered agar plates were kept at 4 °C and restreaked on 10-20 day intervals.

2.2.4. General Recombinant DNA Techniques

All standard DNA manipulations were performed based on general methodology described in various sources (Maniatis et al., 1982; Dillon, 1985; Ausubel et al., 1989; Sambrook et al., 1989; Ausubel et al., 1991). Ampicillin was added at a concentration of 40 $\mu g \cdot ml^{-1}$ for liquid medium and at 50 $\mu g \cdot ml^{-1}$ for agar plates.

i) Plasmid Isolation. Typically, the *E. coli* hosts (3-5 mL) were grown in LB_{amp} (LB medium supplemented with 40 $\mu g \cdot ml^{-1}$ ampicillin) to mid- or late-exponential growth. Cells were pelleted in a microcentrifuge tube, suspended in STET buffer (0.5 mL) and HEWL (35 μL of

a 10 mg/mL solution) was added. The sample was vortexed, heated at 95 °C for 3 min and then centrifuged (13000 rpm, 5 min, Biofuge A). Chromosomal DNA was removed using a sterile toothpick and isopropanol (0.5 mL) was added. The sample was vortexed, centrifuged (13000 rpm, 5 min) and the supernatant removed. The pellet was air dried and resuspended in TE buffer (100 μL) and then was treated with RNase (6 μL of 10 mg/mL solution) for 10 min at 37 °C. EDTA (4 μL of 0.5 M solution) and NaCl (6 μL of 5 M solution) were added. The sample was extracted successively with 1 volume each of TE saturated phenol, 1:1 TE saturated phenol:$CHCl_3$, and $CHCl_3$ and the top aqueous layer from each extraction was removed. The recovered solution was made to 0.1 M NaCl (using 5 M NaCl) and 2 volumes of ethanol were added. The plasmid DNA was precipitated (at -20 °C or lower) for several hours and pelleted (13000 rpm on Biofuge A, 15 min, 4 °C). The pellet was washed with 70% ethanol (cooled at -20 °C, 150 μL), pelleted and air dried. The purified plasmid DNA was then suspended in TE buffer (100-150 μL) and stored at 4 °C or frozen at -20 °C for longer periods of storage.

Alternatively, plasmid DNA was isolated using the Bio-Rad Prep-A-Gene® DNA Miniprep kit following the manufacturer's protocol. The method is based on alkaline lysis of cells followed by adsorption and subsequent elution of plasmid DNA from a binding matrix.

DNA was quantified by measuring the absorbance at 260 nm assuming an absorbance of 1.0 for a concentration of 50 $\mu g \cdot mL^{-1}$ for double stranded DNA, 40 $\mu g \cdot mL^{-1}$ for single stranded DNA, and 20 $\mu g \cdot mL^{-1}$ for oligonucleotides (Sambrook et al., 1989). In some instances, DNA concentration was estimated by comparing the ethidium bromide mediated fluorescence of the unknown sample with that of DNA standards of known concentration after electrophoresis on agarose gels.

ii) Restriction Digestion and Analysis. Plasmid DNA was routinely digested with restriction enzymes to allow characterization of constructs and isolation of linear DNA. The size of all constructs investigated and of the inserts of interest were such that 1-2 μg of plasmid was sufficient for ethidium bromide mediated fluorescence visualization on gels. Typically, plasmid DNA (1-2 μg) was singly or doubly digested with the selected restriction enzyme(s) (1-2 units; 1 unit will digest 1 μg of DNA in 1 hr) in 1× of the appropriate (as suggested by the suppliers) restriction buffer (10 μL). Digestions were conducted for 2-4 hr at 37 °C. Linearized DNA was isolated by phenol/$CHCl_3$ extraction and ethanol precipitation (as described in (*i*)) or on agarose gels as described below (*iv*). For restriction anaylsis, DNA loading buffer (10×, 1 μL) was added to the digests and samples were electrophoresed.

iii) Agarose Gel Electrophoresis. DNA was routinely separated by horizontal electrophoresis using a Bio-Rad DNA Sub Cell™ and Model 200/2.0 power supply. Gels were prepared by the addition of agarose (0.8-1.2% (w/v)) to 1× TBE buffer and then boiled until dissolution of the agarose was complete. The molten agarose was cooled (to ≈50 °C) and cast onto glass plates fitted with a well-forming comb. The gel was allowed to solidify, positioned into the electrophoresis unit and submerged in running buffer (1× TBE). Samples, including DNA markers, were loaded and typically electrophoresed at 20-100 V for 3-12 hr. Gels were stained in 0.5 $\mu g \cdot ml^{-1}$ ethidium bromide for 10-30 min, destained as necessary in DDW and visualized by transillumination.

iv) Isolation of Linear DNA. Desired DNA fragments obtained from restriction of plasmids or λ DNA were either partially purified by phenol/$CHCl_3$ extraction and ethanol precipitation (as described in (*i*)) or purified from agarose gels as follows. The DNA sample was digested with the appropriate enzyme(s) and electrophoresed as described above (*ii* and *iii*) except that the agarose gel contained ethidium bromide (0.5 $\mu g \cdot ml^{-1}$). Following completion of electrophoresis, the DNA bands were visualized using a longwave (365 nm) UV light source and a slit was cut into the gel immediately in front of the DNA to be recovered. A small strip of NA45 membrane (DEAE cellulose) was placed into the slit and electrophoresis was reinitiated at 50 V for 10-15 min. The strip was removed and visualized by longwave UV to ensure elution of the band onto the membrane. Recovery of the DNA from the membrane was achieved using Elutip-D columns and buffers (Schleicher and Schull) as described by the manufacturer followed by ethanol precipitation. The purified DNA was suspended in an appropriate volume of TE buffer (10-20 μL) for subsequent use.

DNA fragments produced during PCR (≈ 500 bp) were run on NuSieve GTG Lo-melt agarose (2%) in 1× TBE buffer. Following electrophoresis, gels were stained in ethidium bromide, visualized with a longwave UV light source and the desired DNA fragment cut from the agarose. The agarose fragment was melted at 60 °C for 15 min and then maintained at 40 °C for subsequent removal of aliquots used during ligations.

v) Ligations and Transformations

Linearized plasmid DNA and insert DNA digested with the desired endonuclease(s) were isolated as described above. Ligation mixtures were prepared containing various ratios of linearized plasmid and insert DNA (total of 8 μL) and 1 μL each of T4 DNA ligase and 10× ligation buffer and the mixtures were incubated at 14 °C overnight.

Competent *E. coli* cells were prepared by the $CaCl_2$ treatment method (Cohen et al., 1972) under aseptic conditions. Cells were grown in LB medium (400 mL) to an $OD_{600} \approx$ 0.4 at 37 °C. The cells were cooled on ice and collected by centrifugation (1600 × g, 7 min, 4 °C). Cell pellets were gently resuspended in $CaCl_2$ buffer (80 mL, Appendix A) and then pelleted by centrifugation (1000 × g, 5 min, 4 °C). The cells were then treated with $CaCl_2$ buffer (80 mL) on ice for 30 min and centrifuged as just described. The cells were resuspended in $CaCl_2$ buffer (8 mL), dispensed into 1 mL aliqouts and stored at -80 °C.

Plasmid DNA aliquots (as prepared in section (*i*) chilled on ice, aliquots of 0.5-4.0 μL, ca. 10-500 ng) or ligation mixtures (10 μL) were added separately to sterile test tubes on ice. Competent cells were rapidly thawed and 100 μL was added to each DNA aliquot and kept on ice for 15 min. Transformation mixtures were heat shocked at 42 °C for 2 min, prewarmed LB or SOC medium (1 mL) was added and samples were incubated at 37 °C for 1-2 hr. Appropriate dilutions of the cells were plated onto LB_{amp} plates (LB plates supplemented with 50 $ug \cdot ml^{-1}$ ampicillin), allowed to soak into the agar plate and then incubated overnight at 37 °C. Selected transformants were screened by restriction anaylsis as described above or by activity measurements (section 2.2.6).

2.2.5. Cloning of the λ *R* Gene and Construction of Expression Vectors

2.2.5.1. E. coli DH5α /pLR1

Lambda DNA (cI_{ss7}, Sam_7) was digested with *Eco*RI and the 3.53 kb band containing the complete *S*, *R* and *Rz* genes was isolated from 0.8% agarose. This fragment was then digested with *Cla*I and the 1.47 kb *Eco*RI-*Cla*I fragment was isolated from 0.8% agarose and ligated into partially purified (phenol/$CHCl_3$ extraction and ethanol precipitation) pBR322 that was digested with *Eco*RI and *Cla*I. The ligation reaction was used to transform *E. coli* DH5α and the recombinant plasmid containing the *R* gene was named pLR1.

2.2.5.2. E. coli TG-1/pLR7

Partially purified plasmid pLR1 was digested with *Eco*RI and then treated with *Bal*31. The exonuclease activity of *Bal*31 will progressively remove nucleotides successively from both 3' and 5' termini of duplex DNA, and the amount of DNA removed is dependent on duration and conditions of incubation. This step was done to generate a series of deletions of the λ DNA insert in an attempt to remove the *S* gene. As such, linearized pLR1 was treated with *Bal*31 and at various times (from 1-60 min) aliquots were removed and the reactions terminated with the addition of EGTA (to 20 mM). The digests

were partially purified by phenol/$CHCl_3$ extraction and ethanol precipitation, treated with *Hind*III and partially purified. The '*S* deletion inserts were ligated with pTTQ18 digested with *Hind*III and *Sma*I (a blunt end generating endonuclease) and used to transform *E. coli* DH5α. Plasmid DNA isolated from individual ampicillin-resistant colonies were analyzed by restriction analysis to confirm the presence of the insert band.

Plasmid DNA from several positive transformants was isolated and used to transform *E. coli* TG-1. Transformants were screened for LaL activity as described in section 2.2.6. One such transformant which demonstrated the highest relative activity was chosen for further analysis and the recombinant plasmid was named pLR7.

2.2.5.2.1 Sequence Determination of the pLR7 Insert

The sequencing of the insert from pLR7 was performed to determine the extent of deletion to the *S* gene and to confirm the presence of the *R* gene.

i) Preparation of Single-Stranded M13 DNA. The 1.14 kb *Eco*RI/*Hind*III insert from pLR7 was purified from 1.2% agarose and ligated into partially purified M13 mp18 and mp19 RF (Replicative Form) DNA digested with the same enzymes. The ligation mixtures (5 μL) were each added to *E. coli* TG-1 competent cells (0.2 mL) and maintained on ice for 45 min. Samples were heat shocked (42 °C, 2 min) after which was added early exponentially growing TG-1 cells (grown in LT medium, 0.3 mL), 0.1 M IPTG (10 μL), 2% X-gal in DMF (30 μL) and soft agar (0.6% agar in LT, 45 °C, 3-4 mL). The mixtures were poured onto prewarmed LT plates (1.5% agar), allowed to solidify and incubated at 37 °C.

A single white plaque from each transformation was added to LT medium (10 mL) that had been innoculated with *E. coli* TG-1 cells (0.1 mL) from an overnight culture grown in M9 minimal medium supplemented with glucose (0.3%) and thiamine (0.005%). The strain TG-1 (Δ*proAB-lac*) is grown on minimal media (lacking proline) in order to maintain the F' factor ($proA^+B^+$-lac^+) which is required for phage infection. Cultures were grown at 37 °C for 4.5 hr and then centrifuged (12000 × g, 4 °C, 5 min). The supernatants were reserved and the phage particles were precipitated with the addition of 5× PEG/NaCl solution (2.5 mL) on ice for 30 min. The phage precipitate was collected by centrifugation (17500 × g, 4 °C, 10 min) and the supernatant was thoroughly removed from the pellet (drain inverted tube on tissue paper and wipe inside of tube). The phage particles were resuspended in phage buffer (0.6 mL) by drawing in a pasteur pipette and subjected to phenol/$CHCl_3$ extraction (as described in (*i*)). The recovered solutions following extraction (≈ 0.4 mL) were made to 0.3 M sodium acetate (by addition from a 3 M solution) and then 2 volumes of cold (-20 °C) ethanol was added. The single stranded DNA was

precipitated at -20 °C overnight, pelleted (13000 rpm on Biofuge A, 4 °C, 15 min) and resuspended in TE buffer (40 μL) and stored at -20 °C.

ii) Sequence Reactions. Sequencing of the recombinant single stranded mp18 and mp19 was performed using the M13 Sequencing Primer (5' d(GTAAAACGACGGCCAGT)) and T7 DNA Sequencing Kit from BioCan Scientific by the dideoxy, chain termination method.

Annealing. Single-stranded M13 template DNA (≈ 0.8 pmol) and primer DNA (2 pmol) were prepared in 1× sequencing buffer (10 μL). Samples were heated at 65 °C for 2 min and then allowed to cool slowly to room temperature. *Labelling.* To the reactions was added labelling mixture (2 μL), ^{32}P-dCTP (1 μL) and T7 DNA polymerase (3 u, 2 μL) and reactions allowed to proceed at room temperature for 2.5-3.0 min. *Termination.* Aliquots of the labelling reaction (3.5 μL) were each added to separate solutions of each ddNTP (2.5 μL), incubated at 37 °C for 5 min and then reactions were ended with the addition of stop solution (4 μL) and stored at -20 °C.

Electrophoresis was performed on 7% acrylamide/50% urea prepared as follows:

20 mL	5 x TBE buffer	- mix reagents with heating until dissolved
50 g	Urea	- cool and add DDW to give a total volume of 100 mL
6.65 g	Acrylamide	- filter through a sterilization unit (0.22 μM)
0.35 g	Bis-acrylamide	- initiate polymerization with addition of 10% ammonium persulfate (0.5 mL) and TEMED (50 μL)

The gel was cast into a Sequi-Gen® Sequencing Cell unit (Bio-Rad) equipped with a Model 3000xi power supply. Sequence reactions were heated to 75 °C for 2 min prior to loading onto the gel. Gels were run at constant power (55-75 watts) for 2-3 hr in 1× TBE buffer. The gel was transferred onto a sheet of Whatmann paper, dried and exposed to X-ray film (Kodak) at -20 °C. Sequences were read following development of the film.

2.2.5.3. E. coli OR1265/pLclR18

Partially purified plasmid pLR1 was digested with *Eco*RI and *Hind*III and the 1.47 kb fragment was isolated from 1.2% agarose. The fragment was further digested with *Hin*fI which cleaves 55 bp 5' to and 7 bp 3' to the *R* gene. The *Hin*fI ends were filled in using Klenow fragment and dNTP's and *Bam*HI linkers (5'd(GGGATCCC)) were blunt end ligated using T4 DNA ligase. After cleavage of the linkers with *Bam*HI the 546 bp *R* gene fragment was purified by agarose gel electrophoresis. Plasmid pCM101 (Chow et al., 1987) was digested with *Bss*HII and partially purified. The *Bss*HII overhang was filled in and

*Bam*HI linkers introduced as described above. The linearized pCM101 was digested with *Bam*HI and treated with alkaline phosphatase. The ca. 5.0 kb fragment was purified by agarose gel electrophoresis and ligated with the 546 bp *R* gene fragment. The ligation reaction was used to transform *E. coli* DH5α and plasmid DNA isolated from individual ampicillin-resistant colonies were subjected to restriction analyis to confirm the presence of the insert band. Plasmid DNA from several positive transformants was isolated, partially purified by phenol/$CHCl_3$ extraction and ethanol precipitation and used to transform *E. coli* OR1265. Transformants were screened for LaL activity and expression as described in sections 2.2.6 and 2.2.7. One transformant was selected that exhibited LaL activity and the recombinant plasmid was called pLcIR18.

2.2.5.4. E. coli TG-1/pHDM10

The 546 bp *Bam*HI *R* gene fragment described above in the construction of pLcIR18 was ligated with partially purified pTTQ18 digested with *Bam*HI. The ligation reaction was used to transform *E. coli* DH5α and plasmid DNA isolated from individual ampicillin-resistant colonies were analyzed by restriction analysis to confirm the presence of the insert DNA. Plasmid DNA from several positive transformants was isolated, partially purified by phenol/$CHCl_3$ extraction and ethanol precipitation and used to transform *E. coli* TG-1. Transformants obtained (32) were screened for LaL activity and expression as described in sections 2.2.6 and 2.2.7. Four transformants with the 546 bp *Bam*HI fragment which exhibited strong and similar lysoyzme activity and expression levels were obtained. The one which was chosen for further study contained the recombinant plasmid that was named pHDM10.

2.2.5.5. E. coli BL21 (λDE3)/pLR102

PCR was employed to amplify the *R* gene and the template used for the reaction was plasmid pHDM10. The synthetic oligonucleotide primers used were:

KpnI *NdeI*
primer 1) 5'-CCA GGTACC CATATG GTAGAAATCAATAATCAA-3' (33 mer) and

HindIII *BamHI* Stop
primer 2) 5'-CCA AAGCTT GGATCC TCATACATCAATCTCTCTGAC-3' (36 mer).

Primers were designed to introduce a 5' *KpnI* restriction site and a *NdeI* restriction site spanning the ATG start codon (primer 1) and a *Bam*HI and *Hind*III restriction sites immediately 3' to the termination codon (primer 2) as indicated above. Sequence complimentary to the LaL sense (primer 1) and antisense (primer 2) strands are

underlined. PCR was performed using a MiniCycler™ thermal unit (MJ Research, INC.). Reaction mixtures were prepared in 80 μL Vent Buffer containing 1 U Vent$_R$ DNA polymerase, 200 μM of each dNTP and with various concentrations of template DNA (either approx. 23 or 56 pg·μL^{-1}), primers (either approx. 16 or 40 μM) and $MgSO_4$ (either 2, 4 or 6 mM). Solutions were topped with parafilm oil to prevent evaporation of the reaction solutions. The samples were allowed to undergo 30 cycles of denaturation (90 s at 94 °C), annealing (60 s at 55 °C) and extension (30 s at 72 °C). Following completion of the reactions, the amplified DNA fragment was partially purified by phenol/$CHCl_3$ extraction and ethanol precipitation. An aliquot of each sample was analyzed on 1.5% agarose to identify successful reactions. The reaction mixture which was successful and that was chosen for subsequent manipulations contained 4.5 ng template DNA (56 pg·μL^{-1}), 40 μM of each primer and 2 mM $MgSO_4$.

The partially purified 504 bp PCR product was treated with *Kpn*I and *Hind*III and the digested DNA fragment was electrophoresed on NuSieve GTG Lo-melt agarose (2%) and isolated from the gel. The fragment was then ligated into plasmid pUC18 digested with *Kpn*I and *Hind*III. The ligation reaction was used to transform *E. coli* DH5α and plated onto LB_{amp} plates containing X-gal. Plasmid DNA isolated from several faint blue to white recombinant colonies were screened by restriction analysis to confirm the presence of the PCR product insert band. Recombinant pUC18 DNA was partially purified from 4 positive transformants and the insert was excised with *Nde*I and *Bam*HI and purified on Lo-melt agarose. The 4 samples of purified DNA were each ligated with plasmid pET-22b that had been previously digested with *Nde*I and *Bam*HI and purified on 0.8% agarose. The ligation reactions were used to transform *E. coli* DH5α and plasmid DNA isolated from 6 ampicillin resistant transformants from each ligation were subjected to restriction analysis. Of the 24 transformants, only one possessed recombinant pET-22b with the correct size insert and plasmid DNA was partially purified from this transformant and used to transform *E. coli* BL21(λDE3). Transformants were assayed for LaL activity and expression as described in section 2.2.6 and 2.2.7. Two transformants obtained exhibited strong activity and expression of LaL and the recombinant plasmids were named pLR102 and pLR103. Plasmid pLR102 was chosen for further study.

Characterization of pLR102. Plasmid pLR102 was isolated and subjected to restriction analysis (*Nde*I/*Bam*HI) and electrophoresed on 1.2% agarose. An estimated size of 0.51 kb was observed for the *Nde*I/*Bam*HI fragment, which has a calculated size of 479 bp. The lysozyme gene was sequenced at the Central Facility of the Institute for Molecular Biology

and Biotechnology (McMaster University, Hamilton, ON) in both directions from pLR102 which revealed that no mutations had occurred during the PCR reaction.

2.2.6. Activity Determinations - Turbidimetric Assay

The activity of LaL was assessed turbidimetrically using chloroform sensitized *E. coli* substrate cells. Details concerning the methodology for the preparation of the bacterial substrate and of the assay are outlined in Chapter 3. Briefly, the procedure involves monitoring the bacteriolytic activity of LaL by measuring the decrease in absorbance (clearing) of a suspension of the substrate cells following the addition of an assay sample. In essence, as LaL degrades the peptidoglycan of the substrate cells, the cells lyse which results in the clearing of the suspension.

i) Typical Assay. A 0.5 mL aliquot of frozen substrate cells was thawed and resuspended in 10 mL of 50 mM KPB, pH 7.0 (to an $OD_{600} \approx 0.7$) and maintained on ice. The sample to be assayed (5-100 µL) was added to 800 µL of the cell suspension that had been warmed to 25 °C in a 1 mL cuvette and the absorbance at 600 nm was monitored. Relative activity was assessed by the rate of decrease in aborbance. Substrate cells that were maintained on ice in suspension for greater than 1 hr were discarded and a fresh aliquot was thawed and prepared for use.

ii) Activity Screening of Transformants. Transformants obtained during construction of expression plasmids were screened as follows. An overnight culture of individual transformants grown in LB_{amp} were diluted 1:50 into the same medium. Cultures were incubated with shaking at 37 °C for 2-3 hr and then induced with 0.5-0.75 mM IPTG (by the addition from a sterile 100 mM solution) and incubated for an additional 2-3 hr as desired. Screening of recombinant *E. coli* OR1265/pLcIR transformants was achieved as described except that cultures were grown at 32 °C, induced by incubation 45 °C for 15 min and further incubated at 40 °C for 1-2 hr. Following growth and induction, cultures were diluted to an OD_{600} of 0.5-0.9 (if necessary) with DDW. Cells were lysed with the addition of $CHCl_3$ (2-3 drops) and 0.1% SDS (1-2 drops) to 2 mL of each (diluted) culture. Samples were vortexed, and allowed to remain at 25 °C until lysis of the cells was visibly detectable. If lysis was not achieved, cells from above were first pelleted and resuspended in 5 mM KPB, pH 7.0 and treated with $CHCl_3$ and SDS as described. An aliquot (50-100 µL) of the lysed sample was removed and assayed turbidimetrically as described above (*i*).

iii) Assays During Protein Purification. Activity determinations of fractions obtained during protein purification were achieved as described above (*i*). Smaller aliquots (5-40 μL) were typically used in the assay of fractions which contained high concentrations of salt since the lysis of the bacterial substrate cells and therefore, evaluation of LaL activity, is decreased in solutions with increasing ionic strength (detailed in Chapter 3, section 3.3.1.2).

2.2.7. SDS-PAGE Determination of LaL Expression Levels

Cultures prepared from strains or transformants harboring recombinant plasmids were grown pre- and post-induction as described above (2.2.6 (*ii*)). Cultures (1-2 mL) were centrifuged (Biofuge A, 8000 rpm, 5 min) and the supernatant was carefully removed. The cell pellets were suspended in protein loading buffer (one-tenth of the original culture volume), vortexed and placed in boiling water for 5 min. Insoluble material was pelleted (Biofuge A, 13000 rpm, 2 min) and an aliquot was removed and subjected to SDS-PAGE analysis and Coomassie staining as described below (section 2.2.8). The LaL expression level in recombinant strains or transformants was qualitatively assessed by the relative intensity of the band corresponding to LaL which migrated with an apparent molecular weight of approximately 18 kDa.

2.2.8. General SDS-PAGE

Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) was routinely performed on a PhastSystem® (Pharmacia) using commercially available PhastGel precast gels and PhastGel SDS buffer strips as described by the manufacturer. In some cases SDS-PAGE was performed by the discontinuous method (Laemmli, 1970) on a Mini-PROTEAN II electrophoresis unit (Bio-Rad) in 0.75 mm polyacrylamide slab gels. Separating gel consisted of either 12% T (11.68% acrylamide and 0.32% bis-acrylamide) or 15% T (14.6% acrylamide and 0.4% bis-acrylamide), 0.1% SDS in 0.375 M Tris-HCl, pH 8.8. Polymerization was initiated with the addition of ammonium persulfate (from a 10% w/v solution) and TEMED to 0.05% each. The separating gel was overlain with butanol saturated with MQW in order to produce a level top surface. Stacking gel consisted of 3.9% T (3.796% acrylamide and 0.104% bis-acrylamide), 0.1% SDS in 0.125 M Tris-HCl, pH 6.8. Polymerization was initiated with the addition of ammonium persulfate to 0.05% and TEMED to 0.1%. The gel solutions were degassed by aspiration prior to initiation of polymerization. Running buffer (pH 8.3) contained Tris-base (3 g/L), glycine (14.4 g/L) and SDS (1 g/L). Samples were electrophoresed at 100 V through the stacking gel and 150 V through the separating gel.

Samples for SDS-PAGE were mixed with one volume of protein loading buffer (Appendix A) and placed in boiling water for 5 min. Low molecular weight protein standards were from Pharmacia. Each vial contained rabbit muscle phosphorylase *b* (94 kDa, 64 μg), bovine serum albumin (67 kDa, 83 μg), egg white ovalbumin (43 kDa, 147 μg), bovine erythrocyte carbonic anhydrase (30 kDa, 83 μg), soybean trypsin inhibitor (20.1 kDa, 80 μg) and bovine milk α-lactalbumin (14.4 kDa, 121 μg). Standards were reconstituted in 150 μL loading buffer, boiled for 5 min and used directly for Coomassie stained gels or diluted 1:25 with loading buffer for silver stained gels. Typically 5-10 μL of the protein standards were loaded for slab gels.

PhastGels were developed by staining with Coomassie brilliant blue R-250 (0.1% w/v in 30% methanol, 10% acetic acid) as described in the PhastSystem manual. Slab gels were stained in 0.1% (w/v) Coomassie brilliant blue R-250 in 40% methanol, 10% acetic acid for 2-4 hr with gentle shaking followed by destaining in 40% methanol, 10% acetic acid as required. Slab gels were preserved in 10% glycerol, 5% acetic acid. Silver staining of PhastSystem gels and of slab gels was performed based on the methodology of Heukeshoven & Dernick (1988) as outlined in Table 2.1.

2.2.9. Protein Purification

2.2.9.1. General Purification Procedures

All manipulations were performed at 4 °C except for precipitation steps (performed on ice) and FPLC (Fast Protein Liquid Chromatography) which was performed at ambient temperature. Absorbance measurements of cultures to assess cell growth were obtained utilizing LB medium as the reference solution. Following completion of the growth of cultures for purification purposes, the cultures were kept on ice during the time period required to harvest the cells by centrifugation. Disruption of cells was performed using a French Press cylinder (cooled at 4 °C prior to use) and a manual Carver Laboratory press operating at 8000-10000 psi. When noted, cells (cooled on ice) were disrupted by sonication with a SONICATOR™ W-225 (Heat Systems-Ultrasonics, Inc.) at output control level 7 and 70% duty cycle. In all cases immediately following cell disruption, PMSF (dissolved in a minimal volume of reagent grade acetone) was added to 1 mM. Lyophilization was performed at -50 °C and at a reduced pressure of 0.1-0.5 mM of Hg using a Lab Con Co. Freeze Dryer-3. Samples for lyophilization were rapidly frozen in either a dry ice/acetone bath or in liquid N_2.

FPLC was performed on a Pharmacia system (LCC-500 Controller, two P-500 Pumps, MV-7 Motor Valve, UV-M Monitor, REC-482 Chart Recorder) monitoring at a

Table 2.1. Silver staining of protein gels.

PhastSystem Gels				Manual Staining of Slab Gels		
Step No.	Solution (100 mL each)	Time (min)	Temp. (°C)	Step No.	Solution (150 mL each)	Time (min)
1.	Solution 1: 30% ethanol, 10% acetic acid	4	50	1.	20% Methanol 10% acetic acid	60
2.	Repeat with 1	4	50	2.	Solution 1	15
3.	Solution 2: 0.5% glutaraldehyde, 0.1% $Na_2S_2O_3$ in 30% ethanol/0.4 M sodium acetate, pH 6.0	8	40	3.	Repeat with 1	15
				4.	Solution 2	15
				5.	MQW	2
				6.	MQW	2
				7.	MQW	2
4.	MQW	3	50	8.	MQW	5
5.	MQW	3	50	9.	Solution 3	20
6.	MQW	3	50	10.	MQW	1
7.	MQW	4	50	11.	Solution 4	<10 sec.
8.	MQW	4	50	12.	Repeat with 4	2-20 as needed
9.	Solution 3: 0.1% $AgNO_3$ containing 25 µL/100 mL of formaldehyde solution†	10	40	13.	10% acetic acid	2
				14.	MQW	2
				15.	10% acetic acid, 10% glycerol	30
10.	Solution 4: 2.5% Na_2CO_3 containing 40 µL/100 mL of formaldehyde solution†	0.4	30			
11.	Repeat with 4	6	30			
12.	5% acetic acid	2	50			
13.	MQW	3	50			
14.	20% glycerol	3	50			
15.	MQW	0.1	40			

All solutions were prepared freshly before use with MQW. Manual staining was performed at room temperature with gentle shaking. †Commercial preparation is 37-40%.

wavelength of 280 nm using commercially available pre-packed columns from Pharmacia (Mono-S HR 5/5, HR 10/10; Phenyl-Superose HR 5/5, HR 10/10; Superdex 75 HR 10/30). The attenuation of the monitor and chart recorder were adjusted to suit the required sensitivity of individual runs. All chromatographic buffers were prepared with Milli-Q water (MQW) and were filtered (0.22 or 0.45 µM) and degassed by aspiration prior to use. Samples for FPLC were filtered using syringe filter units (0.22 or 0.45 µM) immediately before application to columns. Chromatography employing S-Sepharose Fast Flow chromatography resin (Pharmacia) was prepared as follows. The resin was thoroughly washed with MQW, 20% ethanol, MQW and then with equilibration buffer.

The resin was suspended in two parts of equilibration buffer and degassed by aspiration with gentle manual agitation for 20 min. The slurry was then packed into Pharmacia glass columns (HR 5/5, HR 10/10, HR 10/30 or HR 16/50) by gravity flow until half the column was packed and then with mild external pressure using a forced air line to pack the remaining volume of the column. Columns were so packed to give the suggested operating pressure as described by the manufacturer. HPLC (high pressure liquid chromatography) was performed on a Waters 625 LC unit equipped with a Waters 994 photodiode detector and 5200 Printer Plotter using appropriate connector fittings to allow use of Pharmacia columns when necessary.

When total protein quantities are cited, these were determined by Bradford analysis except when noted differently (section 2.2.10).

2.2.9.2. Purification from E. coli TG-1/pLR7

A starter culture of *E. coli* TG-1 harboring pLR7 grown in LB_{amp} was subcultured (1:50 or 1:100) into fresh LB_{amp} media. Cells were grown at 37 °C either at 1.5 L volumes (using a 4 L flask) in an incubator shaker (180 rpm) or at 10 L volumes in a VirTis fermenter set at aeration level 4 and with stirring. Cultures were typically grown for 3.5 hr (to $OD_{600} \approx 0.5$-0.6) prior to the addition of solid IPTG to 0.5 mM. Cultures were grown an additional 3.5 hr post-induction and cells were harvested by centrifugation (10000 × g) or in a Sharples continuous flow centrifuge. In earlier experiments, cells were frozen in liquid N_2 and stored at -80 °C prior to processing while in later purifications, cells were processed immediately after harvesting. Cell yields were typically 2.0-2.2 g (wet weight) per L of culture.

i) Development of a Purification Protocol

Initial attempts to purifiy LaL from *E. coli* TG-1/pLR7 were generally performed on a scale of 2 g cells. Cells were suspended in 50 mM KPB, pH 7.0 (5 mL/g wet weight) and disrupted by 2 successive passages through a French press and cell debris was removed by centrifugation (25000-35000 × g).

The cell free extract was subjected to a 55% ammonium sulfate precipitation (0 °C) and the pellet was collected by centrifugation (15000 × g). In some cases, the cell lysate was first brought to 3% (w/v) streptomycin sulfate and centrifuged as above prior to subjection to 55% ammonium sulfate saturation. The pellet obtained was suspended in a minimal volume of 50 mM KPB, pH 7.0 and desalted by passage through a Sephadex G-25 column (1.5 × 25 cm) equilibrated in the same buffer. Active fractions (2.2.6) were pooled

and concentrated by ultrafiltration (Amicon PM-10 membrane). Alternatively, the suspended pellet was dialysed against 50 mM KPB, pH 7.0. In later experiments when it was realized that products used for concentration of protein samples resulted in poor recovery of LaL (see section 2.3.5.2), concentration of a sample was achieved by lyophilization with the lyophilized material being resuspended in MQW as desired.

Purification of the desalted 55% ammonium sulfate fraction was first explored by cation exchange chromatography over a Mono-S HR 5/5 column. The flow rate employed was 0.5 mL/min. Samples were adsorbed to the column equilibrated in Buffer A (50 mM KPB, pH 7.0). After an initial wash with Buffer A until elution of unabsorbed material was complete, the column was eluted with a linear gradient to Buffer B (1.0 M KCl (w/v) in Buffer A). Gradients of 1, 2 and 3% per min were investigated. In other cases, Buffer A was 50 mM KPB, pH 6.5 and Buffer B was 1.0 M KCl (w/v) in this Buffer A. LaL eluted at approximately 16-18% Buffer B (at pH 7.0) or at 21-23% Buffer B (at pH 6.5).

The Mono-S fraction consisted of approximately 25-30% LaL (estimated by SDS-PAGE) with a major 12 kDa protein contaminant. Therefore, in some experiments, the desalted 55% ammonium sulfate fraction was first subjected to gel filtration chromatography over Sephadex G-50 SF (1.7 × 80 cm) or Superdex 75 (HR 10/30) employing 50 mM KPB, pH 7.0 as the eluent. Active fractions from gel filtration were pooled and lyophilized and then subjected to Mono-S chromatography as described above.

Solid ammonium sulfate was added to the active Mono-S fractions at ambient temperature to a final concentration of 1.7 M. For the following, Buffer A is 50 mM KPB, pH 7.0. The sample was applied to a Phenyl-Superose HR 5/5 column equilibrated with buffer consisting of 1.7 M (w/v) ammonium sulfate in Buffer A. The flow rate employed was 0.5 mL/min. The column was washed with the initial buffer until unabsorbed proteins were eluted and then a linear gradient to Buffer A (over 30 min) was implemented. LaL eluted at approximately 45-50% Buffer A and was deemed homogeneous (>99%) by SDS-PAGE and silver staining.

ii) Large Scale Purification using the Developed Protocol

To 10 L of LB_{amp} (pre-warmed to 37 °C) was added an overnight culture of *E. coli* TG-1/pLR7 (100 mL). Cells were grown to an OD_{600} of 0.52, solid IPTG was added (1.19 g, to 0.5 mM) and the culture further incubated for 3.5 hr at 37 °C. Cells (21 g) were harvested and suspended in 100 mL of 50 mM KPB, pH 7.0. Cells were disrupted by French press, PMSF (17 mg dissolved in 0.1 mL acetone) was added and the sample was centrifuged (25000 × g). The supernatant was brought to 55% saturation with ammonium

sulfate and centrifuged (15000 x g). The pellet (≈ 10.5 g) was suspended in 80 mL of 50 mM KPB, pH 7.0 and dialysed (Spectra/Por® 3, 3500 MWCO tubing) against the same buffer (3 x 4 L). The dialysis retentate (95 mL, 1170 mg total protein) was applied in 8 separate runs to a Mono-S HR 10/10 column (flow rate 2 mL/min) previously equilibrated with 50 mM KPB, pH 7.0 (Buffer A). LaL was eluted in a two step linear gradient from 0-15% (over 5 min) then 15-25% (over 10 min) Buffer B (1.0 M KCl (w/v) in Buffer A). The active fractions from each of the 8 runs were pooled (19 mL, 5 mg total protein) and solid ammonium sulfate (4.27 g, to 1.7 M) was added slowly at room temperature. The sample was applied to a Phenyl-Superose HR 5/5 column (flow rate of 0.5 mL/min) previously equilibrated with 1.7 M (w/v) ammonium sulfate in Buffer A. A linear gradient to Buffer A (over 30 min) was implemented and LaL eluted at approximately 40% Buffer A. The Phenyl-Superose fraction (12 mL, 2.8 mg total protein) was made to 20% glycerol and then stored at -20 °C.

2.2.9.3. Purification from E. coli OR1265/pLcIR18

An overnight culture of *E. coli* OR1265 harboring pLcIR18 grown in LB_{amp} (at 32 °C) was subcultured (1:50) into each of the following: 2 x 0.5 L of LB_{amp} which contained a total concentration of 0.17 M NaCl (10 g/L) and into 2 x 0.5 L of LB_{amp} which contained a total concentration of 0.5 M NaCl (29.2 g/L). Cultures were grown at 32 °C in an incubator shaker (180 rpm) to absorbances at 600 nm of 0.4-0.5 (≈ 5 hr). Induction of cultures was achieved by first immersing the flasks into a 70 °C water bath for approximately 3-4 min and then into a 45 °C water bath for 15 min. Flasks were then cooled in an ice bath for 1 min and then incubated at 40 °C in an incubator shaker (180 rpm). Cultures were manually shaken during the incubations in the water baths. One of the cultures at each NaCl concentration was grown at 40 °C for 1 hr and the other grown at 40 °C for 3 hr. Cells were harvested by centrifugation (10000 x g), frozen in liquid N_2 and stored at -80 °C.

Cells (mass yields were not obtained) from each of the 4 growth conditions (i.e. the two NaCl concentrations at the two post-induction growth times) were processed identically as follows. Cells were thawed, suspended in 50 mM KPB, pH 7.0 (20 mL), disrupted by French press and cellular debris was removed by centrifugation (35000 x g). Continuation of purification of the 4 samples employed the protocol developed for *E. coli* TG-1/pLR7; i.e. a 55% ammonium sulfate precipitation, dialysis of the pellet and sequential chromatography over Mono-S 10/10 and Phenyl-Superose HR 5/5. For the Phenyl-Superose chromatography, the equilibration buffer was 1.2 M (w/v) ammonium sulfate in Buffer A (instead of 1.7 M).

2.2.9.4. Purification from E. coli TG-1/pHDM10

The protocol for purification of LaL from this system was modified from that described above. Development of this procedure was initially performed on 1 L cultures. LaL was then routinely isolated from 6 L cultures and the description which follows is representative of a typical purification. Where quantities are cited following a particular step, they represent a value (or range) averaged from 7 individual purifications.

An overnight culture of *E. coli* TG-1 harboring pHDM10 grown in LB_{amp} medium was subcultured (1:50) into 4 × 1.5 L volumes (using 4 L flasks) of the same medium. Cultures were grown at 37 °C with shaking (180 rpm) to an OD_{600} of 0.6-0.75 (3-4 hr). To each flask was added solid IPTG (180 mg, to 0.5 mM) and the cells were grown an additional 3.5-4 hr to a final absorbance of 1.4-1.6 at 600 nm. Cells (2.3 ± 0.3 g wet weight/L) collected by centrifugation (10000 × g) were suspended in 50 mM KPB, pH 7.0 (70 mL). The cell suspension was subjected to disruption by French press, centrifuged (35000 × g) and the supernatant was reserved. The cellular debris pellet was suspended in 50 mM KPB, pH 7.0 (20 mL), thoroughly vortexed and subjected to a second French press passage and centrifuged. The supernatants from each of the two passages were combined and subjected to another centrifugation step (39200 × g, 30 min) to clarify the cell free extract (≈ 100 mL, 1350-1610 mg total protein representing 240 ± 31 mg per L of culture).

The cell free extract was applied to a S-Sepharose Fast Flow HR 16/50 column (1.6 × 50 cm; void volume of ≈ 30 mL) equilibrated with Buffer A (50 mM KPB, pH 7.0). The sample was applied to the column at a flow rate of 2.5 mL/min and the column was washed with Buffer A until the A_{280} of the eluent decreased below 10% full scale. The flow rate was increased to 5.0 mL/min and a linear gradient (over 40 min, 200 mL) to Buffer B (1.0 M (w/v) KCl in Buffer A) was implemented. The eluent containing lysozyme activity (45-55 ml, 124 ± 24 mg total protein) was collected and dialysed (Spectra/Por® 3 or 7[2.1], 3500 MWCO tubing) overnight against Buffer A (2.5 L). Any precipitate that developed at the end of the dialysis was removed by centrifugation (35000 × g, 20 min).

The dialysed S-Sepharose Fast Flow fraction was divided into two portions and applied in two separate runs to a Mono-S HR 10/10 column equilibrated with Buffer A. A flow rate of 2.0 mL/min was used throughout for this chromatography. After washing the

---

[2.1] During our investigations it was determined that the Spectra/Por® 3 grade of tubing was contributing an unidentified contaminant(s) to solutions (see also 2.3.5.2). A substantial reduction in the contaminant(s) was achieved with use of the Spectra/Por® 7 grade of tubing. Following this realization, this tubing was used in all subsequent dialysis.

column with Buffer A until elution of unadsorbed material was complete, a two step linear gradient from 0-5% Buffer B (over 2 min, 4 mL) and then 5-15% Buffer B (25 min, 50 mL) was implemented. LaL was eluted at approximately 10.5% Buffer B (≈ 0.105 M KCl) and the eluents from both runs were pooled (14-16 mL, 68 ± 9 mg protein). The column was washed with at least 4 column volumes (i.e. > 4 × 8 mL) with the high salt elution buffer between chromatographic runs.

The pooled Mono-S fraction was treated with solid ammonium sulfate (added over 5 min at room temperature) to achieve a final concentration of 1.7 M. The fraction was then equally applied in two separate runs at 1.0 mL/min to a Phenyl Superose HR 10/10 column equilibrated with 1.7 M ammonium sulphate (w/v) in Buffer A. The column was washed with equilibration buffer until unadsorbed material had eluted. The flow rate was increased to 1.5 mL/min and a linear gradient from 0-50% Buffer A (over 30 min, 45 mL) was performed. Lysozyme was eluted from the column typically over 28 to 48% Buffer A (i.e. between approximately 1.2 and 0.9 M ammonium sulphate) and the eluents from both runs were pooled (≈ 30 mL). The column was washed with at least 3 column volumes (i.e. 3 × 8 mL) with Buffer A between chromatographic runs.

The pooled fraction was subjected to a second chromatographic step over Phenyl-Superose to concentrate the protein solution as follows. The concentration of ammonium sulfate in the pooled Phenyl-Superose fraction was estimated to be the average of the concentration at the onset and completion of the fraction collected; eg. for elution between 1.2 and 0.9 M, the concentration was assumed to be 1.05 M. Additional ammonium sulphate was added to bring the final concentration of the pooled sample to slightly over 1.7 M (eg. 0.7 M added to the 1.05 M solution) and the enzyme solution was reapplied to the Phenyl-Superose column as described above (1.0 mL/min). Immediately following the completion of the application of the solution to the column, the flow rate was increased to 1.5 mL/min and a sharp gradient to 100% Buffer A (over 0.5 min, 0.75 mL) was employed to elute lysozyme in a reduced volume (5-7 mL).

The concentrated enzyme was extensively dialysed (Spectra/Por® 3 or 7; 3500 MWCO) against 5 mM KPB, pH 7.0 that was prepared with MQW and filtered (0.45 μM). Dialysis was performed with 4 exchanges against 1.5 L of buffer. Following dialysis, the dialysis retentate solution was lyophilized and stored as a dry powder at -20°C. Protein was quantitated by mass of the lyophilized powder after correcting for mass arising from the buffer salts (refer to section 2.2.10.2).

2.2.9.5. Purification from E. coli BL21(λDE3)/pLR102

An overnight culture of *E. coli* BL21(λDE3) harboring pLR102 grown in LB_{amp} medium was subcultured (1:50) into 4 × 1.5 L volumes (using 4 L flasks) of the same medium. Cultures were grown at 37 °C with shaking (180 rpm) to an OD_{600} of 0.4-0.45 (≈2.5 hr). To each flask was added solid IPTG (268 mg, to 0.75 mM) and the cells were grown an additional 4.5 hr to a final absorbance of 1.5-1.6 at 600 nm. Cells (2.5 g wet weight/L) collected by centrifugation (10000 × g) were frozen in liquid N_2 and stored overnight at -20 °C. Cells were thawed at room temperature and suspended in 50 mL of 50 mM KPB, pH 7.0. The cell suspension was sonicated on ice until disruption was deemed complete (by repeated 20 s pulses spaced by 1 min intervals). The slurry was centrifuged (35000 × g) and the supernatant was reserved. The pellet was suspended in buffer (20 mL), vortexed thoroughly and subjected to a second sonication and centrifugation. The supernatants from each sonication were combined and centrifuged (39200 × g, 30 min) to clarify the cell free extract.

The subsequent purification of LaL from the cell free extract was performed as described in section 2.2.9.4 for *E. coli* TG-1/pHDM10 with the following exceptions:

1) The dialysed S-Sepharose Fast Flow fraction (54 mL) was chromatographed over the Mono-S HR 10/10 column in 6 individual runs (≈ 50 mg total protein loaded for each run). A major active fraction (eluting at ≈ 9.3-10.8% Buffer B, see peak "a" Fig. 2.25 A) was collected from each run and was frozen in liquid N_2 and stored at -20 °C. A minor active fraction from each run (eluting over ≈ 11.7-13.4% Buffer B, see peak "b" Fig. 2.25 A) was also collected, pooled and dialzyed against Buffer A. The dialysed sample was then reapplied to the Mono-S column and the fraction eluting at 10.4% Buffer B was collected (see Fig. 2.25 B).

2) The Mono-S fractions (collected at 10.4% Buffer B) were thawed, combined and were then chromatographed over the Phenyl-Superose HR 10/10 column in 4 individual runs. The active fractions from each of the 4 runs were pooled and then reapplied in 2 individual runs to the Phenyl-Superose column to concentrate the sample. Following this chromatography, the sample was dialysed extensively and lyophilized as described in section 2.2.10.2.

2.2.10. Protein Quantitation Methods

2.2.10.1. Bradford Analysis

During the initial stages of this work, protein determinations were estimated by the dye binding method of Bradford (1976) as described by Dawson et al. (1986) using bovine serum albumin as a standard. The dye reagent was prepared as follows:

Coomassie Brilliant Blue G-250	100 mg
95% Ethanol	50 mL
85% (w/v) Phosphoric acid	100 mL
MQW	to 1.0 L

- dissolve the dye in the ethanol and with continuous stirring, add the phosphoric acid and then the MQW.
- filter the solution through Whatmann No. 1 paper

In some cases, a commercial dye reagent concentrate (Bio-Rad Protein Assay Kit) was used following the suppliers directions for the micro-protein assay. Samples, standard solutions and reference blanks were prepared using 50 mM KPB, pH 7.0. All measurements were performed in triplicate.

2.2.10.2. Quantitation by Mass Following Dialysis and Lyophilization

The final step involved in the purification of LaL from *E. coli* TG-1/pHDM10 and BL21/pLR102 involved extensive dialysis of the concentrated Phenyl-Superose fraction against 5 mM KPB, pH 7.0. Following dialysis, the protein solution (the dialysis retentate) was transferred to a tared 10 dram vial (which was dried in a 140 °C oven and cooled in a dessicator) and an accurate mass of the solution (wet sample) was recorded using an analytical balance. In some cases, the volume of the retentate required the use of two vials. In addition, 3 aliquots (≈ 8-10 mL) of the final dialysis buffer (the dialysate) were also transferred to separate vials and the mass of each aliquot (wet sample) obtained as described. The retentate and dialysate solutions were allowed to warm to room temperature (dialysis was performed at 4°C) prior to their transfer to prevent condensation on the outer surface of the vials. The solutions were frozen in liquid N_2 and lyophilized to dryness. The mass of each dried sample and vial (dry sample) were obtained immediately following their removal from the lyophilizer.

The mass of protein was determined after correcting for the mass arising from the buffer salts carried over from the dialysis buffer. The mass values from a 6 L scale purification from *E. coli* TG-1/pHDM10 are given below to aid in the interpretation of the calculations:

	Dialysate Buffer Aliquots			Retentate
	1	2	3	(Protein)
Mass of Vial:	13.30074 g	13.21795 g	13.30871 g	13.26250 g
+ Wet Sample:	22.33945 g	22.36575 g	22.69867 g	20.52180 g
+ Dry Sample:	13.30830 g	13.22594 g	13.31670 g	13.30855 g
Wet Mass:	9.03871 g	9.14780 g	9.38996 g	7.25930 g
Dry Mass:	7.56 mg	7.99 mg	7.99 mg	46.05 mg
Dry Mass per g of Wet Mass:	0.836	0.873	0.851	

The lyophilized material obtained from the dialysed LaL sample will have a mass contribution from the buffer salts from the dialysis buffer. Therefore, the quantity of buffer salts that are obtained from each gram of dialysis buffer was found by dividing the total dry mass of the buffer salts by the total mass of the buffer before lyophilization (the wet mass). For the example given, an average of 0.853 ± 0.019 mg dry mass of buffer salts are obtained per gram wet mass of buffer. Knowing the wet mass of the protein sample (the retentate), one can calculate the contribution of buffer salts to the mass of the lyophilized protein using this average value as follows:

Protein Mass = total dry mass - mass arising from buffer salts

= total dry mass - [(wet mass of retentate (g))(dry buffer mass per gram wet buffer mass)]

= 46.05 - [(7.25930 g)(0.853 mg/g)]

= 39.86 mg

As such, the yield of protein is taken to be 39.9 mg.

As well, the mass percentage of LaL in the powder can be determined from 39.86/46.05 = 0.8656. That is, 86.56% of the powder is LaL by mass. This was useful since both the concentration of protein and phosphate buffer that would result when a known amount of the LaL/salt powder was dissolved in water could be determined. For example, if 1.5 mg of the powder mixture was dissolved in 1.85 mL water;

i) the protein concentration would be: $\frac{1.5 \text{ mg} \times 86.56\%}{1.85 \text{ mL}} = 0.7018 \text{ mg/mL}$

ii) the buffer concentration would be:

$$\text{ratio of the total dry mass of the retentate dissolved} \times \frac{\text{original vol. of the wet sample (i.e. the wet mass of the retentate)}}{\text{new volume of the solution}} \times \text{original concentration of the dialysis buffer}$$

$$= \frac{1.5\ \text{mg}}{46.05\ \text{mg}} \times \frac{7.25930\ \text{mL}}{1.85\ \text{mL}} \times 5.0\ \text{mM}$$

$$= 0.64\ \text{mM}$$

Then, if desired, the concentration of buffer could be increased with the addition of a stock solution of phosphate buffer.

2.2.10.3. Determination of the Molar Extinction Coefficient

The molar extinction coefficient (ε) of LaL was determined based on the method of Gill & von Hippel (1989) and the amino acid sequence (Sanger et al., 1982).

The method involves calculation of the extinction coefficient of the *denatured* protein in 6.0 M guanidine hydrochloride (Gdn·HCl) from the number of Try, Trp, and Cys residues per protein molecule with the assumption that Cys residues appear as half cystines using:

$$\varepsilon_{280,\text{Gdn.HCl}} = a\,\varepsilon_{280,\text{Tyr}} + b\,\varepsilon_{280,\text{Trp}} + c\,\varepsilon_{280,\text{Cys}}. \qquad [1]$$

The values $\varepsilon_{280,\text{Tyr}}$, $\varepsilon_{280,\text{Trp}}$, and $\varepsilon_{280,\text{Cys}}$ are the molar extinction coefficients of Tyr, Trp, and Cys residues at 280 nm and a, b, and c are the number of each residue per protein molecule. The individual ε_{280} values for each of the amino acids are obtained from those determined previously (Edelhoch, 1967) using the following model compounds in *6.0 M Gdn·HCl, 0.02 M phosphate buffer, pH 6.5:*

Trp	N-acetyl-tryptophanamide	$\varepsilon_{280} = 5690\ M^{-1}\ cm^{-1}$
Tyr	Gly-Tyr-Gly	$\varepsilon_{280} = 1280\ M^{-1}\ cm^{-1}$
Cys	Cystine	$\varepsilon_{280} = 120\ M^{-1}\ cm^{-1}$.

Edelhoch (1967) showed that these residues are the only ones that contribute significantly to the measured optical density of a denatured protein over the 276-282 nm range.

To determine the molar extinction coefficient of the *native* protein (ε_{nat}), the absorbance of the denatured (in 6.0 M Gdn·HCl, 0.02 M phosphate buffer, pH 6.5) and of the native protein are measured at identical protein concentrations. From Beer' law, one may write

$$A_{den(Gdn\cdot HCl)}/\varepsilon_{den(Gdn\cdot HCl)} = C_{den} \quad [2]$$

and

$$A_{nat}/\varepsilon_{nat} = C_{nat} \quad [3]$$

where $A_{den(Gdn\cdot HCl)}$ and A_{nat} are the measured absorbances of the denatured and native protein. If the concentration of denatured (C_{den}) and native (C_{nat}) are set equal experimentally, then equations [2] and [3] can combine to obtain the extinction coefficient of the native protein:

$$\varepsilon_{nat} = (\varepsilon_{den})(A_{nat})/(A_{den}). \quad [4]$$

From equation [1] the extinction coefficient of denatured LaL can be determined knowing that there are 5 Tyr (a=5), 4 Trp (b=4) and 1 Cys (c=1). Therefore, the calculated value for $\varepsilon_{LaL,Gdn\cdot HCl}$ at 280 nm in 6.0 M Gdn·HCl, 0.02 M phoshpate buffer, pH 6.5 is

$$\varepsilon_{LaL,\,den} = (5)(1280) + (4)(5690) + (1)(120)$$
$$= 29280\ M^{-1}\ cm^{-1}.$$

i) Preparation of stock Gdn·HCl solutions. The hygroscopic nature of Gdn·HCl does not permit the accurate preparation of solutions of known concentration by simple measurement of the weight of Gdn·HCl (Pace et al., 1990). An accurate method of determining Gdn·HCl concentration is by refractive index measurements. The following equation can be used to accurately determine the concentration of a Gdn·HCl solution,

$$C = 57.147\ (\Delta N) + 38.68\ (\Delta N)^2 - 91.6\ (\Delta N)^3 \quad [5]$$

where C is the concentration of Gdn·HCl solution (in $mol\cdot L^{-1}$) and ΔN is the difference between the refractive index of the Gdn·HCl solution in an aqueous buffer and the buffer itself (Nozaki, 1972).

Two stock solutions were required. Ultrapure Gdn·HCl was dried over P_2O_5 *in vacuo* and an amount slightly over that needed for a desired concentration was weighed. Refractive index measurements were made on a Mark II Refractometer (Reichart Scientific Instruments, Buffalo, NY). Four independent measurements of the refractive index for each solution were obtained and averaged with observed errors of less than 0.1%.

Stock solution a) Approx. 8.0 M Gdn·HCl, 10 mM DTT, 20 mM KPB pH 6.5. Gdn·HCl (≈ 7.7 g), DTT (15.4 mg) and 1.00 M KH_2PO_4 (200 µL) were mixed and MQW (≈ 1 mL) was added until dissolution. The solution was allowed to equilibrate to 25 °C and the pH adjusted to 6.5 with 5 N KOH. MQW was then added to a final volume of 10.0 mL. The same buffer (i.e. 10 mM DTT, 20 mM KPB, pH 6.5) lacking Gdn·HCl was also prepared. ΔN between the solutions was found to be 0.1324 which correlated to a concentration of Gdn·HCl of 8.0317 M (Eq. 5).

Stock solution b) Approx. 6.7 M Gdn·HCl, 22.2 mM KPB pH 6.5. Gdn·HCl (≈ 32 g) and 1.00 M KH_2PO_4 (1.111 mL) were mixed and prepared to pH 6.5 and a final volume of 50.0 mL as described above. The same buffer (i.e. 22.2 mM KPB, pH 6.5) lacking Gdn·HCl was also prepared. ΔN between the solutions was found to be 0.1108 which correlated to a concentration of Gdn·HCl of 6.6821 M (Eq. 5).

ii) Guanidine induced denaturation of LaL. This study was undertaken to ensure that LaL is denatured under the conditions required such that the $\varepsilon_{LaL,den}$ value of 29280 M^{-1} cm^{-1} calculated above can be used in Eq. [4]. Aliquots (20 µL) of a solution of LaL (purified from *E. coli* TG-1/pHDM10, approx. 2 mg·mL^{-1} in MQW) were added to appropriately mixed solutions (980 µL total) of Gdn·HCl stock solution (a) and 10 mM DTT, 20 mM KPB, pH 6.5 such that the final concentration of Gdn·HCl in the 1.00 mL protein/denaturant solution varied from 0 to 7.5 M in 0.5 M increments. Fluorescence measurements were taken of samples incubated at 25 °C after 1 and 3.5 hr and were made on a LM-480 fluorometer (SLM Instruments, Inc.) with excitation and emission wavelengths of 280 and 344 nm respectively. Emission spectra (over 300-600 nm) were also obtained for all samples at an excitation wavelength of 280 nm. Following the 1 hr incubation, an aliquot of each sample (1-5 µL) was removed and assayed as described previously (section 2.2.6).

iii) A_{280} measurements of equimolar denatured and native LaL samples. Two individual protein preparations were used in the determination of ε_{nat} (Eq. 4) obtained from purification from *E. coli* TG-1/pHDM10 and *E. coli* BL21/pLR102. A stock solution of each protein sample was prepared at approximately 4 mg·ml^{-1} in MQW.

Denatured samples. The following was performed (in triplicate) for both of the 2 protein stock solutions. A protein aliquot (50 or 100 µL), MQW (50 or 0 µL) and Gdn·HCl stock solution (b) (900 µL) were mixed to a total volume of 1.00 mL. Therefore, the final concentration of the denaturant buffer was 6.014 M Gdn·HCl, 20 mM KPB, pH 6.5 (as

required to satisfy conditions to determine C_{den}, Eq. 2). The samples were incubated for 1 hr at 25 °C and the A_{280} was measured for each solution. Denaturant buffer (6.014 M Gdn·HCl, 20 mM KPB, pH 6.5) was used as a reference solution.

Native Samples. The following was performed (in triplicate) for both of the 2 protein stock solutions. A protein aliquot (50 or 100 μL) and either MQW (950 or 900 μL) or 52.6 mM KPB, pH 7.0 (950 μL) or 55.6 mM KPB, pH 7.0 (900 μL) were mixed to a total volume of 1.00 mL. In this manner, the concentration of phosphate buffer in protein samples was either 0 or 50 mM. The A_{280} was measured using the appropriate reference solution (MQW or 50 mM KPB, pH 7.0).

iv) Determining the concentration of unknown LaL samples

The sample was diluted (in either MQW or KPB (< 50 mM), pH 7.0) to give an A_{280} reading between 0.3-0.7. The concentration of the original sample was then determined using the experimentally determined ε_{280} value of 31712 $M^{-1}\text{-}cm^{-1}$ (refer to section 2.3.3) and correction by the dilution factor.

2.2.11. Electrospray Mass Spectrometry

Electrospray mass spectrometry was performed on a Fisons VG Quatro II triple quadrupole mass spectrometer. Data acquisition and processing was performed with the MassLynx software. The sample delivery solvent consisted of 1:1 H_2O/CH_3CN containing 0.1% TFA or 0.1 % acetic acid and was used at a flow rate of 10-50 μL/min. Samples of LaL were prepared by dissolving a small portion of the lyophilized powder in delivery solvent and followed by centrifugation of the sample (13000 rpm on Biofuge A, 10 min). When prepared at 0.1-0.5 mg/mL, excellent ion counts were obtained when 10-20 μL of the sample was injected into the spectrometer. Desalting of protein samples was achieved by chromatography over a 2.0 × 150 mm Delta Pak™ C_{18} column (Waters) as follows. The column was eluted at 0.2 mL/min employing linear gradients from Buffer A (0.1% TFA in MQW) to Buffer B (0.1% TFA in CH_3CN) with detection at 215 nm. The column was equilibrated in 80% Buffer A and the gradient was 0-10-35 min, 20-40-65% Buffer B. A solution of LaL in MQW (≈ 1 mg/mL) was prepared and 20 μL of this solution was injected onto the column. LaL eluted at ca. 19 min and the peak was collected from baseline to baseline and used directly for ESMS (typical injections of 20 μL of the eluent collected gave excellent ion counts).

Typical operating parameters for data acquisition were:

MS1 Scanning		Tuning	
Range:	600-1200 Da	Ionization Mode	ES+
Ion energy:	0.3 Volts	Capillary	3.39 KVolts
Ion Energy Ramp:	0.0 Volts	HV Lens	0.52 KVolts
LM Resolution:	15.0	Cone	39 Volts
HM Resolution:	15.0	Skimmer Offset	5 Volts
Lens 5:	100 Volts	Skimmer	0.9 Volts
Lens 6:	7 Volts	RF Lens	0.3 Volts
Multiplier 1:	650 Volts	Source Temp.	80 °C
Scan Duration:	3-5 s		

These parameters gave a very good multiply charged ion distribution series for LaL when the sample was prepared as described above. The only parameters that were varied to optimize each sample were the Low Mass (LM) and High Mass (HM) resolution values (less than ± 1.0) and the cone voltage (less than ± 5 volts).

The raw multiply charged data were then analyzed using the following typical processing parameters:

Background Subtract		Spectrum Smooth		Spectrum Center Method	
Polynomial order:	0-2	Peak width:	0.6-0.8 Da	Channels:	2-4
Below curve (%):	33.00	No. of smooths:	2	Centroid Top:	80%
Tolerance:	0.010	Smoothing method:	mean		

At times, the qualitiy of the data dictated the choice of the processing parameters. The polynomial order was chosen for either a flat baseline (0) or a more curved baseline (2). The peak width is the peak width at half height measured from the raw data spectrum and for samples examined, was typically in the range given. All three of the processing functions (performed in the order from left to right shown above) are required to measure the mass of the components in a spectrum.

The MaxEnt algorithm (a brief description of the algorithm is given in 2.3.2.2) was used to produce true molecular mass spectra from the raw data. To obtain a MaxEnt reconstructed spectrum, the raw data was subjected to a background subtraction and then the peak width at half height was measured. The Uniform Guassian damage model software option was selected and the width at half height was entered as a value 0.05-0.1 Da higher than that measured. The resolution (in Da/channel) was set according to that which was deemed conducive to the quality of the raw data and was typically set from 0.25-1.0 Da as desired.

2.2.12. Estimation of the Isoelectric Point (pI)

A cell free extract obtained from a 2 L culture of *E. coli* TG-1/pHDM10 (≈ 30 mL) was dialysed overnight (4 °C) against 2 L of 20% glycerol. The volume of the dialzyed sample was brought to approx. 52 mL with 20% glycerol after which BioRad BioLyte 8/10 ampholytes (2.75 mL of a 20% solution, to a final concentration of 1%) were added. The sample was loaded into a Rotofor® (BioRad) preparative isoelectric focusing unit and focused at constant power (12 W) until the voltage levelled (ca. 4 hr). The sample cell holds 20 fractions and following their collection, the pH of each fraction was measured. A sample of each fraction was subjected to SDS-PAGE on a 10-15% gradient PhastGel and silver stained.

2.2.13. Crystallization Attempts

Initial attempts at crystallization of LaL were performed in our laboratory using the Crystal Screen™ Reagent Kit from Hampton Research (Riverside, CA). The kit consists of 50 separate reagents formulated from a broad mixture including various salts, buffers and precipitants which are listed in Appendix B. A sample of LaL (purified from *E. coli* BL21/pLR102) was exchanged into MQW (using a Centricon-3 unit) to a concentration of 12.12 mg/mL and then filter sterilized. The hanging drop method (McPherson, 1982; Ducruix & Giegé, 1992) was used and was performed under aseptic conditions. Equal volumes of each reagent and LaL solution (2 μL) were mixed on a cover slide (sterilized in ethanol) which was then inverted over a reservoir of a Linbro box tissue culture plate containing 1 mL of the same Crystal Screen reagent. A thin layer of vacuum grease was applied to the rim of each reservoir to provide a good seal with the cover slide. The screen was performed in duplicate plates incubating one plate at 4 °C and the other at ambient temperature. The droplets from each of the plates were examined over time under a stereo microscope recording all observations.

2.3. RESULTS AND DISCUSSION

2.3.1. Cloning of the λ *R* Gene and Purification of LaL from Strains Harboring Recombinant LaL Expression Vectors

2.3.1.1. Construction of pLR7 and Purification from E. coli TG-1/pLR7

Construction of plasmid pLR7 was our first attempt to create a system for overexpression of LaL. The complete lysis cluster containing the *S*, *R* and *Rz* genes was obtained from genomic lambda DNA as a 1.47 kb *Eco*RI-*Cla*I fragment. This fragment was cloned into pBR322 generating pLR1 (Fig. 2.1). Plasmid pLR1 was cleaved with *Eco*RI and the linearized plasmid was treated with exonuclease *Bal*31 (Fig. 2.2). The *Bal*31 digestion of pLR1 was performed in order to generate a series of deletions of the λ DNA insert in an attempt to remove the *S* gene with the intent of generating a λ DNA fragment with little DNA 5' to the *R* gene. The '*S* deletions were then cloned into plasmid pTTQ18 as illustrated in Fig. 2.2 and the recombinant plasmids were transformed into *E. coli* TG-1. Several active transformants were obtained, each of which would be expected to have recombinant plasmid DNA with varying degrees of deletion to the *S* gene. The pLR7 transformant was selected as it exhibited the highest relative activity.

The size of the *Eco*RI-*Hin*dIII insert from pLR7 was approximately 1.14 kb as determined from its migration in 1.0% agarose. The 1.14 kb fragment was cloned into the M13 vectors mp18 and mp19, allowing the sequence to be determined from each end (Fig. 2.3). The entire sequence of the insert was not obtained but correct sequence was obtained for i) 130 bp into the *Rz* gene from the 3' end and ii) 149 bp into the 5' end of the insert. The latter revealed that only 89 bp of the *S* gene had been deleted; therefore approximately 219 bp of unwanted *S* gene remained between the *R* gene start codon and the 5' insertion site (*Sma*I) of the fragment in the polylinker region of the pTTQ18 vector (Fig. 2.4). Confirmation of the *R* gene within the insert is established since sequence of both the *S* and *Rz* genes flanking the *R* gene was obtained although no sequence data for the *R* gene itself was acquired.

The 1.14 kb insert DNA fragment in pLR7 is under transcriptional control of the strong and regulable *tac* promoter (a hybrid of the *trp-lac* promoters). Repression of transcription during initial cell growth is achieved by the plasmid encoded $lacI^{Q}$ gene product, the *lac* repressor, which binds to a *lac* operator sequence immediately downstream from the promoter sequence. Derepression is achieved by the inclusion of IPTG in the growth media. Based on LaL activity, the highest level of LaL expression was achieved when the TG-1 cells harboring pLR7 were grown to an $OD_{600} \approx 0.5$-0.6 prior to

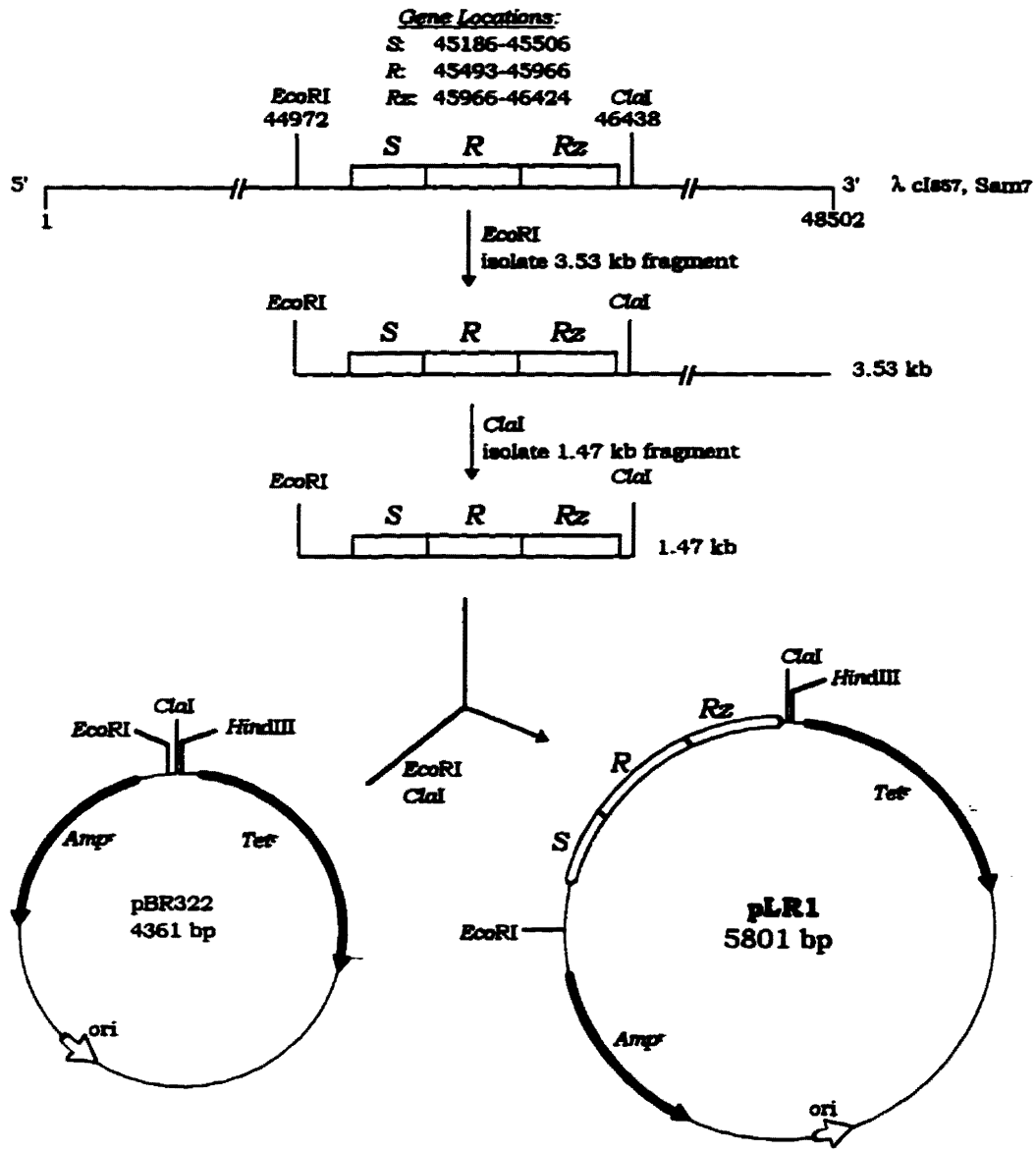


Figure 2.1. Construction of plasmid pLR1 containing the complete lysis genes *S*, *R*, and *Rz*. The 3.53 kb *Eco*RI fragment of λ DNA was isolated and further cleaved with *Cla*I. The 1.47 kb *Eco*RI-*Cla*I fragment was isolated and ligated into the *Eco*RI and *Cla*I sites of plasmid pBR322 yielding pLR1.

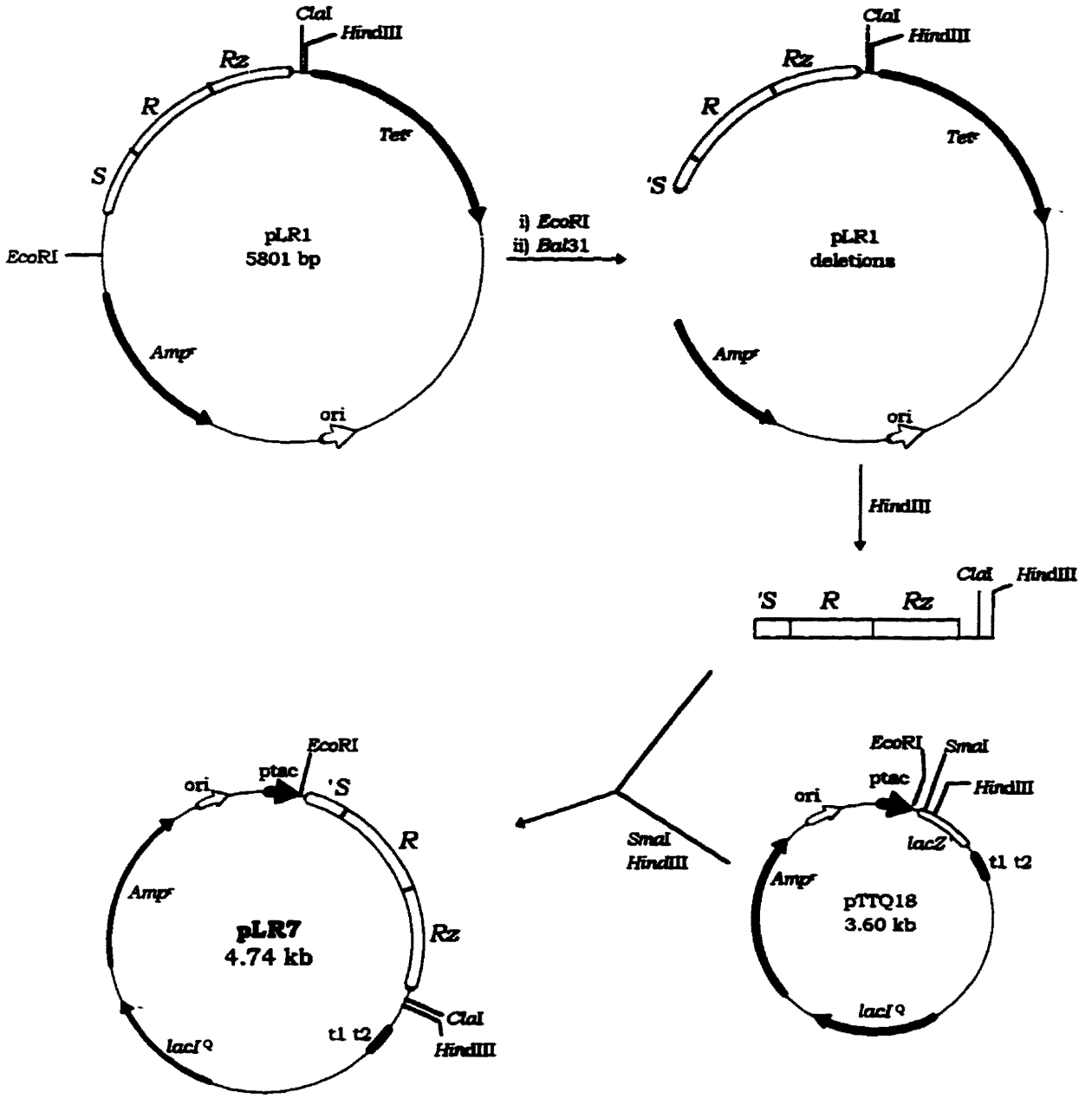


Figure 2.2. Construction of plasmid pLR7. Plasmid pLR1 was digested with *Eco*RI, treated with exonuclease *Bal*31 and then cleaved with *Hind*III. The 'S deletion/*Hind*III fragment was isolated and ligated into the *Sma*I and *Hind*III sites of plasmid pTTQ18 yielding pLR7.
Symbols: ptac, *tac* promoter; t1 t2, *rrnB* terminator fragment; *lacI*^{q}, *lacZα* coding region.

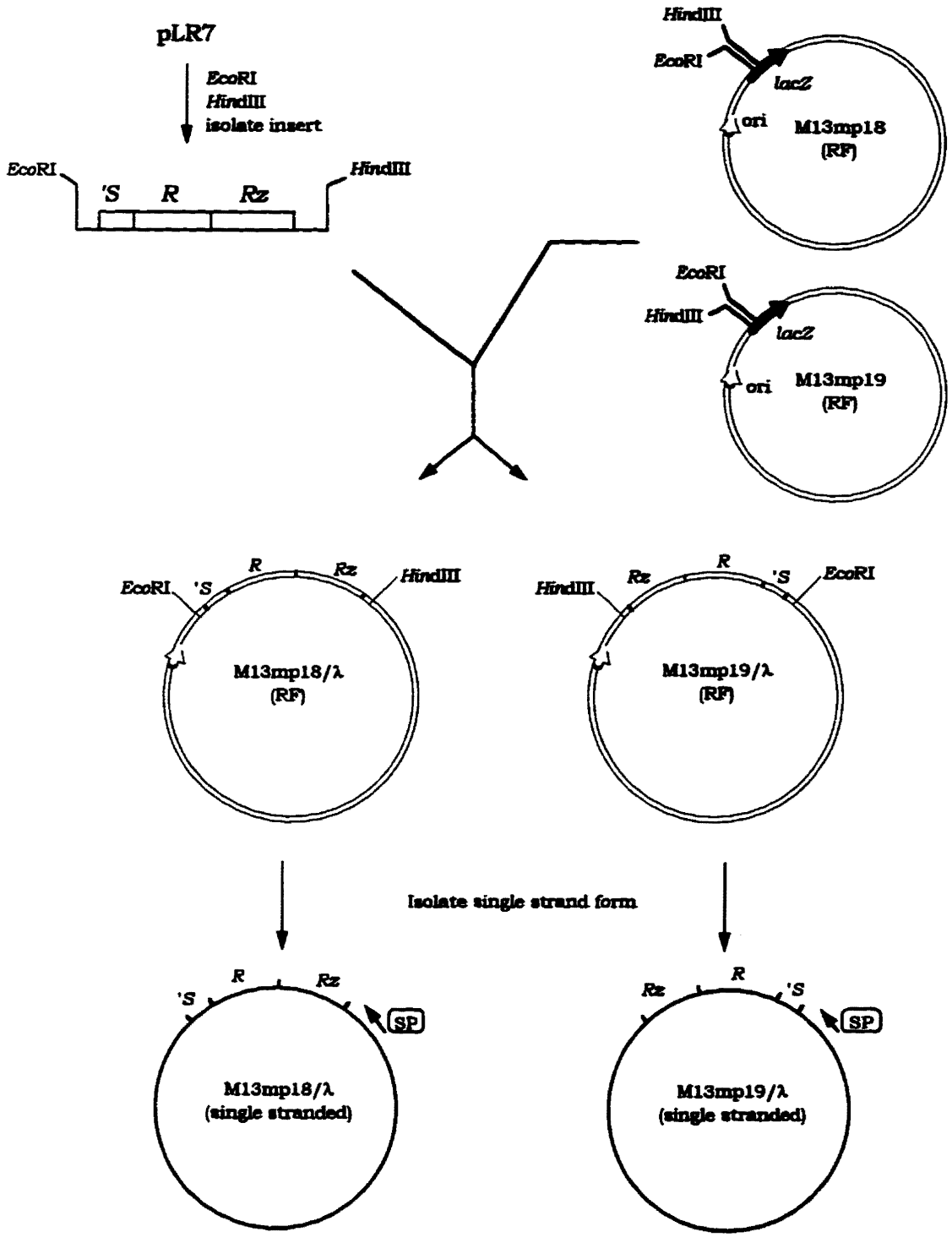


Figure 2.3. Schematic of the cloning of the pLR7 λ DNA fragment into M13 vectors for sequencing. The insert was excised from pLR7 and ligated with replicative form (RF) M13mp18 and M13mp19 and single stranded forms of the recombinant plasmids were produced and isolated. The M13 Sequencing Primer (SP) was used to obtain sequence into the *Rz* gene (130 bp) from M13mp18 and into the *'S* gene (149 bp) from M13mp19.

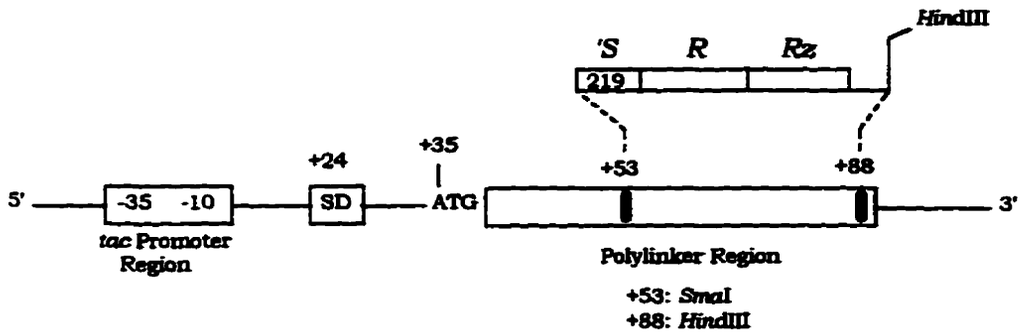


Figure 2.4. Schematic of the *tac* promoter and polylinker regions of pTTQ18. Only relevant restriction sites in the polylinker are shown. The mode of insertion of the λ DNA fragment (not drawn to scale) into the polylinker for pLR7 is also shown.

induction with IPTG and then followed by a 3-4 hr post-induction growth time before harvesting of the cells.

The initial purification step attempted from the cell free extract involved ammonium sulfate precipitation. This step was not very efficient. The concentration of ammonium sulfate required for optimal precipitation of LaL and in which LaL could be purified to homogeneity could not be achieved. Although saturation of the cell free extract with 65% ammonium sulfate resulted with near quantitative recovery of activity from the cell free extract, LaL could not be purified to homogeneity following further processing of the 65% ammonium sulfate fraction. Lysozyme was incompletely precipitated at 35% ammonium sulfate saturation (i.e. substantial activity remained in the 35% supernatant). Saturation of the cell free extract with 55% ammonium sulfate did result in good recovery of activity in which LaL could be subsequently purified. A very crude estimation of 60-80% recovery of LaL from the cell free extract is suggested from comparison of their respective activities. LaL maintained in the 55% ammonium sulfate pellet and storage at 4 °C retained activity for prolonged periods (a period of 3 months was monitored).

Subjecting the cell lysate to streptomycin sulfate (3% w/v) to remove nucleic acid cellular material prior to the 55% ammonium sulfate precipitation was also attempted. Substantial LaL activity was lost during the streptomycin sulfate precipitation. Samples were obtained from both a streptomycin/ammonium sulfate or solely an ammonium sulfate precipitation. SDS-PAGE analysis of these samples further purified by Mono-S chromatography revealed that the relative amount of LaL (to contaminating protein) following Mono-S purification was greater for the sample subjected to only the ammonium sulfate precipitation. Inclusion of a streptomycin sulfate precipitation was therefore abandoned as it decreased the overall recovery of LaL.

The 55% ammonium sulfate fraction was desalted by gel filtration or dialysis. The basic nature of LaL (pI 9.7) prompted exploitation of cation exchange chromatography. As such, purification was continued by Mono-S chromatography. This chromatography was extremely efficient in that greater than 99% of the contaminating protein from the ammonium sulfate fraction was removed as estimated from (i) the area under the peak corresponding to LaL activity to that of the total area under the Mono-S chromatogram (Fig. 2.5) and (ii) from the ratio of the total protein in the Mono-S fraction to the desalted ammonium sulfate fraction (data not shown). LaL was eluted from the column with an increasing gradient to buffer containing 1.0 M KCl (Buffer B, see legend to Fig. 2.5). A gradient to Buffer B of 1%/min was chosen as application of this gradient offered minimal, yet slightly improved resolution of LaL than application of steeper gradients of 2 or 3%/min. When the column buffer (Buffer A, Fig. 2.5) was 50 mM KPB, pH 7.0, less contaminating protein co-eluted with LaL than when the pH of Buffer A was 6.5 and for this reason, the pH 7.0 buffer was chosen. A prior gel filtration step of the ammonium sulfate fraction over Sephadex G-50 or Superdex 75 was investigated in an attempt to remove a major 12 kDa protein contaminant (as estimated by SDS-PAGE) that co-eluted with LaL from the Mono-S chromatography. The gel filtration step was not successful in removing the 12 kDa contaminant and for this as well as practical considerations, was withdrawn as a pre-Mono-S purification step.

Purification of LaL was achieved by hydrophobic chromatography of the Mono-S fraction over Phenyl-Superose. The majority of contaminating proteins were not adsorbed to this column. LaL was retained on the column and could be separated from other adsorbed protein during application of the decreasing salt gradient (Fig. 2.6). LaL obtained from the Phenyl-Superose chromatography constituted approximately 50-60% of the total protein of the Mono-S fraction and was homogeneous (>99%) by SDS-PAGE and silver staining. Figure 2.7 shows the SDS-PAGE results of the stages of the purification from *E. coli* TG-1/pLR7.

LaL purified from 2 g scales of cells typically resulted in Phenyl-Superose fractions that contained from 0.05-0.1 mg/mL LaL (by Bradford analysis). LaL activity was lost over a period of 5 days when these fractions were stored at 4 °C. Freezing of these samples in liquid N_2 also resulted in loss of enzyme activity. The instability of these samples can be attributed to their dilute protein content as was realized during optimization of the substrate cell assay (Chapter 3, see also section 2.3.5.1). Addition of glycerol (to a final concentration of 20%) and storage at -20 °C or freezing in liquid N_2 and storage at -80 °C resulted in indefinite stability of the enzyme (2.3.5.1). These samples also contained

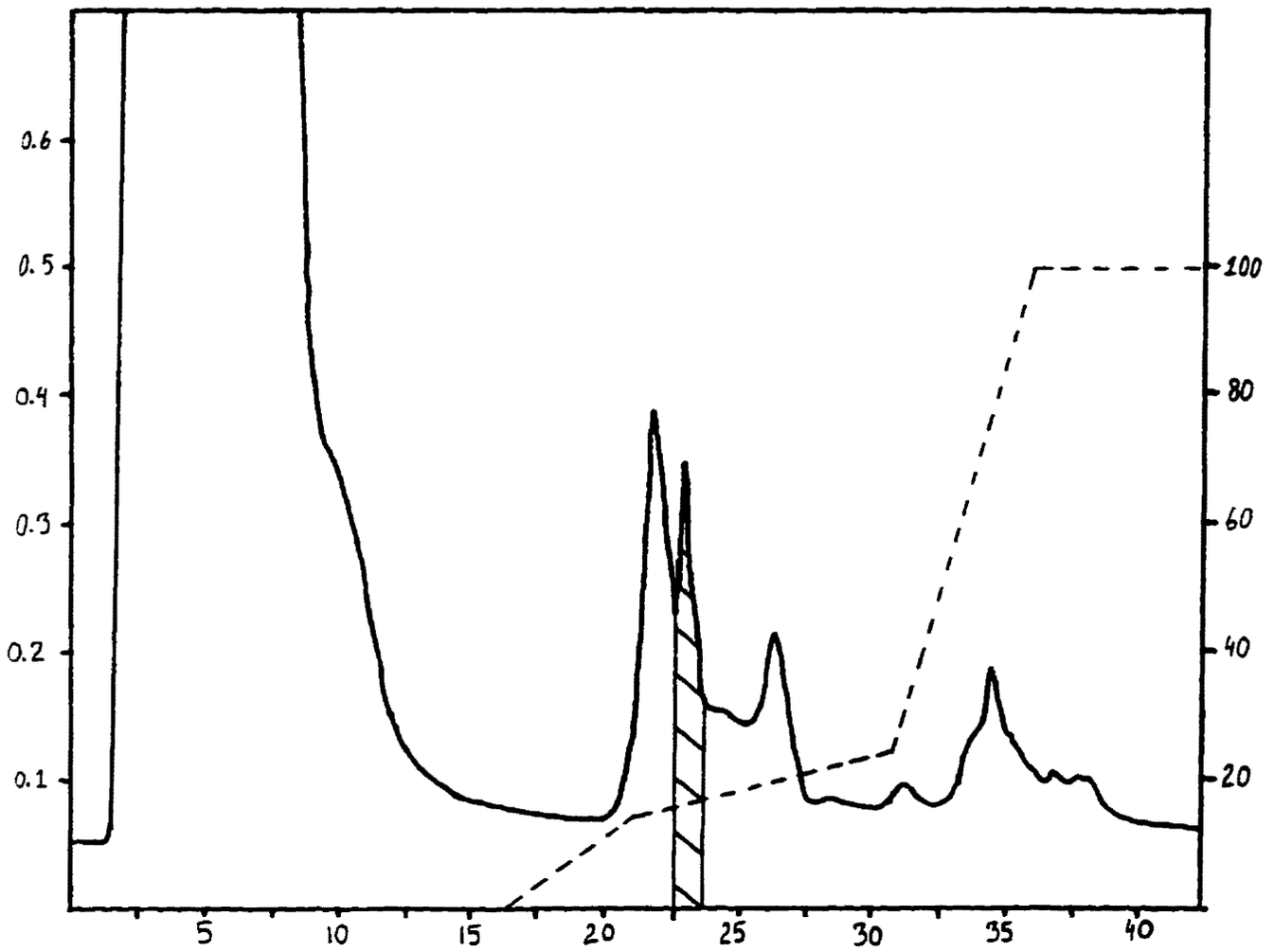


Figure 2.5. Mono-S chromotogram from purification of LaL from *E. coli* TG-1/pLR7. Shaded area represents the active fraction collected.

Sample:	55% ammonium sulfate fraction (Dialysed)
Column:	Mono-S HR 10/10
Initial Buffer:	Buffer A
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.0 M KCl (w/v) in Buffer A
Flow Rate:	2.0 mL/min
Left Axis:	Absorbance at 280 nm, indicated with the solid line
Right Axis:	Percentage Buffer B, indicated with the dashed line
Bottom Axis:	Time in minutes
Gradient:	0-15 % Buffer B over 5 min
	15-25% Buffer B over 10 min
	25-100% Buffer B over 5 min

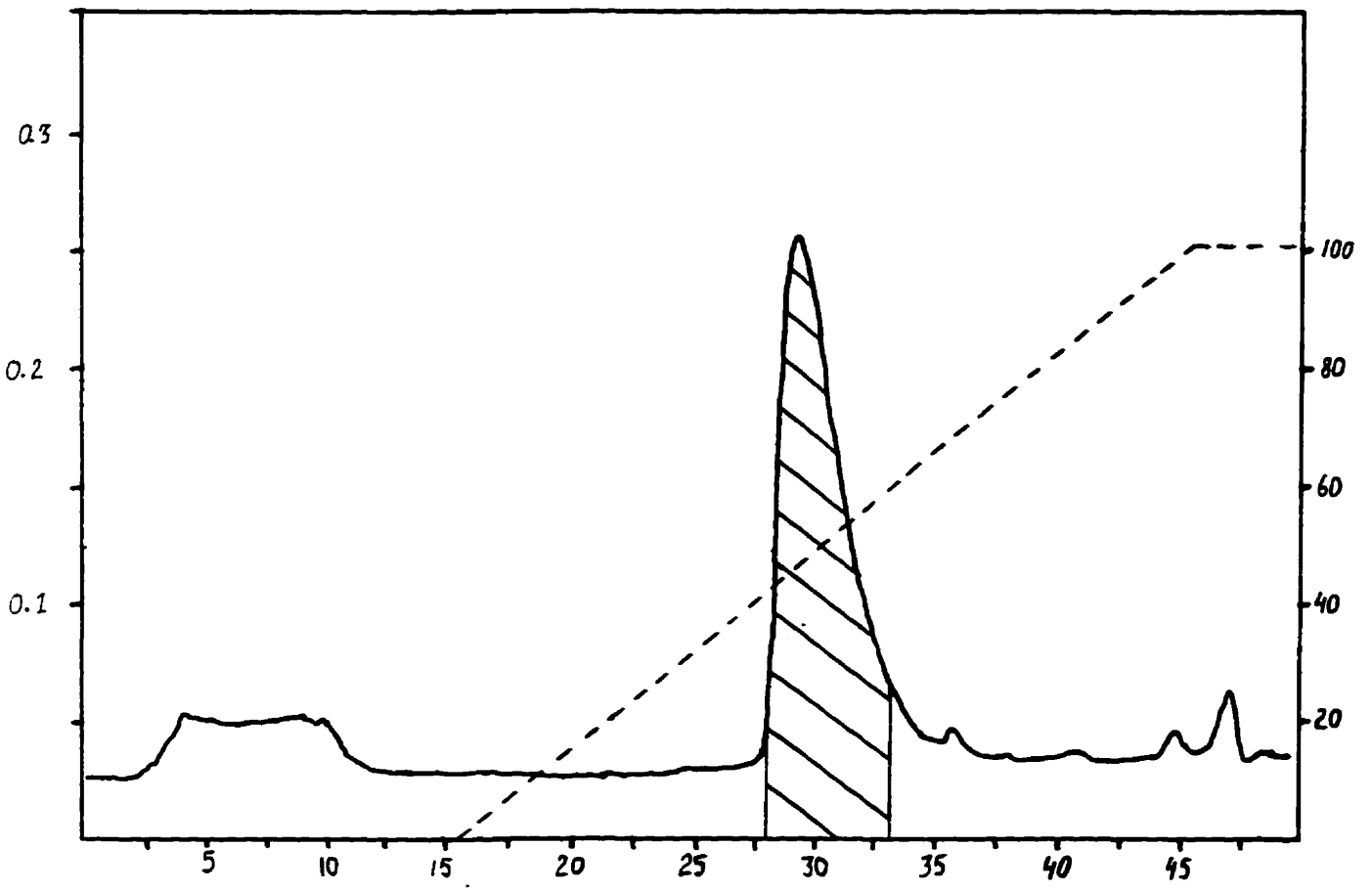


Figure 2.6. Phenyl-Superose chromatogram from purification of LaL from *E. coli* TG-1/pLR7. Shaded area represents the active fraction collected.

Sample:	Pooled Mono-S fractions, adjusted to 1.7 M $(NH_4)_2SO_4$
Column:	Phenyl-Superose HR 5/5
Initial Buffer:	Buffer B
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.7 M ammonium sulfate (w/v) in Buffer A
Flow Rate:	0.5 mL/min
Left Axis:	Absorbance at 280 nm, indicated with the solid line
Right Axis:	Percentage Buffer A, indicated with the dashed line
Bottom Axis:	Time in minutes
Gradient:	0-100% Buffer A over 30 min

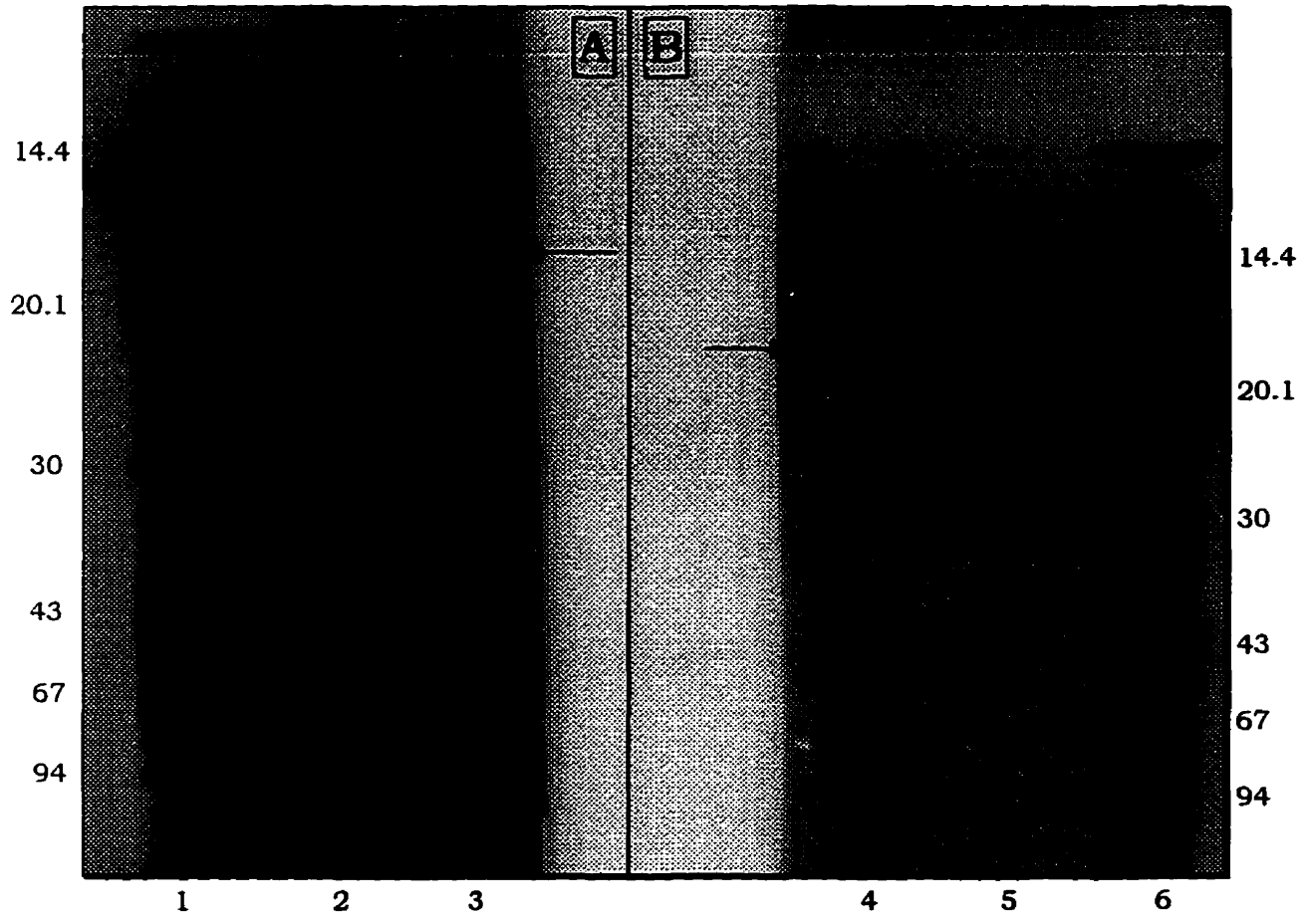


Figure 2.7. SDS-PAGE analysis of the stages of purification of LaL from *E. coli* TG-1 harboring pLR7. The samples were electrophoresed on two separate Homogeneous 20% Phast Gels (Gel A and B) followed with silver staining. The arrows indicate the position of LaL in each gel. The figure is best viewed at arm's length.

Lane:

1: Protein Standards (in kDa)
2: 55% Ammonium Sulfate fraction (dialysed)
3: Mono-S fraction
4: Phenyl-Superose fraction
5: Phenyl-Superose fraction (1:10 dilution)
5: Protein Standards (in kDa)

ammonium sulfate carried over from the Phenyl-Superose chromatography. Other methods developed for prolonged storage of LaL are described later (2.3.5.1).

The protocol developed for the purification of LaL from *E. coli* TG-1/pLR7 was used for two large scale purifications of LaL starting from 10 L of culture. Bradford analysis of the purified protein obtained from each suggested yields of approximately 0.2 and 0.3 mg/L (these are actually overestimates, see section 2.3.4). The low yields obtained were surprising considering the levels of protein expression others have attained using recombinant plasmids derived from pTTQ18 (Stark, 1987) and two explanations for this are offered. Firstly, the protocol developed at this stage for purification of LaL from TG-1/pLR7 was not as highly efficient as those developed later, and the levels of purified protein obtained cannot be expected to fittingly reflect the true cellular expression levels.

The low yields, however, could be the result of an expression phenomena. It is a common hypothesis that when low expression levels result from a gene introduced into *E. coli*, it is not a transcriptional problem, yet a translational one in which the messenger RNA (mRNA) transcript is inefficiently translated (de Boer & Hui, 1990). It is assumed that transcription of the pLR7 insert proceeds with the efficiency expected for the *tac* promoter, generating high mRNA copies of the insert. Reduced expression may be a consequence of the amount of DNA present between the promoter region and upstream to the *R* gene. The *Sma*I site into which the λ DNA insert of pLR7 was cloned is located 29 bp downstream from the Shine-Dalgrano (SD) sequence engineered into the polylinker; therefore, including the 219 bp of *S* gene DNA places the *R* gene start codon 248 bp downstream from the engineered site of translation initiation (Fig. 2.4). For this reason, it is expected that initiation of translation of the *R* gene mRNA occurs with the endogenous λ *R* ribosome binding site (RBS) and not the one engineered into the promoter/polylinker region of pTTQ18 (the pTTQ18 encoded RBS comes from pUC18, the template used in construction of pTTQ18 (Stark, 1987)). Comparative analysis with statistical models (Stormo et al., 1982) have led others (Jespers et al., 1991) to suggest that the *R* gene RBS presents negative features including a short SD sequence (4 nt) which is too close to the start codon (6 nt) and unfavourable G and C nucleotides in the immediate vicinity of the SD sequence (Fig. 2.8). As such, the wild-type *R* gene RBS is far from optimal which could lead to low translational efficiency of the *R* gene mRNA. As will be discussed in the development of pHDM10, which is also a recombinant plasmid derived from pTTQ18 (see 2.2.5.4 and 2.3.1.3), much of the DNA 5' to the *R* gene was eliminated which led to substantially enhanced expression levels.

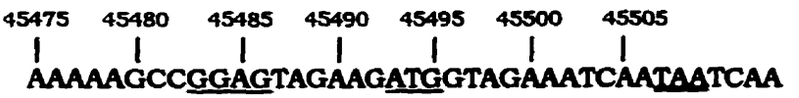


Figure 2.8. Sequence of the *R* gene ribosome binding site. The typical size of the RBS extends 17-23 bp 5' to and 12-18 bp 3' to the start codon (Schoner et al., 1985). The *R* gene start codon and SD sequence are underlined. The stop codon for the *S* gene is double underlined. Sequence data and nucleotide numbering are related to the λ genome and are taken from Sanger et al. (1982).

LaL purified from *E. coli* TG-1/pLR7 was not used in many subsequent studies and was therefore, not fully characterized. Apart from its homogeneity on SDS-PAGE, a sample of the enzyme was kindly brought to Sciex by Dr. Gilles Lajoie for ESMS analysis. The results provided a mass of 17822 ± 3 Da (MW calculated from sequence = 17825 Da) offering implicit evidence that purification of LaL was indeed accomplished. Although the protein yields were disappointing for the studies we had envisaged, the considerable effort dedicated towards purification methodology development was worthwhile. The purification strategy was readily adapted with minor modifications to the stronger LaL expression systems.

2.3.1.2. Construction of pLcIR18 and Purification of LaL from E. coli OR1265/pLcIR18

The construct pLcIR18 is a recombinant plasmid derived from pCM101. Plasmid pCM101 was previously constructed for the very effective overproduction of the two subunits, gp*Nu1* and gp*A* (the *Nu1* and *A* gene products) of λ DNA terminase (Murialdo et al., 1987; Chow et al., 1987). The coding region of pCM101 for the terminase genes was removed and replaced with a λ R gene DNA fragment. As illustrated in Fig. 2.9, the *R* gene was excised from pLR1 utilizing the *Hin*fI restriction sites, which are located 55 bp 5' to and 7 bp 3' to the *R* gene. Addition of *Bam*HI linkers to the 535 bp *Hin*fI fragment allowed the introduction of the 546 bp *R* gene fragment into the truncated pCM101 vector. The host strain used for pLcIR18 was *E. coli* OR1265.

In this construct, the insert containing the *R* gene is under the transcriptional control of the efficient lambda P_L promoter. This operon is subjected to regulation by the plasmid encoded lambda *cI*857 repressor. The *cI*857 gene is also provided by the OR1265 host, which contains this and other λ genes (Reyes et al., 1979). The *cI*857 repressor is

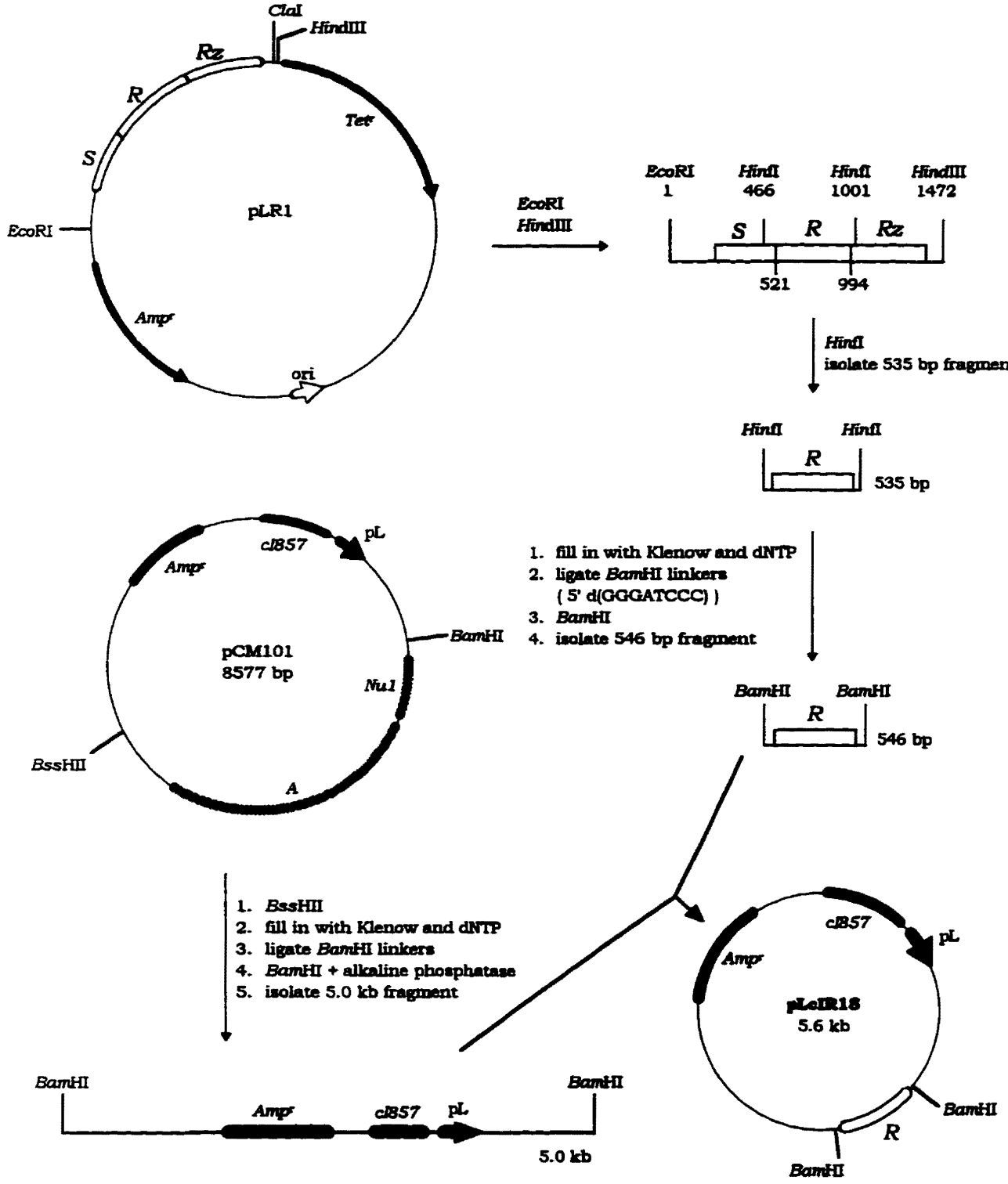


Figure 2.9. Construction of plasmid pLcIR18. Plasmid pCM101 was restricted with *Bss*HII and then *Bam*HI linkers were introduced. The *Nu1* and *A* genes were removed by subsequent cleavage with *Bam*HI and isolation of the 5.0 kb fragment. This fragment was ligated with the 546 bp *R* gene fragment yielding pLcIR18.
Symbols: *Nu1* and *A*, subunits of λ DNA terminase; cI857, λ repressor; pL, λ promoter.

thermolabile. During initial growth of the strain at 32 °C, the functional lambda repressor blocks transcription under control of the P_L promoter. A temperature shift to 45 °C (heat induction) causes thermal inactivation of repressor function which is sustained during post-induction incubation at 40 °C. As such, protein expression is achieved by growth of the cells at 32 °C to early logarithmic growth followed by a 15 min incubation at 45 °C and post-temperature shift incubation at 40 °C.

Purification of LaL was attempted from cultures grown under different growth conditions. The temperature shift inducing protein expression results in increased levels of the La protease (encoded by the *lon* gene). The La protease is one of seventeen *E. coli* proteins, termed heat shock proteins, that are induced under stressful conditions during what has been called the "heat shock response" (Goff et al., 1984; Goff & Goldberg, 1985). The La protease, whose induction increases 2-3 fold under stress or with the production of a cloned foreign protein, functions to degrade unfolded or "abnormal" proteins (Goff & Goldberg, 1985). The protease has also been shown to be inhibited *in vitro* by a high concentration of NaCl but is rather insensitive to PMSF (Waxmann & Goldberg, 1982; Tabor & Tabor, 1976). As such, media was prepared in which the total NaCl concentration was either 0.17 M (the standard concentration in LB medium) or elevated to 0.5 M. As well, purification of LaL was performed on cells that were subjected to either a 1 or 3 hr post-induction period to observe how expression levels were affected by La protease mediated degradation of LaL or possible cellular exhaustion.

The purification of LaL was attempted using the protocol developed for TG-1/pLR7, that is a 55% ammonium sulfate precipitation followed by sequential chromatography over Mono-S and Phenyl-Superose. However, purification following this methodology was not successful in obtaining purified LaL. It became obvious that induction and expression from the OR1265/pLcIR18 strain produced a considerably different cellular protein composition than for TG-1/pLR7. As seen for the Mono-S elution profile for TG-1/pLR7, the majority of the applied fraction does not adsorb to the column and elutes early during chromatography and LaL elutes in a relatively narrow peak (Fig. 2.5). The Mono-S elution profile for OR1265/pLcIR18 reveals a larger percentage of basic proteins. There is considerably less early eluting protein and a higher proportion of protein is adsorbed and eluted during application of the salt gradient (Fig. 2.10). More protein was found to co-elute with LaL in a broader peak. Chromatography of the Mono-S fraction over Phenyl-Superose resulted in contaminating protein co-eluting with LaL. As well, more protein from the Mono-S fraction was adsorbed to the Phenyl-Superose column (Fig. 2.11) than occurred with application of the same fraction from TG-1/pLR7 (Fig. 2.6). With the

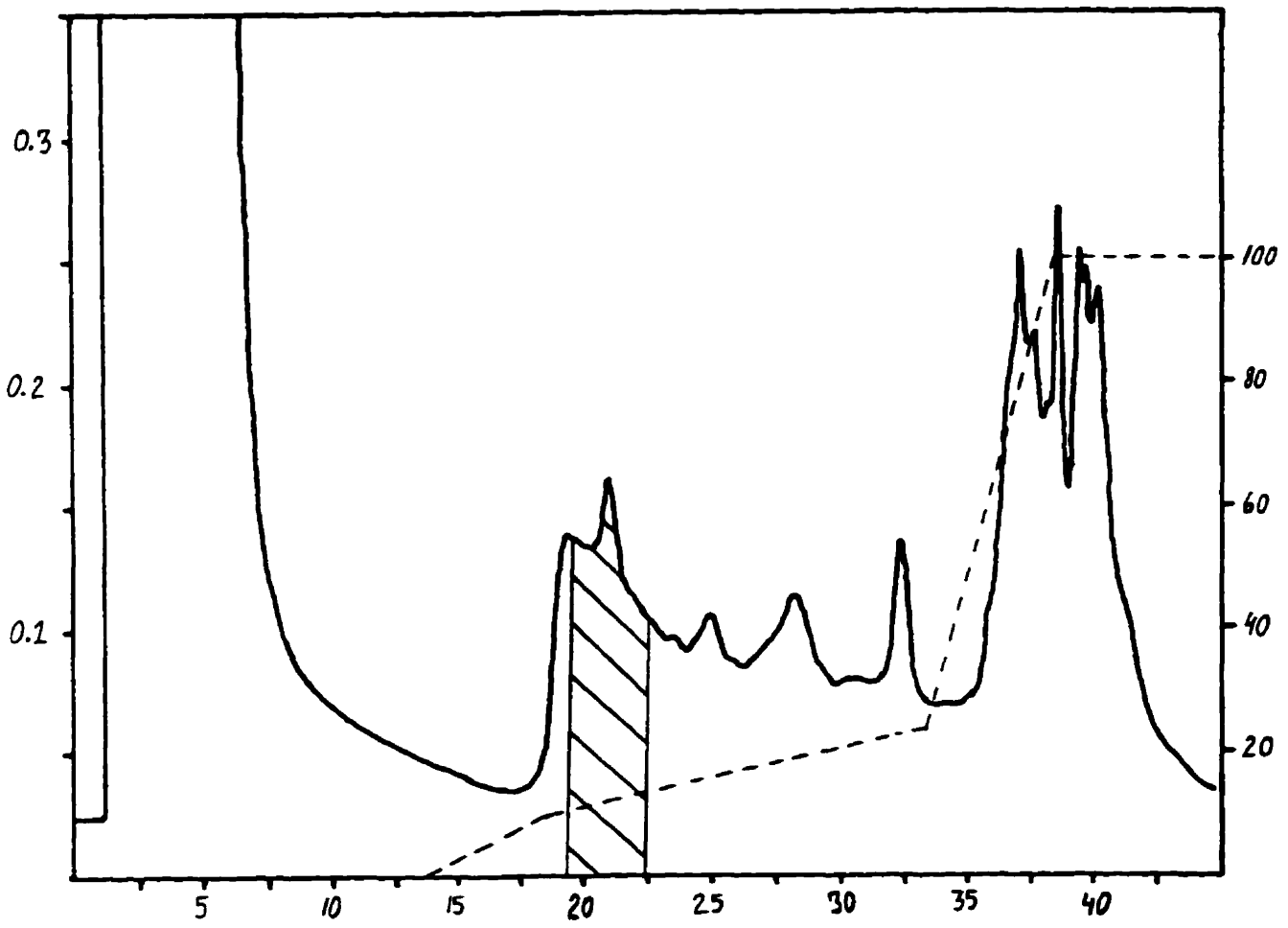


Figure 2.10. Mono-S chromatogram from purification of LaL from *E. coli* OR1265/pLcIR18 (0.5 M NaCl, 3 hr post induction). Shaded area represents the active fraction collected.

Sample:	55% ammonium sulfate fraction (Dialysed)
Column:	Mono-S HR 10/10
Initial Buffer:	Buffer A
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.0 M KCl (w/v) in Buffer A
Flow Rate:	2.0 mL/min
Left Axis:	Absorbance at 280 nm, indicated with the solid line
Right Axis:	Percentage Buffer B, indicated with the dashed line
Bottom Axis:	Time in minutes
Gradient:	0-10 % Buffer B over 5 min
	10-25% Buffer B over 15 min
	25-100% Buffer B over 5 min

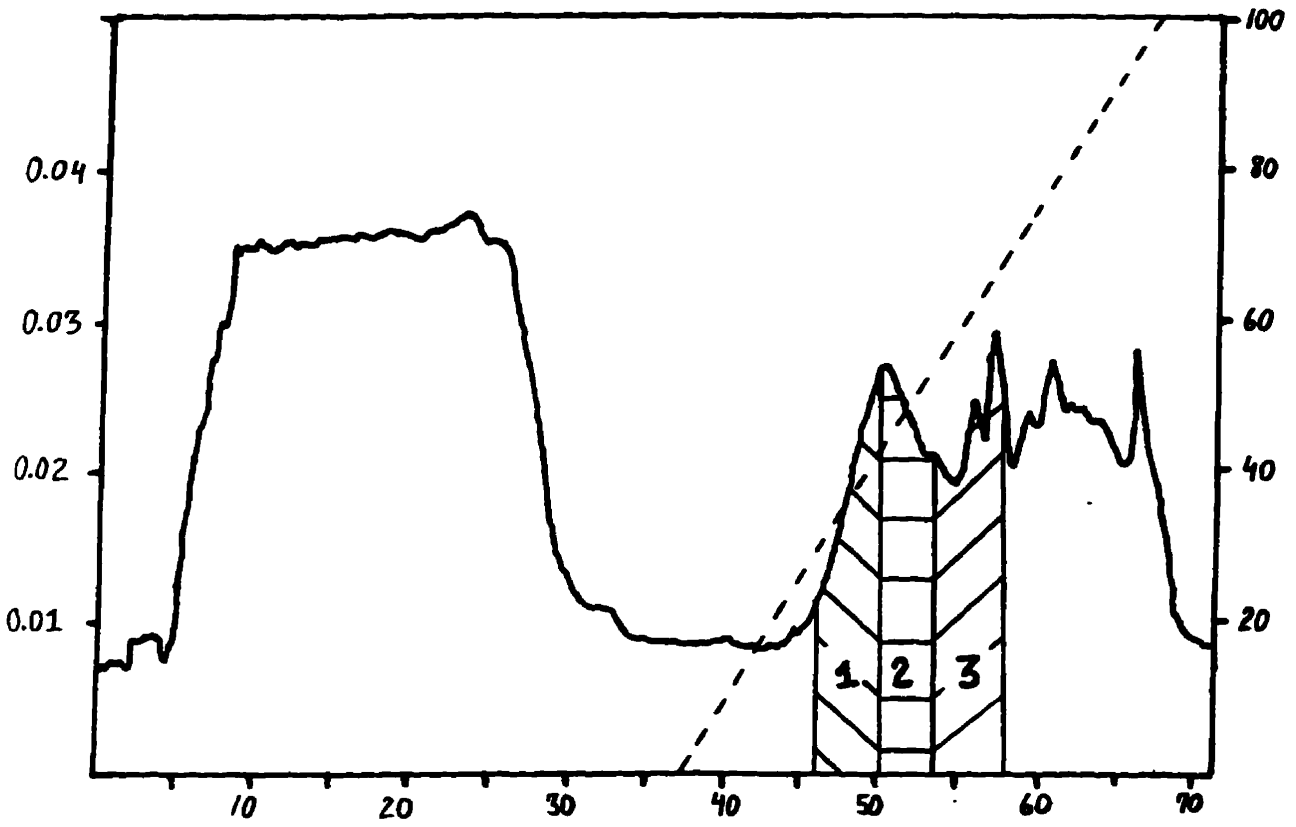


Figure 2.11. Phenyl-Superose chromatogram from purification of LaL from *E. coli* OR1265/pLcIR18 (0.5 M NaCl, 3 hr post induction). Shaded areas represent the active fractions collected.

Sample:	Mono-S fraction, adjusted to 1.2 M $(NH_4)_2SO_4$
Column:	Phenyl-Superose HR 5/5
Initial Buffer:	Buffer B
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.2 M ammonium sulfate (w/v) in Buffer A
Flow Rate:	0.5 mL/min
Left Axis:	Absorbance at 280 nm, indicated with the solid line
Right Axis:	Percentage Buffer A, indicated with the dashed line
Bottom Axis:	Time in minutes
Gradient:	0-100% Buffer A over 30 min

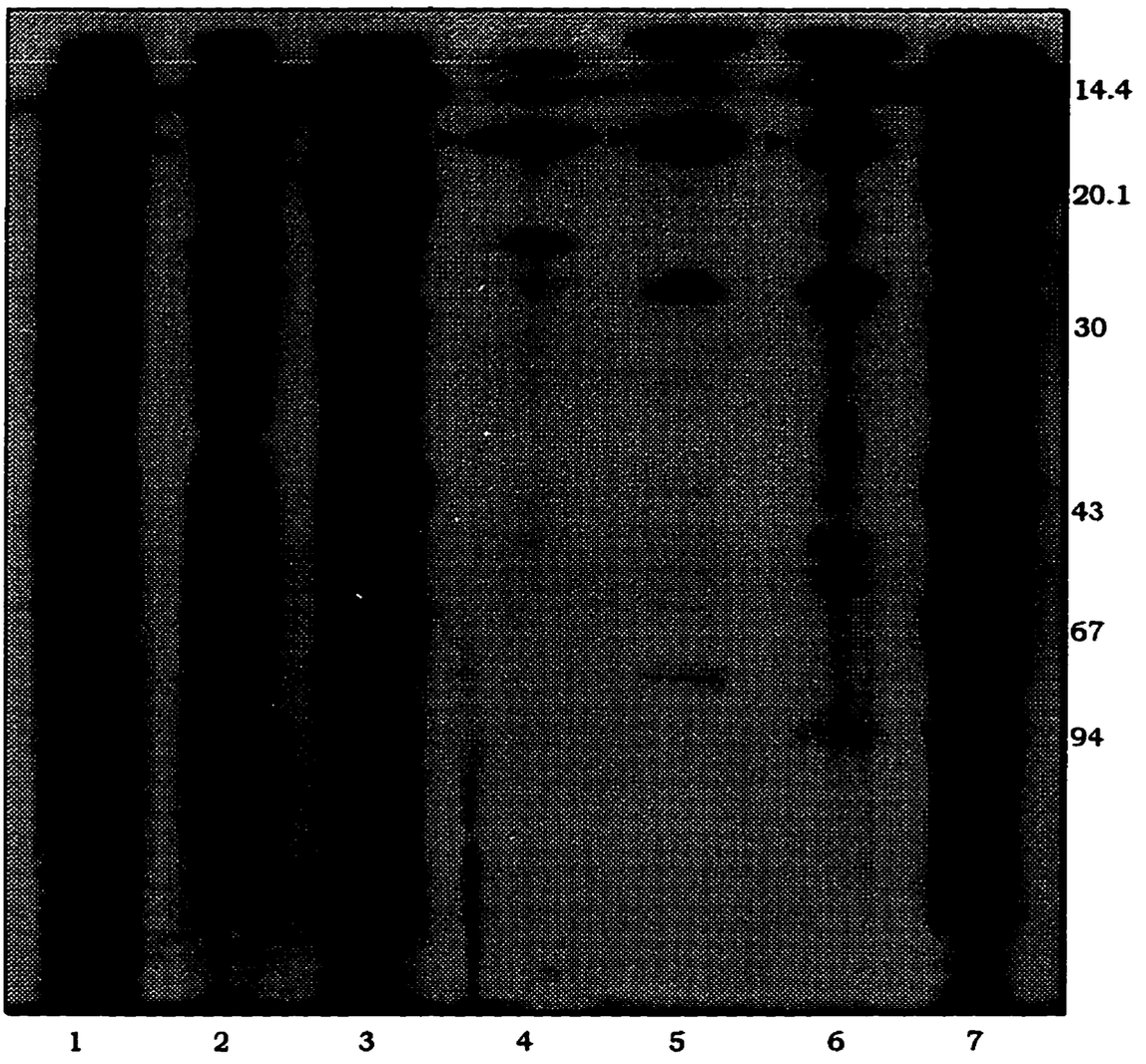


Figure 2.12. SDS-PAGE analysis of the stages of purification of LaL from *E. coli* OR1265 harboring pLcIR18 (0.5 M NaCl LB, 3 hr post induction). Protein samples were electrophoresed on a Gradient 10-15% PhastGel followed with silver staining. The arrow heads indicate the position of LaL. The figure is best viewed at arm's length.

Lane:

1: 55% Ammonium sulfate fraction (dialysed)
2: Mono-S fraction
3: Protein Standards (in kDa)
4: Phenyl-Superose fraction 1
5: Phenyl-Superose fraction 2
6: Phenyl-Superose fraction 3

} Refer to Fig. 2.11 for the respective fractions.

7: Protein Standards (in kDa)

intention of decreasing the load of adsorbed protein that would elute during the gradient, the ammonium sulfate concentration in the binding buffer was reduced to 1.2 M (as opposed to 1.7 M used previously). The extent of purification of intermediates obtained from the 0.5 M NaCl/3 hr post-induction growth period is shown in Fig. 2.12.

The thermal conditions required for protein induction would be accompanied with induction of heat shock proteins as well as a corresponding repression and/or degradation of other cellular proteins, i.e. the heat shock response. Because of this different cellular protein composition and possibly for other reasons, LaL could not be purified using the protocol developed for TG-1/pLR7. In addition, expression of LaL in this system was very low which may be a consequence of i) inability to produce a rapid and uniform temperature shift which is needed for efficient induction but is cumbersome to achieve with large volumes, ii) proteolysis of LaL by the La protease or iii) some other factor limiting expression.

Similar vectors have been used previously in *E. coli* to overproduce other proteins (Yoakum et al., 1982; Young et al., 1983) while others have also described unexpectedly low expression levels achieved with the λ P_L promoter using heat induction with the λ *cI*857 repressor (Mott et al., 1985). The pCM101 directed expression of the λ DNA terminase subunits resulted with gp*Nu1* and gp*A* representing 13% and 7% of the total protein following a 45 min post-induction period (Chow et al., 1987). However, in this construct, the wild-type ribosome binding sites of each gene were substituted with more efficient ones. The *R* gene RBS still mediates translational initiation from the pLcIR18 derived mRNA transcript, and it appears that the low expression levels could also be due to inefficient translation.

No appreciable difference was observed for purification from cells grown in the presence of the two NaCl concentrations or post-induction incubation periods. The amount of total protein obtained following the attempted purification was less than 0.2 mg/L (by Bradford analysis) in each case. No further studies with this system were investigated.

2.3.1.3. Construction of pHDM10 and Purification of LaL from E. coli TG-1/pHDM10

Following the unsuccessful results obtained with the pLcIR18 system, our attention again focused on the pTTQ18 vector. As illustrated in Fig. 2.13, the *R* gene was isolated in a 546 bp *Bam*HI fragment and ligated into the *Bam*HI site of pTTQ18 giving pHDM10. The expression host chosen for pHDM10 was again *E. coli* TG-1.

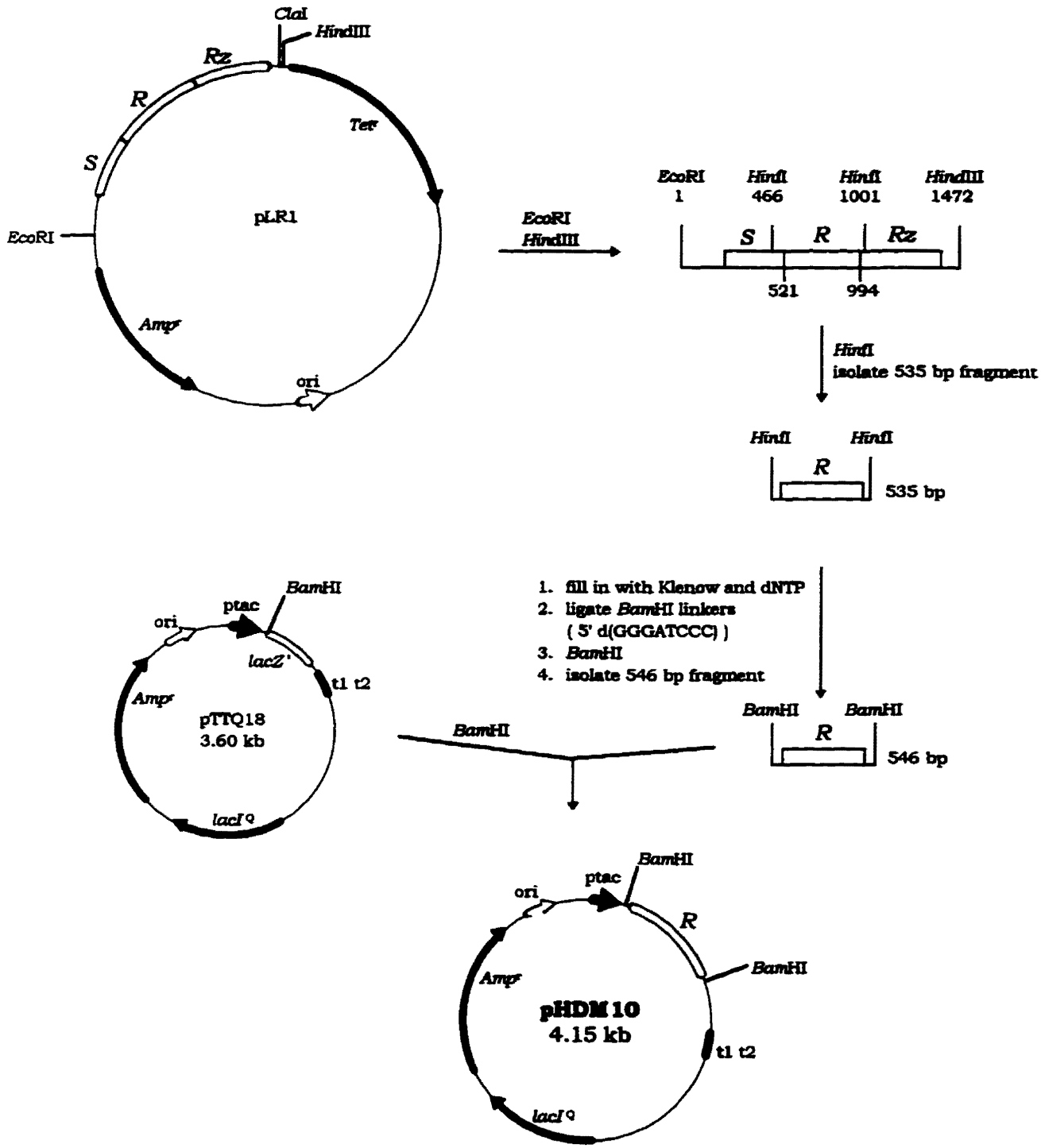


Figure 2.13. Construction of plasmid pHDM10. The *R* gene was isolated from pLR1 in a 1.47 kb *Eco*RI/*Hin*dIII fragment. This fragment was further digested with *Hin*fI, *Bam*HI linkers introduced and ligated into the *Bam*HI site of pTTQ18 giving pHDM10. Symbols: ptac, *tac* promoter; t1 t2, *rrnB* terminator fragment; *lacZ'*, *lacZα* coding region.

As with pLR7, the insert DNA fragment of pHDM10 is under transcriptional control of the strong hybrid *tac* promoter and transcriptional repression during initial cell growth is mediated by the plasmid encoded *lacI*^{Q} allele of the lac repressor gene. Derepression is achieved with the inclusion of IPTG in the growth media. In both pLR7 and pHDM10, the *R* gene mRNA transcript includes a portion of the *S* gene which precedes the *R* gene. In essence, the *R* gene is transcribed as part of a "dual" cistron in which the mRNA produced consists of the portion of the *S* gene and the complete *R* gene (in pLR7, the *R* gene is transcribed as part of a three gene unit that also includes the *Rz* gene). Taking advantage of the *Hin*dIII restriction sites upon construction of pHDM10 resulted in a reduction in the amount of DNA present between the pTTQ18 encoded RBS and the *R* gene (Fig. 2.14) compared to pLR7 (Fig. 2.4). Taking into account the remaining *S* gene DNA (55 bp) and the added *Bam*HI linker (6 bp), cloning of the λ DNA fragment into the *Bam*HI site of pTTQ18 places the *R* gene start codon 96 bp from the pTTQ18 encoded SD sequence as compared to 248 bp in pLR7. This resulted in a dramatic increase in the levels of LaL expressed from *E. coli* TG-1 harboring pHDM10. The amount of purified LaL obtained per litre of culture increased to 6.5 ± 0.6 mg from *E. coli* TG-1/pHDM10 as opposed to 0.2-0.3 mg for *E. coli* TG-1/pLR7 (a 20-30 fold increase).

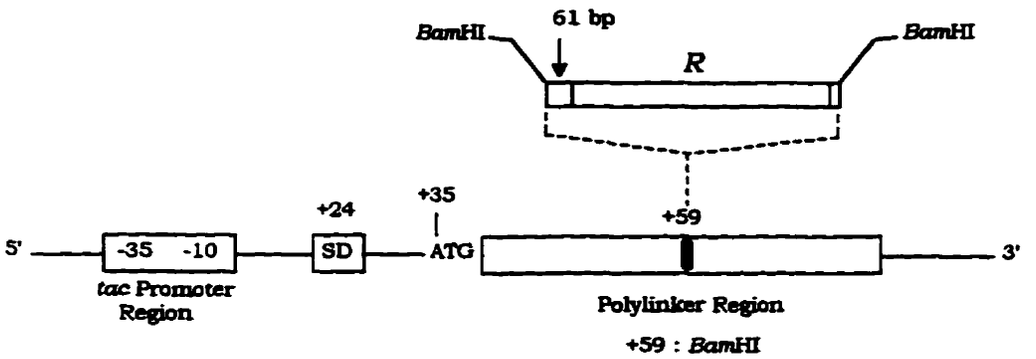


Figure 2.14. Schematic of the insertion mode of the λ DNA fragment of pHDM10 into the *Bam*HI site of pTTQ18. The 61 bp preceding the *R* gene comes from the remaining portion of the *S* gene (55 bp) and the *Bam*HI linker (6 bp) introduced.

As was described previously, the *R* gene does not have an optimal natural RBS. In pHDM10, the reduced distance of the *R* gene from the promoter region may result in enhanced translational efficiency of the *R* gene from what has been described as a two-cistron expression system (Schoner et al., 1987; Schoner et al., 1990). In this system,

ribosomes are assumed to bind to the RBS of the first cistron, initiate translation at the AUG codon, and translate the first cistron. Following termination of translation at the end of the first cistron, these ribosomes may not dissociate, but rather reinitiate translation at the AUG codon of the second cistron and proceed with translation of the second cistron. The stop codon of the first cistron should be close to the restart codon of the second cistron and a second SD sequence should be present in front of the restart codon. To be successful, the first cistron should code for an RBS for efficient initiation and therefore insure an increased local concentration of ribosomes for the second cistron RBS (Schoner et al., 1990).

In both pLR7 and pHDM10, the first cistron begins with the pTTQ18 encoded start codon and includes the polylinker region preceding the cloning (insertion) site of the λ DNA fragment and that portion of the remaining *S* gene (see Figs. 2.4 and 2.14) and translation would proceed to the *S* gene stop codon. Other requirements for the two-cistron expression system are met. The first cistron has a strong RBS (encoded by pTTQ18, derived from pUC18) and the stop codon of the first cistron (i.e. the *S* gene stop codon) is only 14 nucleotides downstream from the restart start codon of the *R* gene (Fig. 2.8). As such, it is possible that translation of the *R* gene from both pHDM10 and pLR7 occurs in this manner. However, in pLR7, the amount of DNA (248 bp, 219 bp from the *S* gene) between the pTTQ18 SD sequence and the *R* gene start codon may increase the possibility of undesirable mRNA elements being present. Although no analysis of the *S* gene was performed, it has been established that runs of certain codons can decrease synthesis of a protein by causing the ribosomes to stall on the transcript or even to dissociate (Robinson et al., 1984). In pHDM10, the decreased DNA present upstream from the *R* gene (96 bp, 55 bp from the *S* gene) could reduce the chance of such elements. One would expect similar expression levels from pLR7 and pHDM10 if translation for both was initiated with only the wild-type *R* gene RBS. It is expected that translation is initiated in both plasmids by a combination (but to different extents) of initiation by the two-cistron model as well as initiation at the natural *R* gene RBS. However, the 20-30 fold increase in LaL purified from pHDM10 suggests that translational initiation of the *R* gene by the two-cistron model described is more efficient for pHDM10 than for pLR7.

Several aspects of the purification protocol of LaL from TG-1/pHDM10 (which was typically performed on a 6 L culture scale) were modified from that developed for TG-1/pLR7. Firstly, it was observed (for both strains) that during recovery after disruption of cells, a considerable amount of LaL activity was found in the cellular debris pellets,

presumably from adsorption of LaL to the bacterial wall fragments or by "trapping" of the enzyme in the cellular debris. To circumvent this loss in enzyme, cells were subjected to a single French press passage. Following centrifugation, the supernatant was reserved and the cell pellet was resuspended in buffer and subjected to a second French press passage. This was done (as opposed to 2 successive French press passages without an intervening centrifugation) to recover enzyme from the first cell debris pellet in the second volume of buffer added for the second passage. Although there was still LaL activity associated with the ensuing second cellular debris pellet, the amount was reduced which improved the recovery of LaL following cell disruption. Although not explored, it may be possible to recover additional enzyme from the cellular debris if a buffer with a higher concentration of salt is used to reduce interactions of LaL with cellular components.

When the cell free extract was subjected to ammonium sulfate precipitation, inconsistencies were observed. As was discussed for the purification of LaL from TG-1/pLR7, a practical concentration of ammonium sulfate affording optimal recovery of LaL could not be achieved. To avoid this inefficient procedure, an alternative was pursued for the initial purification step. It was possible to subject the cell free extract directly to Mono-S chromatography. However, for practical considerations, this was not performed. Firstly, it is difficult and tedious to filter large volumes of a concentrated cell extract sample (even following extensive centrifugation to remove particulate matter) which is an absolute requirement of sample preparation for chromatography using any commercially available pre-packed column. Secondly, the longevity and performance of the Mono-S column would be compromised by the irreversible adsorption of material from such a crude sample, an expensive consequence. In one instance, a small sample of the cell free extract (≈ 5 mg total protein) was directly chromatographed over the Mono-S column. LaL could be subsequently purified by chromatography of the active fraction over Phenyl-Superose demonstrating that a two column procedure could be employed to purify LaL from TG-1/pHDM10.

Purification of LaL from the cell free extract was initiated by cation exchange chromatography over S-Sepharose Fast Flow packed in a HR 16/50 column. The column can hold approximatey 100 mL of the S-Sepharose resin which has a loading capacity of approximately 50-75 mg of protein/mL. This chromatography proved to be very advantageous. Firstly, the crude extract could be applied without any reservation. If the sample was "dirty", material collected at the top of the column and the contaminated resin could be easily and inexpensively replaced. Secondly, large volumes and amounts of

protein could be loaded in a single chromatographic run, with elution of LaL during application of the salt gradient (Fig. 2.15) in a desirable volume for subsequent purification. Although this column can withstand high flow rates (> 5 mL/min), the sample was applied using a slower rate of 2.5 mL/min. A very minimal, yet detectable amount of lysozyme activity was found to elute during application of the sample in the early fractions (i.e. the void volume). This leaching was minimized using the reduced flow rate. This chromatographic step was also a very efficient first purification stage, not only in its capacity to eliminate > 90% of the contaminating cell free extract protein, but also to provide a favorably clean sample for the ensuing Mono-S chromatography.

Following dialysis of the S-Sepharose Fast Flow fraction, purification was continued by Mono-S HR 10/10 chromatography. The salt gradient was altered slightly from that used previously, and LaL eluted at approximately 0.11 M KCl (Fig. 2.16). The suggested manufacturer's loading capacity for the Mono-S HR 10/10 column (8 mL) is 20-50 mg/mL and 40 mg per single peak. For this reason, the dialysed S-Sepharose Fast Flow fraction (with an average content of 124 ± 24 mg protein determined from 7 individual purifications on a 6 L culture scale) was applied in two individual runs. In this manner, the fraction from each run containing LaL comprised approximately 30 mg of protein.

The pooled Mono-S fraction was made to 1.7 M ammonium sulfate and was then subjected to Phenyl Superose HR 10/10 chromatography. This column (8 mL) has a suggested loading capacity of 5-10 mg/mL (40-80 mg total), and the chromatography of the Mono-S fraction was again performed in two separate runs. LaL eluted during the decreasing salt gradient at approximately 1.0 M ammonium sulfate (Fig. 2.17). The pooled Phenyl-Superose fraction was made to 1.7 M ammonium sulfate and was subjected to a second chromatography over the Phenyl-Superose column and eluted with a sharp gradient (Fig. 2.18). This was done to concentrate the protein sample into a smaller volume of buffer to reduce the amount of buffer salts that would be carried over following exhaustive dialysis and lyophilization of this final fraction (section 2.2.10.2).

The progression of LaL purification from TG-1/pHDM10 is summarized in Table 2.2 and illustrated in the SDS-PAGE analysis of the stages of purification (Fig. 2.19). The values in Table 2.2 are a very subjective assessment of the stages of purification. Firstly, unfortunate inherent limitations associated with the turbidimetric assay employing substrate cells presents a degree of uncertainty in activity measurements. Also, because there exists a "batch to batch" variation in the performance or response of the substrate

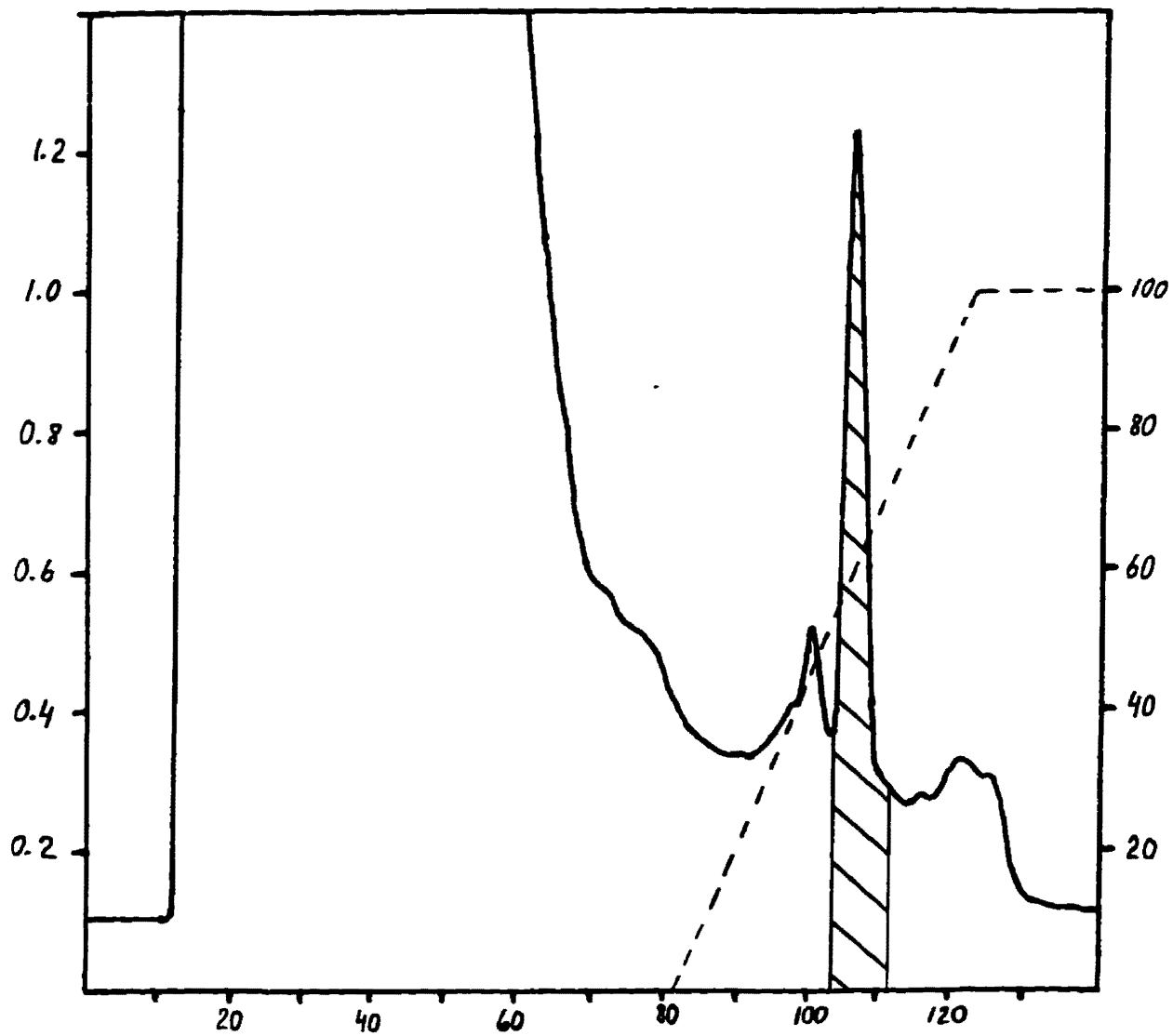


Figure 2.15. S-Sepharose Fast Flow chromatogram from purification of LaL from *E. coli* TG-1/pHDM10. The shaded area represents the active fraction collected.

Sample:	Cell free extract (from a 6 L culture)
Column:	S-Sepharose Fast Flow packed in an HR 16/50 column
Initial Buffer:	Buffer A
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.0 M KCl (w/v) in Buffer A
Flow Rate:	2.5 mL/min during sample application (before ≈ 80 min)
	5.0 mL/min during gradient application (after ≈ 80 min)
Left Axis:	Absorbance at 280 nm, indicated with the solid line
Right Axis:	Percentage Buffer B, indicated with the dashed line
Bottom Axis:	Time in minutes
Gradient:	0-100 % Buffer B over 40 min

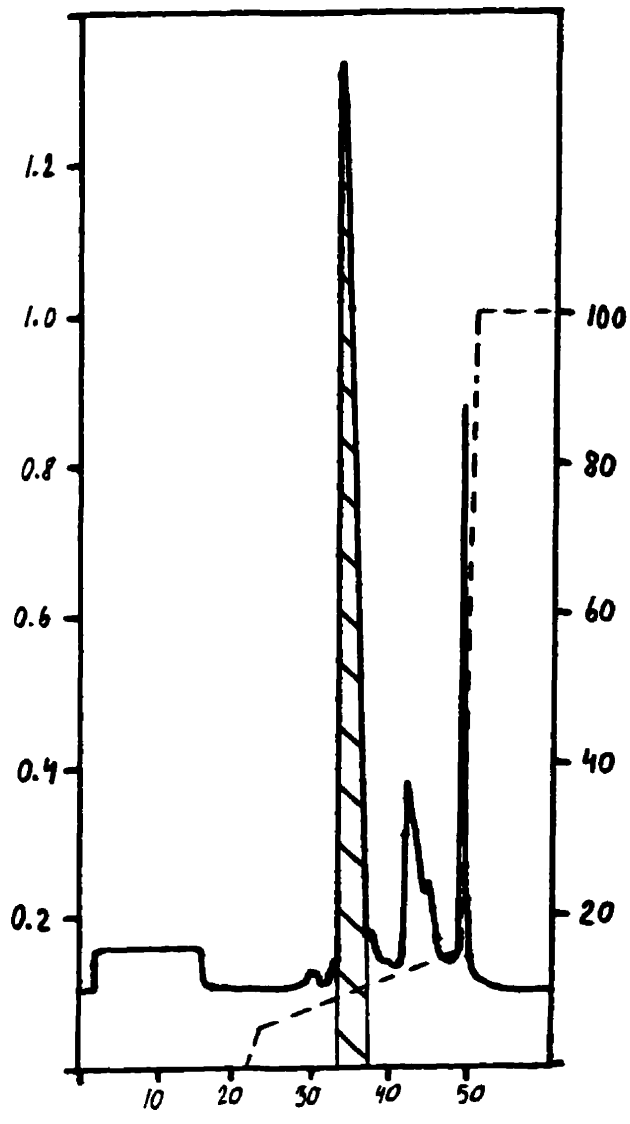


Figure 2.16. Mono-S chromatogram from purification of LaL from *E. coli* TG-1/pHDM10. The shaded area represents the active fraction collected.

Sample:	One-half of the S-Sepharose Fast Flow fraction
Column:	Mono-S HR 10/10
Initial Buffer:	Buffer A
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.0 M KCl (w/v) in Buffer A
Flow Rate:	2.0 mL/min
Left Axis:	Absorbance at 280 nm, indicated with the solid line
Right Axis:	Percentage Buffer B, indicated with the dashed line
Bottom Axis:	Time in minutes
Gradient:	0-5 % Buffer B over 2 min
	5-15% Buffer B over 25 min
	15-100% Buffer B over 2 min

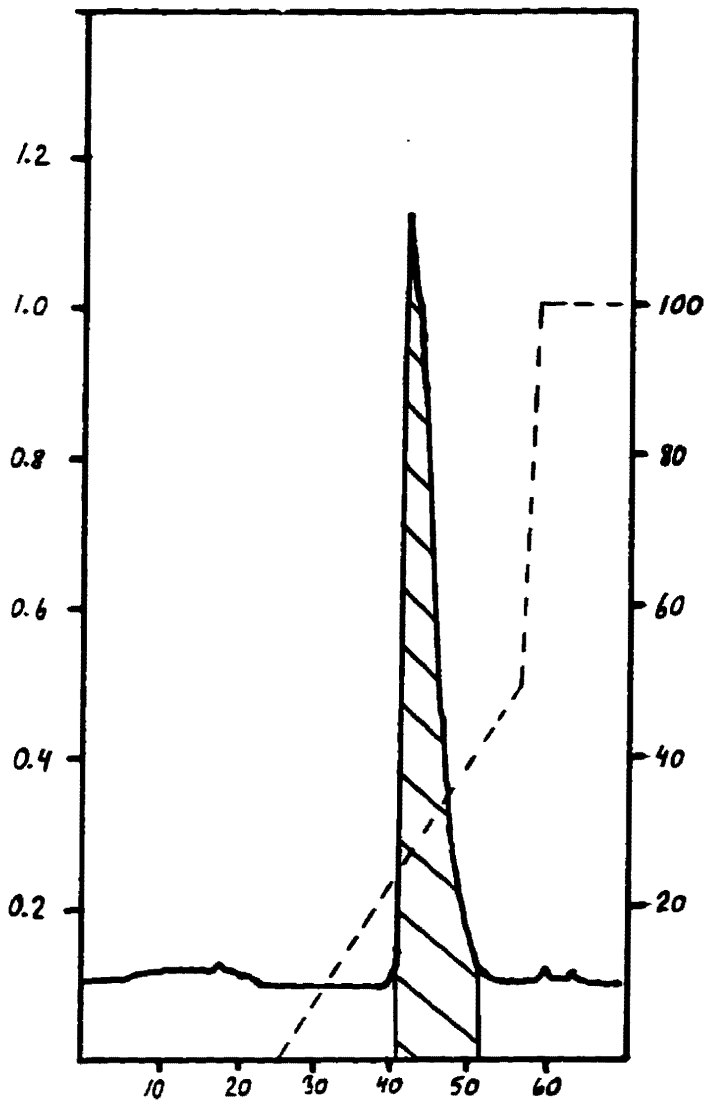


Figure 2.17. Phenyl-Superose chromatogram from purification of LaL from *E. coli* TG-1/pHDM10. The shaded area represents the active fraction collected.

Sample:	One-half of the pooled Mono-S fractions, adjusted to 1.7 M $(NH_4)_2SO_4$
Column:	Phenyl-Superose HR 10/10
Initial Buffer:	Buffer B
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.7 M ammonium sulfate (w/v) in Buffer A
Flow Rate:	1.0 mL/min during sample application (before ≈ 25 min)
	1.5 mL/min during gradient application (after ≈ 25 min)
Left Axis:	Absorbance at 280 nm, indicated with the solid line
	The scale is 10× the sensitivity shown before 25 min
Right Axis:	Percentage Buffer A, indicated with the dashed line
Bottom Axis:	Time in minutes
Gradient:	0-50% Buffer A over 30 min
	50-100% Buffer A over 3 min

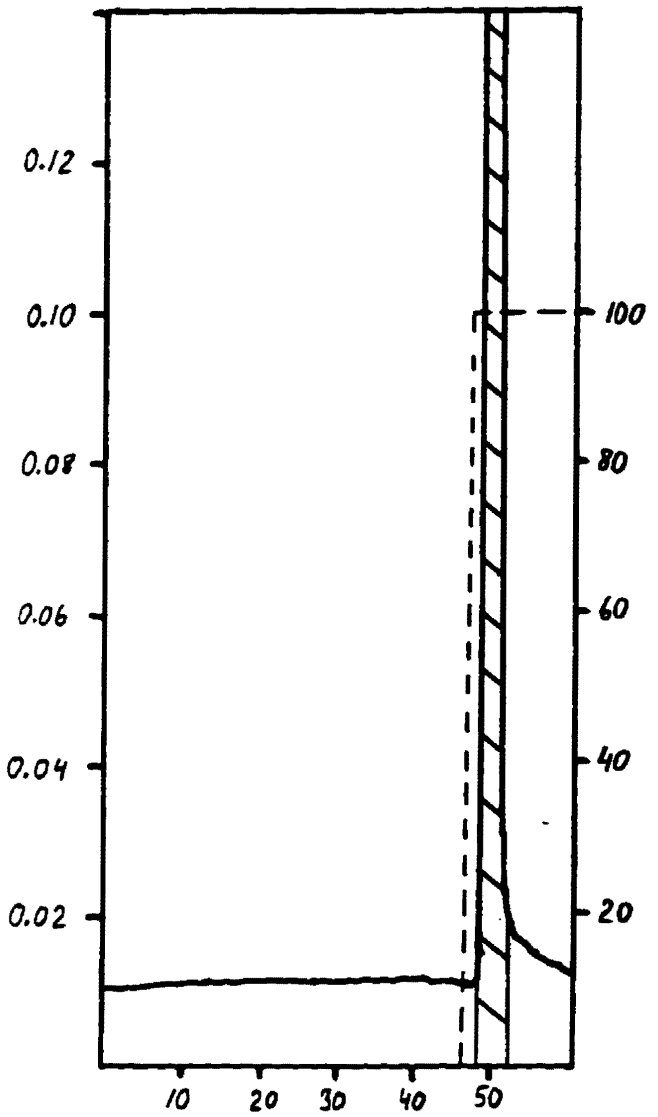


Figure 2.18. Phenyl-Superose chromatogram for the concentration of LaL from *E. coli* TG-1/pHDM10. Shaded area represents the active fraction collected.

Sample:	Pooled first Phenyl-Superose fraction, made to ≈ 1.7 M $(NH_4)_2SO_4$
Column:	Phenyl-Superose HR 10/10
Initial Buffer:	Buffer B
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.7 M ammonium sulfate (w/v) in Buffer A
Flow Rate:	1.0 mL/min during sample application (before ≈ 46 min)
	1.5 mL/min during gradient application (after ≈ 46 min)
Left Axis:	Absorbance at 280 nm, indicated with the solid line
Right Axis:	Percentage Buffer A, indicated with the dashed line
Bottom Axis:	Time in minutes
Gradient:	0-100% Buffer A over 0.5 min

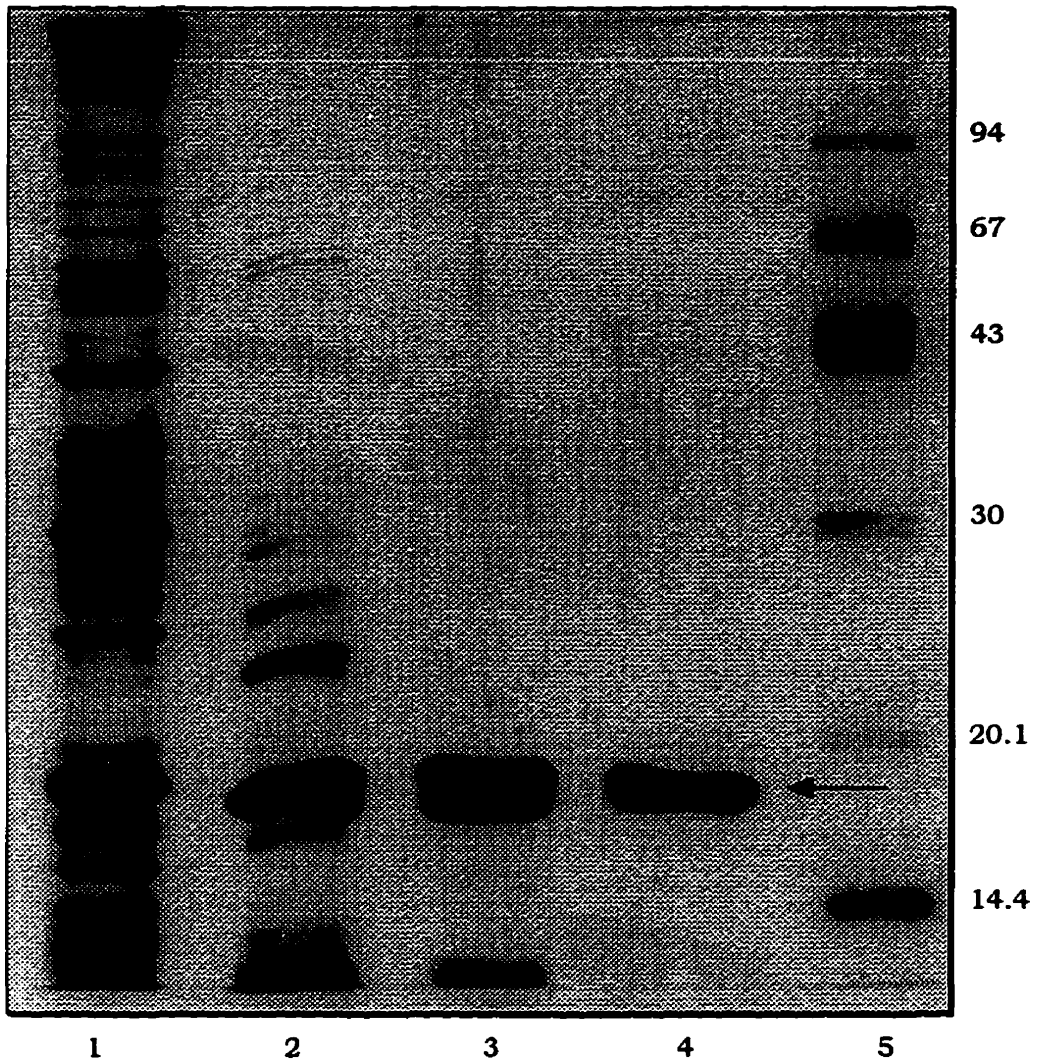


Figure 2.19. SDS-PAGE analysis of the stages of purification of LaL from *E. coli* TG-1 harboring pHDM10. Protein samples (≈ 2 μg/lane) were electrophoresed on a 15% T gel followed with silver staining. The arrow indicates the position of LaL.

Lane:

1: Cell free extract
2: S-Sepharose Fast Flow fraction (dialysed)
3: Mono-S fraction
4: Phenyl-Superose fraction
5: Protein Standards (in kDa)

Table 2.2. Representative purification of LaL from a 6 L culture growth scale of *E. coli* TG-1/pHDM10.

Fraction (Stage of Purification)	Volume (mL)	Total Protein (mg)[a]	Specific Activity[b]	Enzyme Recovery (%)	Purification Fold
Cell Free Extract	86.0	1606	59 ± 5	100	1
S-Sepharose Fast Flow					
(1) column fraction	45.0	128.3	690 ± 42	93	12
(2) after dialysis	47.5	110.7	762 ± 69	89	13
Mono-S	16.4	63.1	1088 ± 93	73	18
Phenyl-Superose					
(1) Pooled 1st column fraction	31.2	47.6	nd.[c]		
(2) after 2nd column (concentrated fraction)	5.0	50.2	nd.[c]		
(3) after dialysis	7.3	58.2	752 ± 70	46	13
(4) after lyophilization	-	39.9	1098 ± 102	46	19

[a] Determined by Bradford analysis (2.2.10.1) except for the Phenyl-Superose fraction (4) determined by mass of the lyophilized powder following correction for buffer salts (2.2.10.2).

[b] expressed in units·μg^{-1}. In this case, the unit is arbitrarily defined as the amount of sample causing a decrease in OD_{600} of 0.01/min of *E. coli* substrate cells using the optimized assay described in Chapter 3.

[c] not determined as the concentration of ammonium sulfate in these samples would cause some inhibition of the lysis of the substrate cells due to ionic strength effects (see 3.3.1).

NOTE: Calculations of enzyme recovery and purification fold values for each stage of purification require the specific activity determined for the cell free extract. In addition to LaL expressed from pHDM10, the cell free extract will also contain endogenous soluble *E. coli* peptidoglycan degrading enzymes. The activity of these enzymes would contribute to increasing the specific activity of this fraction. The activity of the subsequent fractions more accurately reflect the specific activity attributable to LaL since it is expected that the endogenous enzymes would be progressively eliminated with purification. As such, the enzyme recovery and purification values for the subsequent fractions are somewhat lower (see below) than would be obtained if the sole activity detected in the cell extract was due to LaL.

Unfortunately, suitable controls to determine the contribution of measured activity of these endogenous enzymes were not performed. However, in earlier experiments involved in the preparation of pLR7, controls were performed. Control samples of wild type *E. coli* TG-1 grown under inducing conditions and *E. coli* TG-1 cells harboring pLR7 grown under non-inducing conditions typically displayed 15-20% of the activity of cells harboring pLR7 grown under induction. Considering that a 20-30 fold increase in LaL was obtained from pHDM10 over pLR7, the contributing activity from endogenous enzymes in the former cell free extract would therefore be less than 1%.

cells, the definition of a "unit" of activity is conditional. For example, defining a unit as a 0.01/min absorbance unit decrease, different amounts of LaL would be required to achieve this unit for different batches of substrate cells. Therefore, an absolute definition of "unit" is not possible (this is made more clear in Chapter 3 with reference to experimental data). Secondly, the total protein of each fraction was determined by Bradford analysis. This method exaggerated the quantity of *purified* LaL (discussed in 2.3.4) and is affected by buffer composition (compare Phenyl-Superose fractions (1) and (3) in Table 2.2 which contain ≈ 1 M and 0 M ammonium sulfate respectively). Because the fractions contain varying percentages of LaL to contaminating proteins, the total protein in each fraction will be overestimated to varying degrees. For this reason, the specific activities will also contain a hidden and conjectural error. Confidence is placed in the total protein quantitated by mass following dialysis and lyophilization (2.3.4). Despite these unavoidable shortcomings, the analysis does permit a qualitative evaluation of the purification procedure.

Based on the results from 7 separate purifications on a 6 L culture scale, the average yield of purified LaL obtained was 6.5 ± 0.6 mg/L of culture (determined by mass following lyophilization (section 2.2.10.2)). A total of 312 mg of LaL was purified from TG-1/pHDM10 over the course of this study following the development of the current protocol. All of the kinetic, inhibition and biophysical investigations conducted utilized LaL purified from this system.

2.3.1.4. Construction of pLR102 and Purification of LaL from E. coli BL21(λDE3)/pLR102

The last of the expression vectors created for the overexpression of LaL was pLR102. This plasmid was constructed using an *R* gene DNA fragment prepared using the polymerase chain reaction (PCR), an effective technique for not only amplifying a DNA fragment but also in generating one that is flanked with desired restriction sites.

A schematic of the PCR reaction used to amplify the λ *R* gene is shown in Fig. 2.20. The template chosen for PCR was plasmid pHDM10 and with the fortunate availability of the *R* gene sequence (Sanger et al, 1982), permitted the design of suitable primers. Primers for PCR were prepared following a strategy developed by Dr. G. Wright (McMaster University, Hamilton, ON) and were designed to introduce unique restriction sites. Primer 1 (with sequence complimentary to the *R* sense strand) introduced a *Kpn*I and a *Nde*I

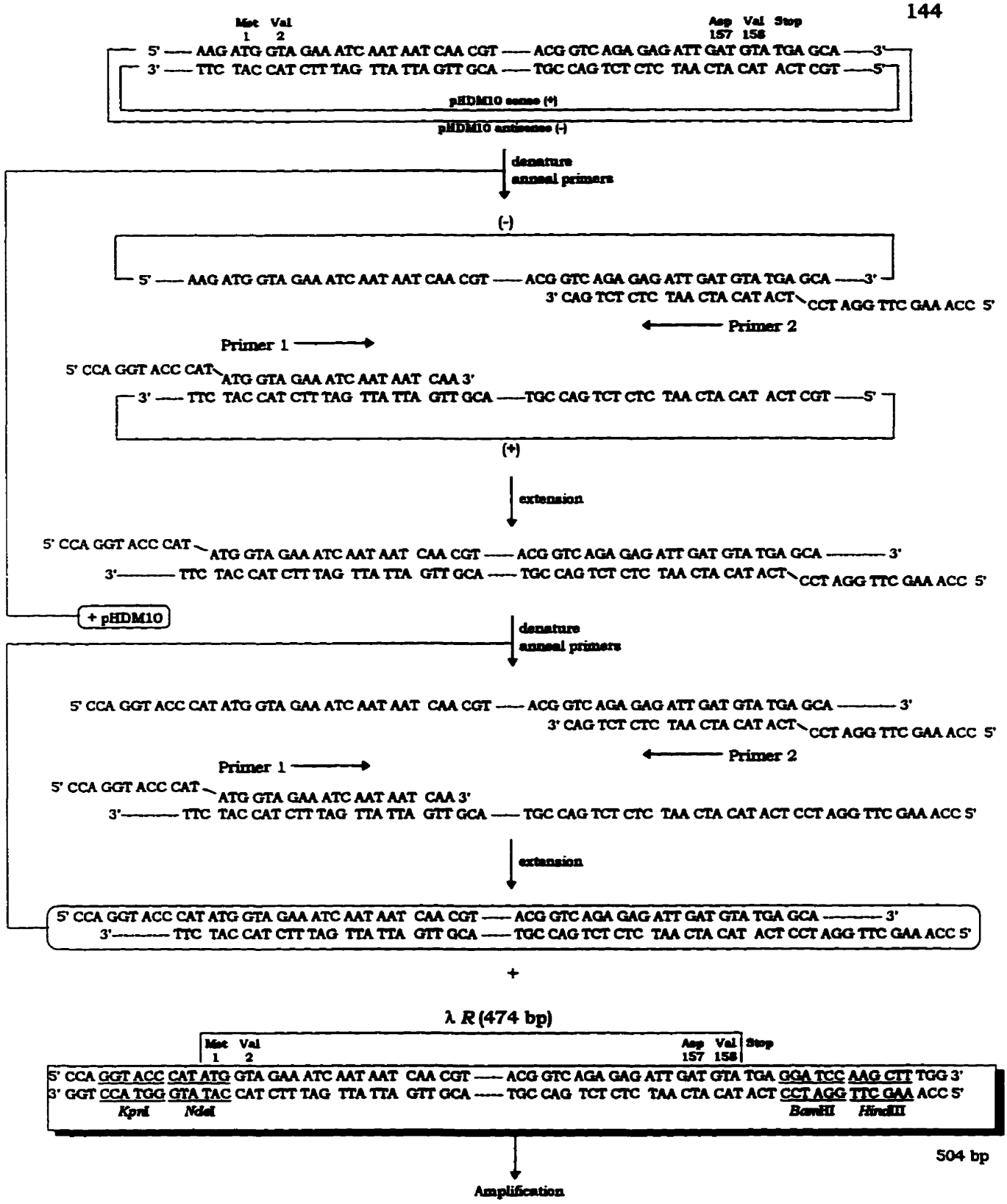


Figure 2.20. Schematic of the polymerase chain reaction for the amplification of the lambda *R* gene from pHDM10. Indicated are the synthetic primers employed (33 mer, primer 1; 36 mer, primer 2) and relevant sequences of the *R* gene to illustrate the complimentary regions with the primers. Also shown are the designed restriction sites in the amplified product.

restriction site (the *Nde*I site spans the *R* gene start codon) 5' to the *R* gene while primer 2 (with sequence complimentary to the anti-sense strand) included both a *Bam*HI and *Hin*dIII restriction site immediately 3' to the *R* gene stop codon. As illustrated in Fig. 2.20, two cycles of denaturation, primer annealing and extension will generate the desired product DNA fragment (enclosed in the box at the bottom of Fig. 2.20) which is subsequently amplified in the ensuing cycles of the PCR reaction.

The construction of pLR102 is illustrated in Fig. 2.21. The PCR product fragment was first incorporated into the *Kpn*I/*Hin*dIII sites of pUC18, allowing for selection of recombinant transformants (white colonies on X-gal media) and isolation in good yields of the recombinant plasmid. The *R* gene was further excised from pUC18/*R* with *Nde*I and *Bam*HI and ligated into the expression vector pET-22b yielding pLR102.

The host chosen for pLR102 and one that is optimized for expression from pET vectors was *E. coli* BL21(λDE3), hereafter referred to as BL21. This bacterial strain contains a chromosomal copy of the gene for T7 RNA polymerase that is under control of the inducible *lacUV5* promoter. The *R* gene in pLR102 is under transcriptional control of what is termed a T7*lac* promoter (Studier et al., 1990). This promoter contains a *lac* operator sequence immediately downstream from the T7 promoter sequence (Fig. 2.22). The pET22-b plasmid also encodes for the lac repressor (*lacI*) which represses transcription at both the *lacUV5* promoter on the chromosome and the T7promoter on the plasmid. Induction with IPTG will cause derepression at both promoters permitting expression of T7 polymerase which subsequently acts to transcribe the *R* gene on the plasmid from the T7 promoter.

Relevant regions of the transcription and expression region of plasmid pET22-b are illustrated in Fig. 2.22. Cloning of the 5' end of the *R* gene PCR product into the *Nde*I site optimally places the *R* gene start codon only 9 bp from the SD sequence. Plasmid pLR102 therefore possesses extremely favourable elements for transcription and translation of the *R* gene. Expression of LaL in *E. coli* BL21/pLR102 was very effective. The gel illustrated in Fig. 2.23 clearly suggests that induction of *E. coli* BL21/pLR102 resulted with the expression of LaL at levels comprising at least 10% of the total cellular protein. As was mentioned in section 2.2.5.5, two recombinant pET22-b plasmids, pLR102 and pLR103, were obtained that were able to direct expression of LaL in *E. coli* BL21 (Fig. 2.23); however only the characterization of pLR102 and of LaL purified from *E. coli* BL21/pLR102 was performed (section 2.3.2).

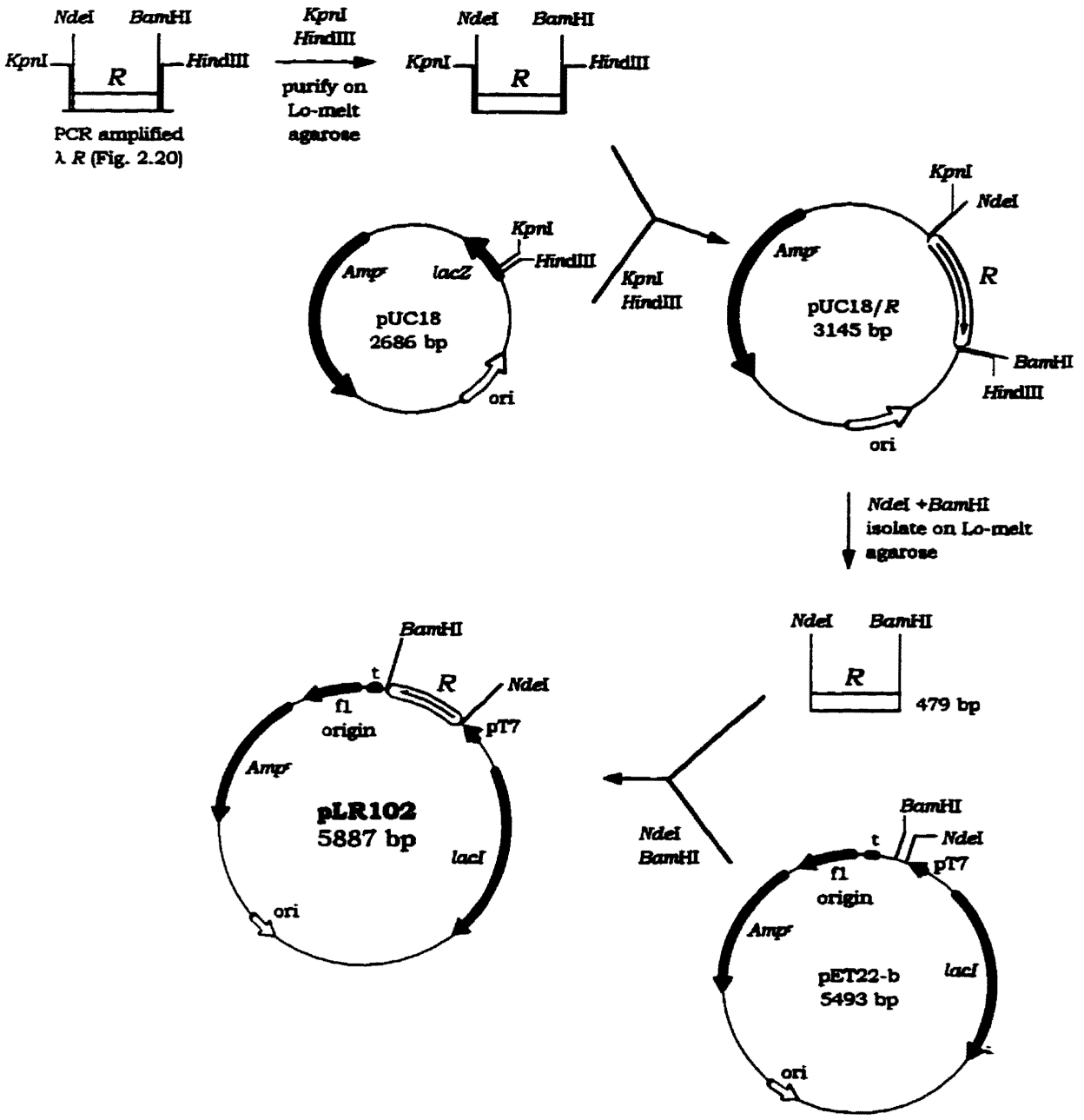


Figure 2.21. Construction of plasmid pLR102. The PCR amplified λ *R* gene was restricted with *Kpn*I and *Hind*III and ligated with plasmid pUC18 cleaved with the same enzymes. The *R* gene was recovered from pUC18/*R* in a 479 bp *Nde*I/*Bam*HI fragment and ligated into the *Nde*I and *Bam*HI sites of plasmid pET22b yielding pLR102.
Symbols: pT7, T7 promoter region; t, T7 terminator fragment.

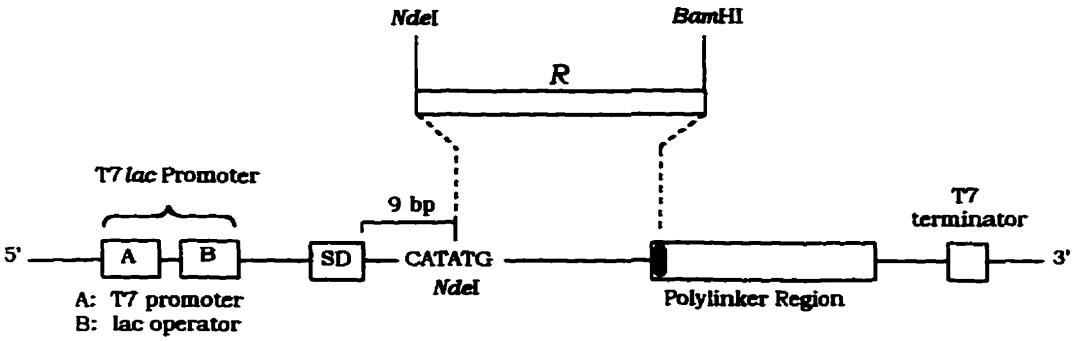


Figure 2.22. Schematic of the transcription and expression region of pET-22b. Shown is the mode of insertion of the *R* gene PCR product into the *Nde*I and *Bam*HI sites.

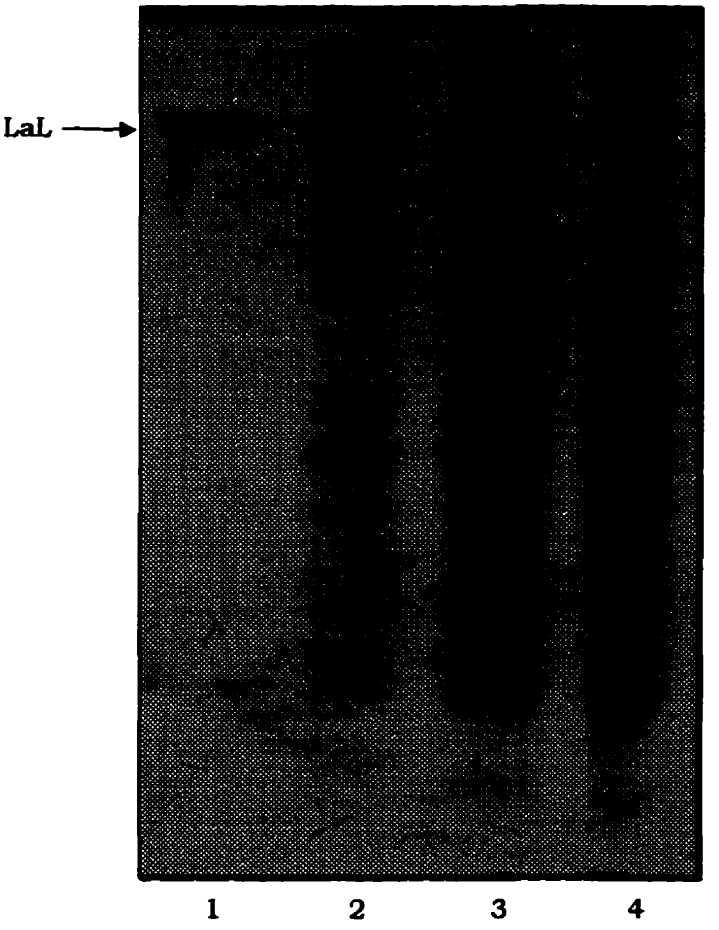


Figure 2.23. Expression of LaL directed by *E. coli* BL21(λDE3)/pLR102 and pLR103. Cultures were brought to early log growth, induced (or not induced) with IPTG (0.7 mM) and grown additionally for 2.5 h. Culture samples were withdrawn as described in 2.2.7 and were subjected to SDS-PAGE on a Gradient 10-15% PhastGel followed with Coomassie staining.

Lane: (1) Purified LaL; (2) induced BL21/pLR103; (3) induced BL21/pLR102; (4) non-induced BL21/pLR102.

Purification of LaL from BL21/pLR102 followed the same procedure used for TG-1/pHDM10 with only slight modifications. With the pLR102 system, cell disruption was achieved with the more convenient method of sonication. Sonication was attempted with the earlier systems but was not as effective as disruption by French press. It appears that the high cellular presence of LaL actually aided in the lysis of the cells during the sonication process. Sonication was performed in two stages for the same rationale as described previously for French press disruption (2.3.1.3) in order to minimize loss of LaL adsorbed to the cellular debris. That is, the harvested cells were subjected to a first round of sonication and centrifugation. The pellet obtained was suspended in a second volume of buffer and subjected to a second round of sonication and centrifugation.

Purification by column chromatography was again achieved by sequential chromatography of the cell free extract over S-Sepharose Fast Flow HR 16/50, Mono-S HR 10/10 and Phenyl-Superose HR 10/10 (see Figs. 2.24-2.26) bearing in mind the loading capacities of the latter two columns. To process protein from 6 L of cell culture, six individual runs of the dialysed S-Sepharose Fast Flow fraction were performed over the Mono-S column. As shown in Fig. 2.25 (A), LaL was eluted in both a major (a) and minor (b) fraction, with the major eluting typically over 9.3-10.8% Buffer B and the minor over 11.7-13.4% Buffer B. When the minor fractions from each run were pooled and re-chromatographed over the Mono-S column, LaL once again eluted in the same two peak profile (Fig. 2.25 (B)).

This phenomena where a single protein species elutes in two fractions has been observed previously in our laboratory using ion-exchange columns containing resin formulated with Mono beads (i.e. Mono-S, Mono-Q) even when the total protein loaded is far below the suggested capacity of the column. The chitinase from *Neurospora crassa* appeared to elute in two closely related fractions on a Mono-Q column (Mohd. Salleh, 1994). As well, experience with various pre-packed columns has frequently revealed that single species will elute in more than one fraction, perhaps suggesting lack of uniformity or deterioration of the packing material or the involvement of two or more physical separation modes operating during chromatography. These difficulties were not encountered with the Phenyl-Superose column at any time.

The progression of LaL purification from BL21/pLR102 is illustrated in the SDS-PAGE analysis of the purification intermediates (Fig. 2.27). All the column fractions appear homogeneous for LaL if stained by Coomassie. Although not apparent in the scanned image of the silver stained gel photograph, the original gel did reveal contaminating protein bands in both the S-Sepharose and Mono-S fractions. As seen in Fig 2.26 (A), the Phenyl-Superose column still removes contaminating protein that eluted

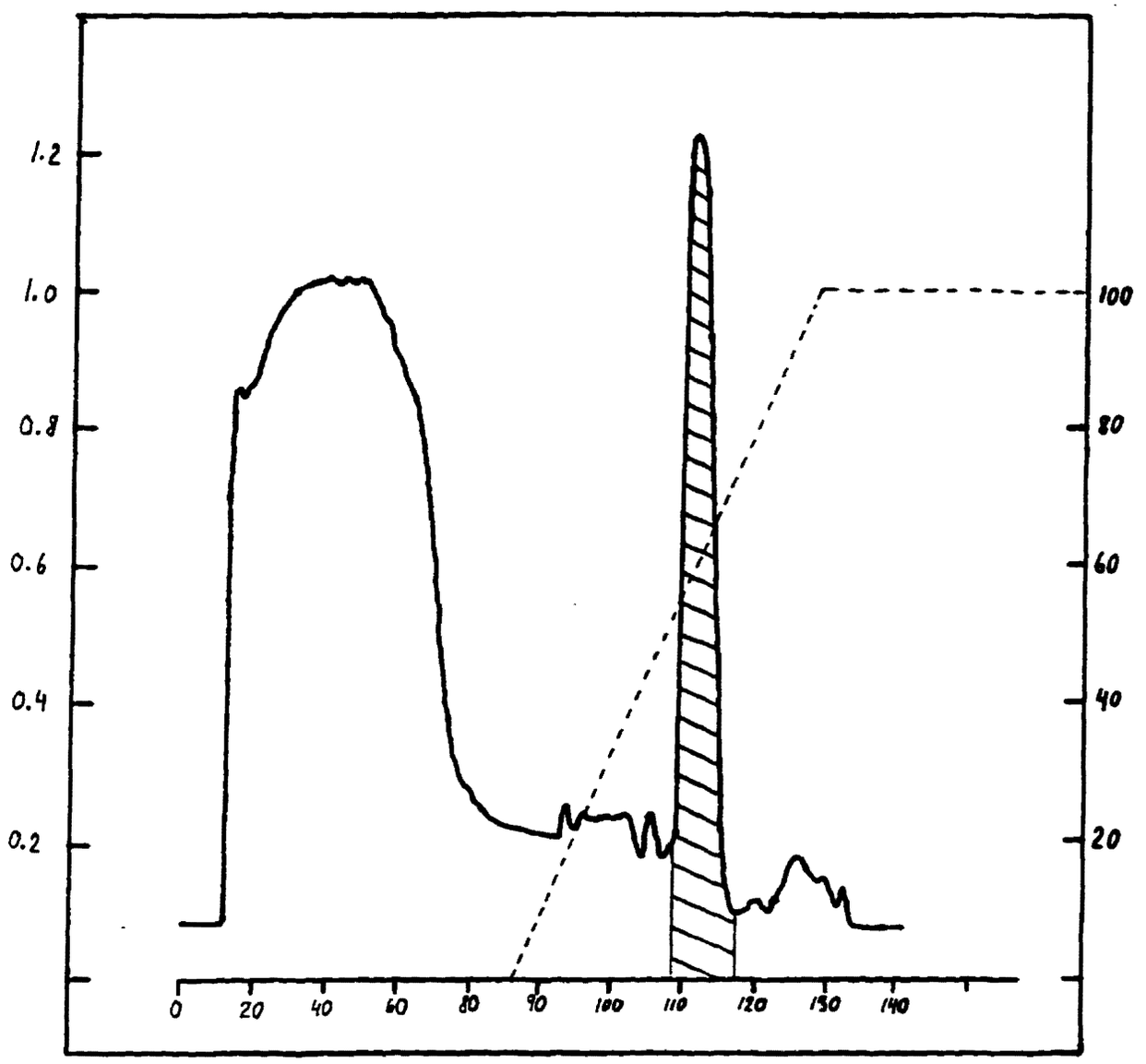


Figure 2.24. S-Sepharose Fast Flow chromatogram from purification of LaL from *E. coli* BL21(λDE3)/pLR102. Shaded area represents the active fraction collected.

Sample:	Cell free extract (from a 6 L culture)
Column:	S-Sepharose Fast Flow packed in an HR 16/50 column
Initial Buffer:	Buffer A
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.0 M KCl (w/v) in Buffer A
Flow Rate:	2.5 mL/min during sample application (before ≈ 86 min)
	5.0 mL/min during gradient application (after ≈ 86 min)
Left Axis:	Absorbance at 280 nm, indicated with the solid line
Right Axis:	Percentage Buffer B, indicated with the dashed line
Bottom Axis:	Time in minutes
Gradient:	0-100 % Buffer B over 40 min

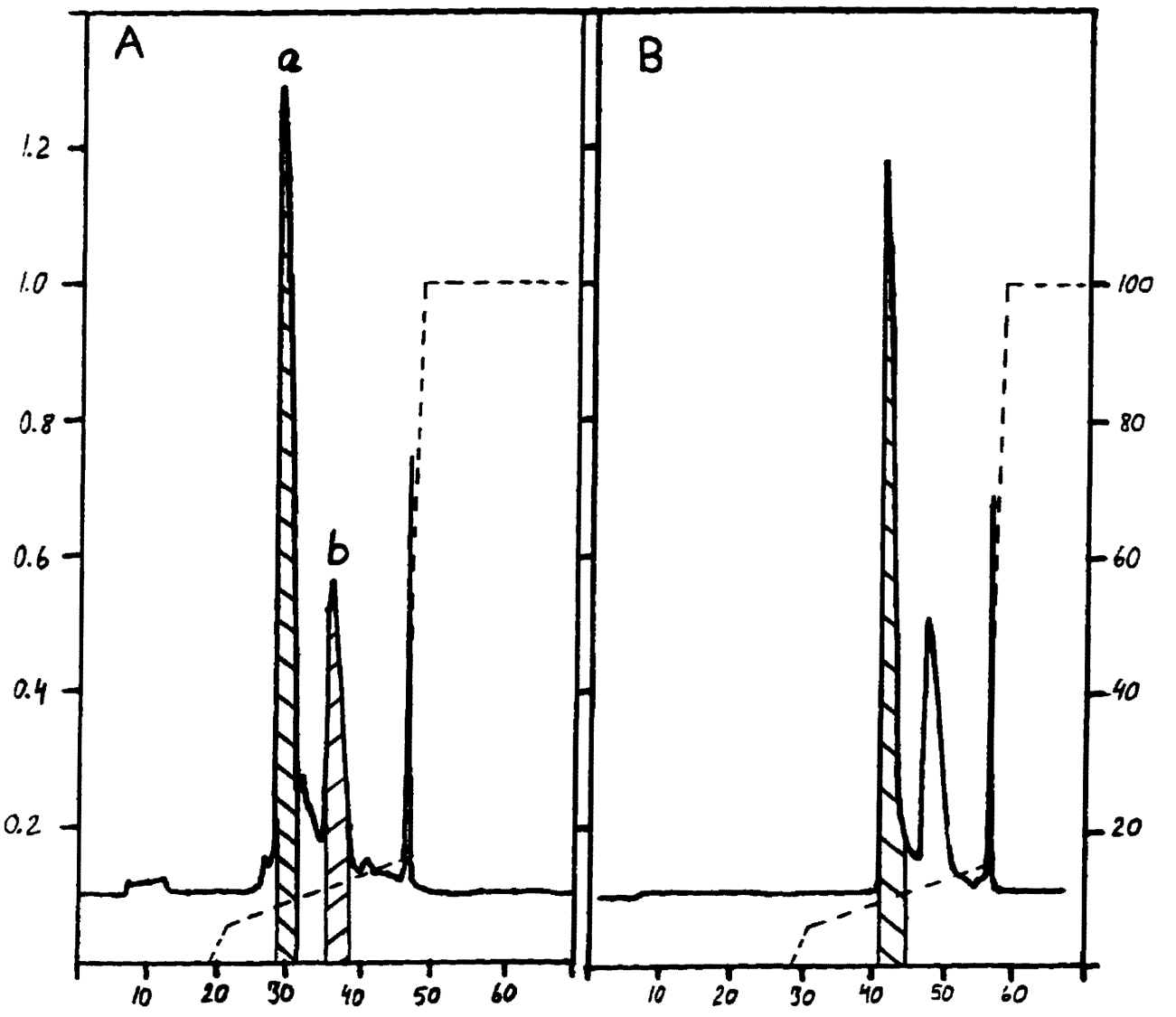


Figure 2.25. Mono-S chromatograms from purification of LaL from *E. coli* BL21(λDE3)/pLR102. Shaded areas represent the active fractions collected. Both peaks a and b were collected in (A) and then the pooled peak b fractions were reapplied collecting the active fraction indicated in (B).

Sample:	(A) One-sixth of the S-Sepharose Fast Flow fraction (B) Re-application of pooled peak b fractions (after dialysis) collected in (A)
Column:	Mono-S HR 10/10
Initial Buffer:	Buffer A
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.0 M KCl (w/v) in Buffer A
Flow Rate:	2.0 mL/min
Left Axes:	Absorbance at 280 nm, indicated with the solid line; same scale for A and B
Right Axes:	Percentage Buffer B, indicated with the dashed line; same scale for A and B
Bottom Axes:	Time in minutes
Gradient:	0-5 % Buffer B over 2 min 5-15% Buffer B over 25 min 15-100% Buffer B over 2 min

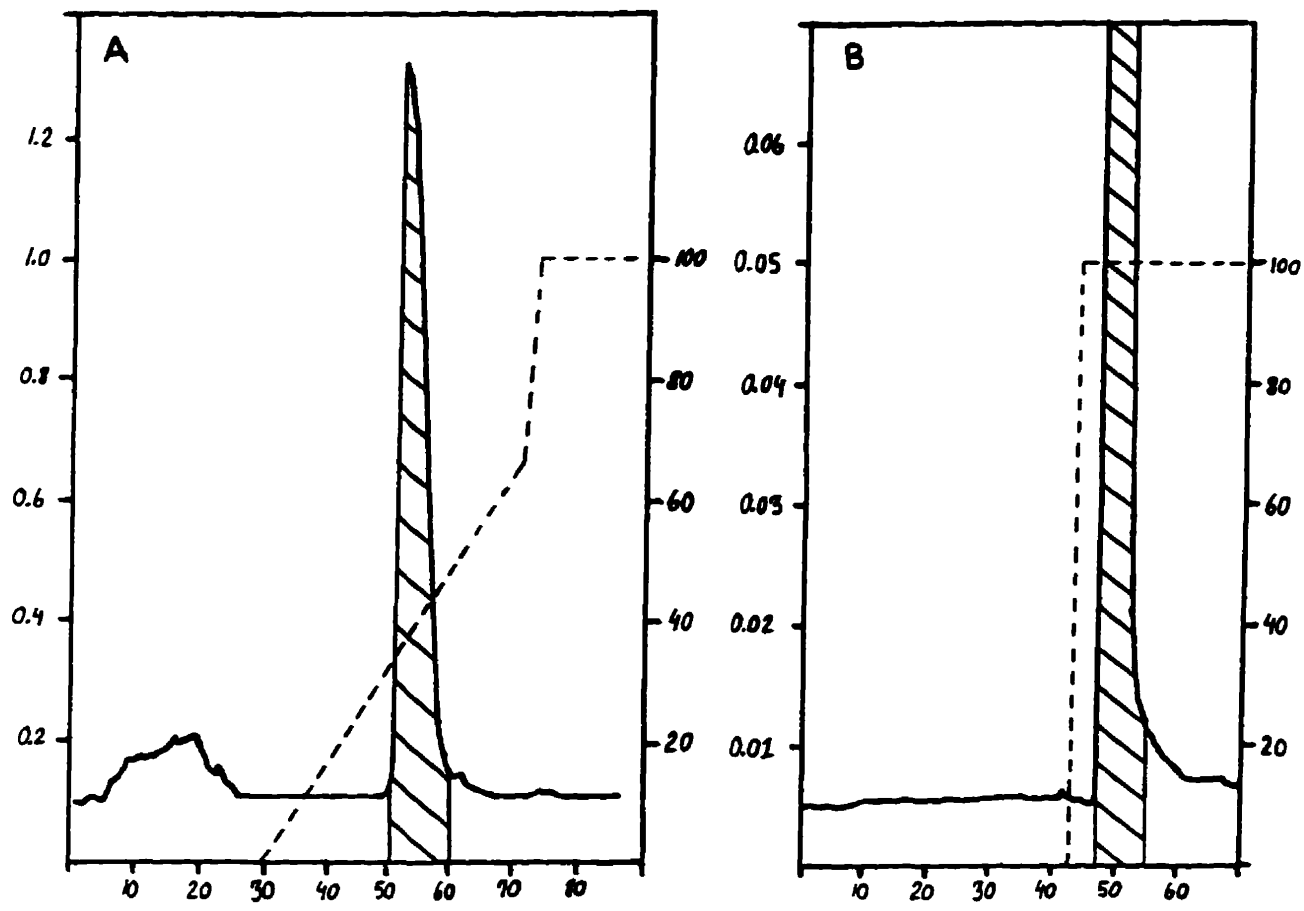


Figure 2.26. Phenyl-Superose chromatograms from purification of LaL from *E. coli* BL21(λDE3)/pLR102. Shaded areas represent the active fraction collected. Chromatogram (B) is from the concentration run.

Sample:	(A) One-fourth of the pooled Mono-S fractions, adjusted to 1.7 M $(NH_4)_2SO_4$ (B) One-half of the pooled Phenyl-Superose fractions collected in (A), adjusted to ≈ 1.7 M $(NH_4)_2SO_4$
Column:	Phenyl-Superose HR 10/10
Initial Buffer:	Buffer B
Buffer A:	50 mM KPB, pH 7.0
Buffer B:	1.7 M ammonium sulfate (w/v) in Buffer A
Flow Rate:	1.0 mL/min during sample application (before ≈ 30 min in (A) and ≈40 in (B)) and then 1.5 mL/min during gradient application
Left Axes:	Absorbance at 280 nm, indicated with the solid line In (A) the scale is 40× the sensitivity shown before ≈ 27 min
Right Axes:	Percentage Buffer A, indicated with the dashed line
Bottom Axes:	Time in minutes

Gradient:	Chromatogram (A)	Chromatogram (B)
	0-66.7% Buffer A over 40 min	0-100% Buffer A over 0.5 min
	66.7-100% Buffer A over 3 min	

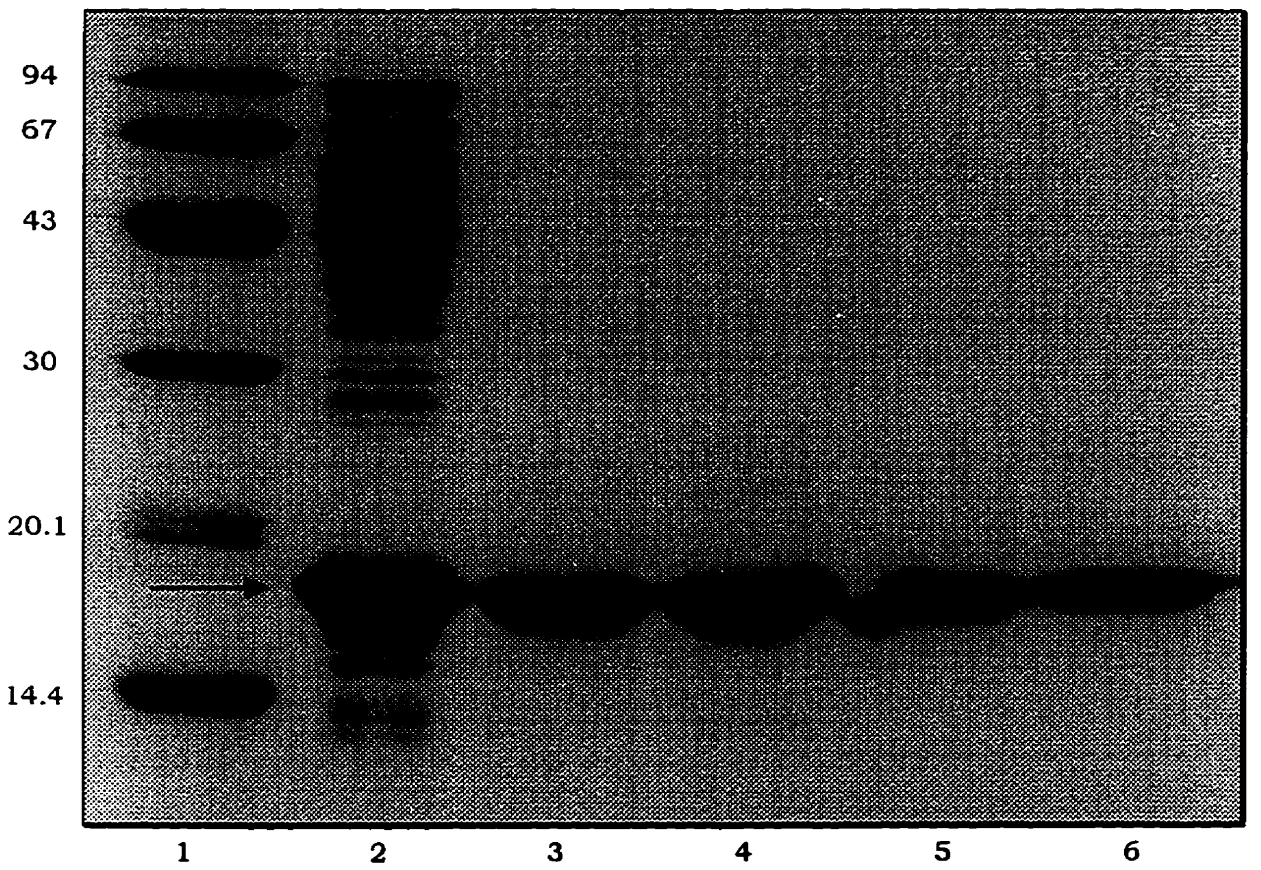


Figure 2.27. SDS-PAGE analysis of the stages of purification of LaL from *E. coli* BL21(λDE3) harboring pLR102. Protein samples were electrophoresed on a 12% T gel followed with silver staining. The arrow indicates the position of LaL. Other very faint bands were visible in the original gel in lanes 3-5 that are not obvious from the figure.

Lane:

1: Protein Standards (in kDa)
2: Cell free extract
3: S-Sepharose Fast Flow fraction
4: S-Sepharose Fast Flow fraction (dialysed)
5: Mono-S fraction
6: Phenyl-Superose fraction

early during this chromatography, thereby still necessitating this step. The strength of expression of LaL from this system is again demonstrated in the lane corresponding to the cell free extract (Lane 2, Fig. 2.27).

A summary of the total protein of each purification intermediate is presented in Table 2.3. It is given as a guide so that others who will conduct the purification, have an estimation as to the amount of protein that is obtained from each step (dependent on the scale of culture growth) and the number of chromatographic steps necessary to process the fractions. The total protein of the fractions are given as determined by Bradford analysis and by the experimentally determined molar extinction coefficient at 280 nm for LaL. Since the majority of protein in the column fractions is LaL (Fig. 2.27), the absorbance at 280 nm of these fractions will primarily arise from and reflect the concentration of LaL. As mentioned previously, Bradford analysis overestimates the protein content which is exemplified on comparison of the total protein determined by the two methods. It is interesting to note that the total protein in the cell free extract obtained from a 6 L culture of *E. coli* BL21/pLR102 is only 1001 mg, as compared to 1606 mg for TG-1/pHDM10 (Table 2.2). Although *E. coli* BL21 and TG-1 are different strains with some expected differences in protein profile and composition, the comparison does indicate how the cellular resources are channeled towards the production of LaL with a concomitant reduction in the total protein synthesized during induction of BL21/pLR102.

For future reference, it is noteworthy to suggest that purification of LaL from BL21/pLR102 could be performed on a very large scale to generate gram quantities of LaL. Approximately 100 mg of protein from 1 L of culture is obtained in the S-Sepharose Fast Flow fraction. Since the loading capacity of S-Sepharose Fast Flow is 50-75 mg/mL resin, the volume of the HR 16/50 column (≈ 100 mL) would allow for accommodation of 5-7.5 g of protein in a single chromatographic run. This would allow the processing of cell extract from 50-75 L of culture. LaL would be recovered in a convenient volume from the column which could be frozen, stored and then further processed by Mono-S and Phenyl-Superose chromatography as required.

To date, purification of wild type LaL from BL21/pLR102 was performed from 1 L and 6 L of culture and yielded 28.0 and 24.4 mg of LaL per L of culture respectively. In addition, [*methyl*-^{13}C]methionine labelled LaL was purified with a yield of 31 mg per L of culture from a methionine auxotroph (*E. coli* B834, a derivative strain of BL21) harboring pLR102 grown in minimal media supplemented with [*methyl*-^{13}C]methionine (see section 3.2.4.3.3). As such, excellent and reproducible yields are obtained upon purification of

Table 2.3. Purification of LaL from a 6 L culture growth scale of *E. coli* BL21(λDE3)/ pLR102.

Fraction	Number of Runs Perfomed[a]	Volume (mL)	Total Protein (mg) by Bradford[b]	Total Protein (mg) by ε_{280}[c]
Cell Free Extract		110	1001	
S-Sepharose Fast Flow	1			
(1) column fraction		48.5	556	
(2) after dialysis		54.0	497	322
Mono-S	6 + 1	49.6	321	188.5
Phenyl-Superose				
(1) Pooled 1st column fraction	4	59.4	285	159.4
(2) after pooling of 2nd column (concentration) and dialysis	2	20.0	236	146.5
(3) after lyophilization		-	145.0	

[a] The number of individual runs performed over each column. For the Mono-S column, 6 runs were performed of the S-Sepharose Fast Flow fraction (collecting the major and minor fraction, Fig. 2.25)) and then 1 run was performed to process the minor fractions collected.
[b] Determined by Bradford analysis except for the Phenyl-Superose fraction (3) determined by mass of the lyophilized powder following correction for buffer salts (2.2.10.2).
[c] Determined using the experimentally determined (2.3.3) molar extinction coefficient at 280 nm (ε_{280}) for LaL. Note that the Mono-S and Phenyl-Superose (1) fractions contain higher concentrations of salt which may affect the 280 nm absorbance of LaL (see 2.3.3).

LaL from the *E. coli* BL21/pLR102 system. Of the 175 mg of wild type LaL purified from this system, 120 mg has been provided to Dr. Albert Berghuis (McMaster University, Hamilton, ON) in a collaborative effort for crystallographic studies.

2.3.2. Characterization of LaL Purified from *E. coli* TG-1/pHDM10 and *E. coli* BL21(λDE3)/pLR102

2.3.2.1. Nucleotide Level

The Vent$_R$ DNA polymerase used for the amplification of the λ R gene by PCR is a high-fidelity polymerase, with an error rate of 57 per 10^6 base insertions (or approximately 1 error will occur in the daughter strand for each 17544 bases of template DNA copied

(New England BioLabs Catalog)). Even though the likelihood of an error arising on copying the 504 bp PCR product (Fig. 2.20) is therefore quite small given this fidelity, the absence of errors was established by nucleotide sequencing. Commercially available primers designed for the pET-22b vector were used to obtain the *R* gene sequence from both directions from pLR102. There was complete agreement of the sequence obtained with that previously determined (Sanger et al., 1982) confirming that no mutations had occurred during the PCR reaction.

Although the *R* gene from pHDM10 was itself not sequenced, pHDM10 was used as the template for the PCR reaction. Since pLR102 possessed the correct sequence for the *R* gene, then in principle, pHDM10 should also (although it is possible that a mutation in pHDM10 was incorrectly copied to pLR102 giving the correct sequence). In addition, construction of pHDM10 involved standard manipulations (i.e. restrictions and ligations) to clone the *R* gene fragment from λ DNA into pTTQ18. Therefore, any possibility of mutation would only come from the natural spontaneous mutation rate for maintaining plasmids in *E. coli* hosts.

2.3.2.2. Protein Level

i) N-Terminal Analysis. A sample of LaL purified from each system was subjected to *N*-terminal analysis. The correct amino acid sequence was obtained without ambiguity for 21 cycles of Edman degradation for LaL purified from BL21/pLR102 and 20 cycles for LaL purified from TG-1/pHDM10.

ii) Electrospray Mass Spectrometry. Accurate mass measurements of LaL preparations were made possible following the very fortunate and timely acquisition of an electrospray mass spectrometer by Dr. G. Lajoie (Dept. of Chemistry, University of Waterloo, ON). Based on the amino acid composition of the protein (Sanger et al., 1982), the calculated molecular weight of LaL with neutral (unionized) functional groups is: (a) 17814.1346 $g \cdot mol^{-1}$, using the exact mass of the most abundant isotope of the elements (the monoisotopic mass) and (b) 17825.2174 $g \cdot mol^{-1}$, using the atomic mass of the elements (the average mass). In the case of large molecules such as proteins, the average mass is determined by mass spectrometry. For example, the 158 amino acids comprising LaL include a total of 788 carbons. Because of this large number, each protein molecule will essentially have the same relative composition of ^{12}C and ^{13}C based on their natural abundances (i.e. statistically, each molecule should have about 779-780 ^{12}C (≈98.9%) and

8-9 ^{13}C (≈1.1%)). Since the mass spectrometer detects individual ions of the protein molecules, the mass of the ion arising from the carbons will represent the relative abundance of each isotope (i.e. the mass average).

Typical mass spectra (obtained using the data acquisition parameters given in section 2.2.11) of LaL purified from the two systems are illustrated in Fig. 2.28. In all cases of wild type LaL examined, the mass measured was in excellent and complete agreement with and never varied more than 2 Da from the calculated mass average value. Therefore, taken together with the correct N-terminal sequence obtained, the ESMS results provide almost unambiguous evidence that LaL purified from TG-1/pHDM10 and BL21/pLR102 consisted of the correct primary sequence.

The spectra shown in Fig. 2.28 are computational reconstructions of the multiply charged raw data performed using the MaxEnt algorithm. This algorithm finds the simplest molecular mass spectrum (spectrum of maximum entropy) that could account for the observed multiply charged raw data and produces true molecular mass spectra from the raw data. The reconstruction has the advantage that all (or most) of the components in a sample containing a mixture of analytes can be more readily identified than from analysis of the multiply charged spectra of the same multi-component sample. In addition, MaxEnt spectra are quantitative and the area under individual peaks can be obtained as a measure of the relative amounts of the species present. The shape and width of the peaks generated partially describes the distribution of molecular isotopes which has a characteristic shape that the conversion can "detect". The associated error reflects errors involved in this computed reconstruction.

The multiply charged raw data spectra for the MaxEnt spectra in Fig. 2.28 are shown in Fig. 2.29. The raw data was always examined to confirm the MaxEnt determined mass. The MaxEnt analysis only takes into account the data acquired over the specified molecular weight range; for LaL, data over a MW range of 600-1200 Da gave the optimal ion distribution based on the acquisition parameters used. However, if this optimal range was not known and did not encompass good quality and sufficient multiply charged peaks, the reconstruction may not produce an accurate description of the species present. Therefore, analysis of the raw data provides a feeling as to the quality of the MaxEnt results. Table 2.4 lists the calculated masses for the individual ions that would be obtained from LaL. Comparison of the values listed in Table 2.4 to those indicated on the spectra in Fig. 2.29 demonstrate that the series distribution could have only arisen from LaL (or a protein with identical mass). The accuracy and precision in the measurement is also demonstrated in the mass and associated error for LaL calculated solely from the ion series (see value at top right corner of the spectra in Fig. 2.29).

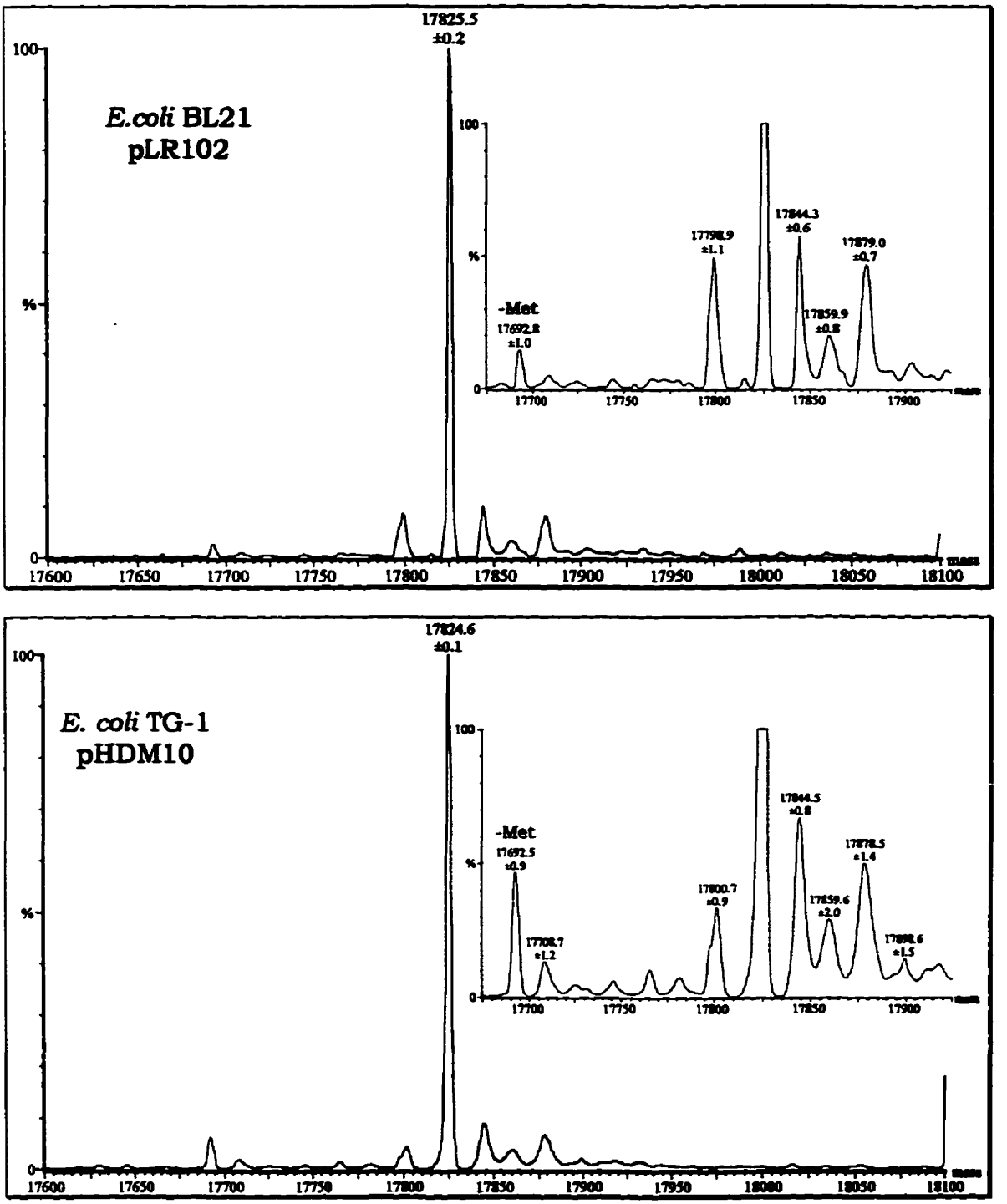


Figure 2.28. MaxEnt reconstructed ESMS spectra of LaL purified from *E. coli* BL21/pLR102 and TG-1/pHDM10. The samples were desalted by reverse phase chromatography. The spectra were resolved to a resolution of 0.25 Da. The insets show the same expanded region from each spectra. The -Met denotes a species corresponding to the mass of LaL minus the mass of a methionyl residue.

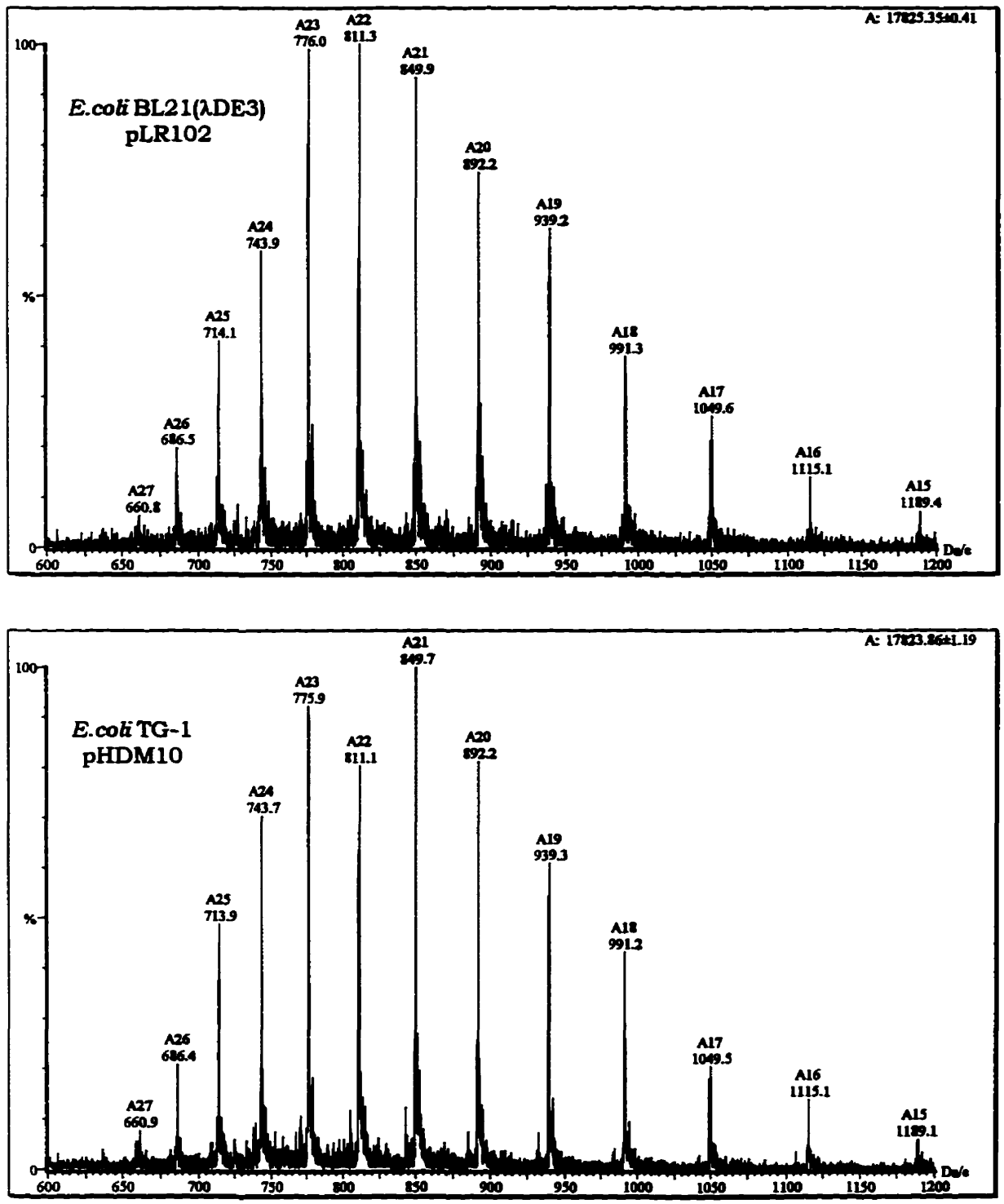


Figure 2.29. Multiply charged raw data ESMS spectra of LaL purified from *E. coli* BL21(λDE3)/pLR102 and TG-1/pHDM10. The observed masses for the molecular ion series can be compared to those calculated and listed in Table 2.4.

Table 2.4. Masses of the Molecular Ion Series for LaL.

Molecular Ion (A + x)		Mass	Molecular Ion (A + x)		Mass
A	A	17825.217	A + 19	A19	939.177
A + 1	A1	17826.225	A + 20	A20	892.269
	⋮	⋮	A + 21	A21	849.828
			A + 22	A22	811.245
A + 14	A14	1274.238	A + 23	A23	776.017
A + 15	A15	1189.356	A + 24	A24	743.725
A + 16	A16	1115.084	A + 25	A25	714.017
A + 17	A17	1049.550	A + 26	A26	686.593
A + 18	A18	991.298	A + 27	A27	661.201

Calculated using mass = [(A) + (x)(1.0079)]/x, where A is the calculated mass of unionized LaL and x is the number of H^+ (proton ions) associated with the xth charged molecular ion.

There are additional peaks seen in Fig. 2.28 wherein most can be explained. The species indicated in the insets of Fig. 2.28 as "-Met" are believed to arise from LaL in which the initiator methionyl residue has been removed by methionine aminopeptidase. This peak was always observed and was typically present at amounts of 3-5% to that of the major species (i.e. LaL retaining the N-terminal Met). Loss of the N-terminal methionine would reduce the mass of LaL by 131.1986 mass units (i.e. 17825.2174 - 131.1986 = 17694.0188), the approximate mass observed for this species in the spectra.

A series of species having molecular weights higher than that of LaL are present in the spectra (see insets, Fig. 2.28). These species have the correct mass that would agree with the formation of multiple water adducts with LaL; the peaks have an approximate mass increment which corresponds to the mass of water (18.015 Da). Even the "-Met" species appears to have associated water adducts. The mass spectra in Fig. 2.28 were obtained from samples of LaL that had been desalted by reverse phase chromatography. Recall that when LaL was purified from TG-1/pHDM10 or BL21/pLR102, the final lyophilized enzyme preparation will contain potassium phosphate salts. When the enzyme was subjected to ESMS without prior desalting, the presence of both potassium and water adducts becomes more evident (Fig. 2.30). The overall appearance of the spectra for the non-desalted sample (Fig. 2.30) is very similar to the desalted ones (Fig. 2.28) except for the peaks appearing at molecular weights higher than that of LaL. For the non-desalted sample, the higher molecular weight species can be rationalized as a combination of multiple potassium only or potassium/water adducts of LaL as depicted in Fig. 2.30. Therefore, this analysis suggests that these species do not result from contaminating protein but from the adducts with LaL described.

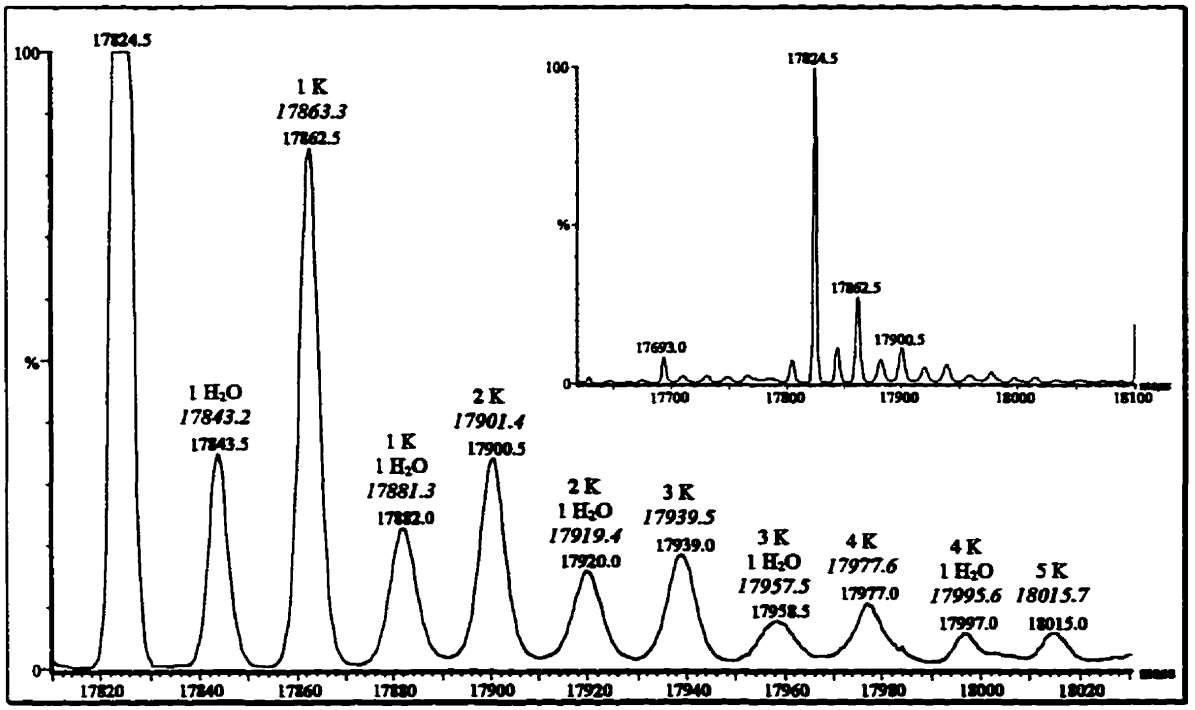


Figure 2.30. ESMS spectrum of non-desalted LaL illustrating the potassium (K) and water adducts. The sample was purified from *E. coli* TG-1/pHDM10. Shown is the observed mass (directly above each peak) and the calculated mass (in italics) expected for the different adducts listed. Molecular weights used for the calculations are: LaL, 17825.2174; H_2O, 18.0152; K, 38.0904 (since K^+ will replace H^+ in the molecular ion, the increase in mass due to K is 39.0983 - 1.0079 = 38.0904). The inset is the unexpanded spectrum and the intensity of the adducts can be compared to those in Fig. 2.28.

There is another contaminating peak present in the spectra presented in Fig. 2.28 with an associated mass that was found to vary in different spectra from 17797-17806 Da. No explanation can be offered describing a modification in LaL that could account for a decrease in mass of 18-27 Da. Unfortunately, it appears that this species might originate from a protein co-purified with LaL, which, because of its similar mass, would not be detected by SDS-PAGE analysis. One cannot discredit the possibility that the peak may be an artifact from the MaxEnt reconstruction. No other species in the many repeated spectra obtained of LaL demonstrated a variability in mass to the degree that was observed for that particular peak. Indeed, altering the MaxEnt parameters involved in processing the raw data had the greatest affect on this peak, not only in its observed mass but also in its intensity which varied from approximately 3-8% of that for the peak corresponding to LaL.

iii) Estimation of the Isoelectric Point (pI)

A cell free extract obtained from *E. coli* TG-1/pHDM10 was subjected to preparative isoelectric focusing. Using the ampholyte BioLyte 8/10 and SDS-PAGE analysis of the fractions, LaL could be detected in fractions 10-18 which encompassed a range of pH 9.25 to 10.50. The majority of LaL was focused into fractions 12-14 in which the pH range was 9.51 to 9.80. This latter range agrees well with the calculated pI of 9.7 (PC/Gene software, IntelliGenetics). Fractions in which LaL was focused also contained other proteins, with LaL representing roughly 50% of the total protein in fractions 12-14. As such, preparative isoelectric focusing would not permit the purification of LaL. Furthermore, addition of the ampholytes resulted in loss of enzyme activity which could not be restored even after dialysis of the fractions.

2.3.3. Guanidine Induced Unfolding and Molar Extinction Coefficient of LaL

The guanidine unfolding curve for LaL is shown in Fig. 2.31. No difference was found between the curves obtained from samples of LaL incubated with the denaturant for 1 hr or for 3.5 hr. Analysis of the data suggests that concentrations of guanidine in excess of 2 M result in essentially complete denaturation of LaL. The denatured state of LaL at these concentrations of guanidine is substantiated by the following observations. The fluorescence of a protein is usually dominated by the Trp residues. As a protein unfolds and the Trp residues are exposed from the interior of the protein to the more polar solvent, there is an accompanying red shift (to higher wavelength) and decrease in intensity of the emission spectrum (Freifelder, 1982). As well, both increases and decreases in fluorescence intensity can occur upon protein unfolding (Schmid, 1990). This is what appears to occur. At denaturant concentrations of 0-1.0 M, the emission maximum remains at 344 nm; however, the increase in fluorescence intensity suggests the occurrence of some conformational change (without complete unfolding) at these low guanidine concentrations which affect the exposure of some of the Trp residue(s). At 1.5 M guanidine, a sharp decrease in intensity takes place (the ensuing gradual increase in the intensity of samples can be attributed to the increase in concentration of guanidine, which exhibits a fluorescence of approximately 0.04 units at 7.5 M) and the emission maximum is 352 nm. This maximum increases to 354 nm for concentrations of denaturant of 6.0 M and higher. This agrees with the expected redshift of emission maximum from lower wavelength to around 350 nm for an unfolded protein which corresponds to the fluorescence maximum of Trp in aqueous solution, although the exact

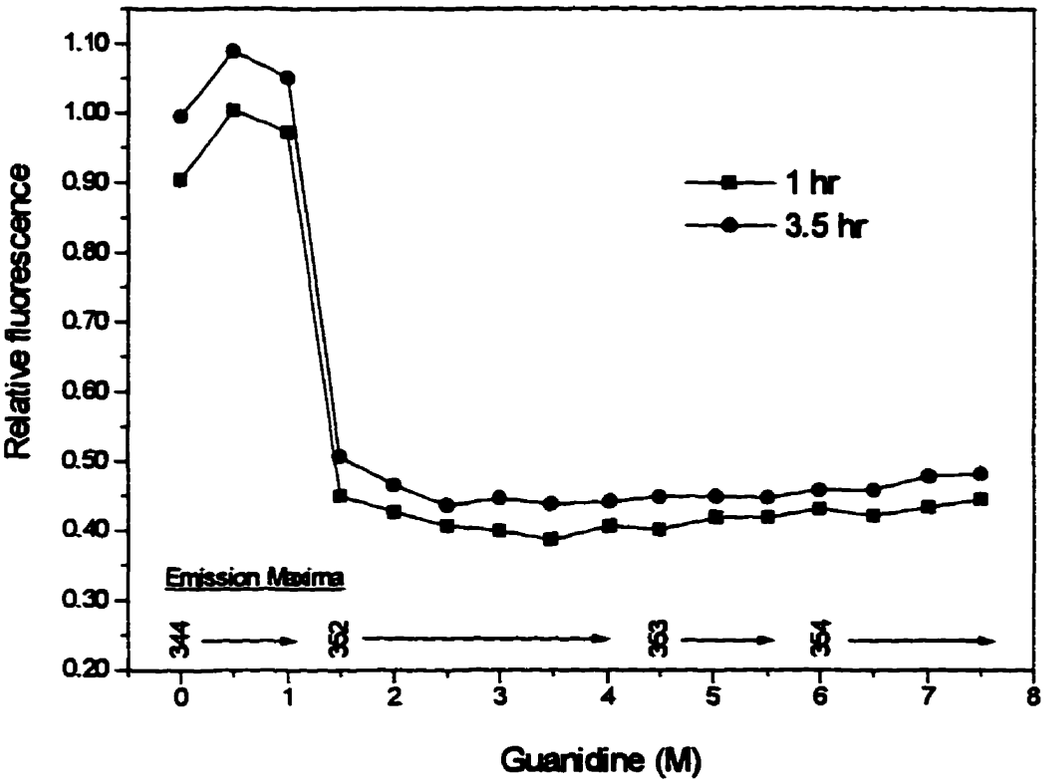


Figure 2.31. Guanidine induced unfolding curve of LaL. Samples contained 2.2 μM LaL in 10 mM DTT, 0.02 M KPB, pH 6.5 and guanidine hydrochloride as indicated. Fluorescence was measured at excitation and emission wavelengths of 280 and 344 nm respectively. The emission maxima of samples (280 nm excitation) are also given.

location of this maximum will depend on the nature and concentration of the buffer (Schmid, 1990).

Therefore, the properties of the unfolding curve and fluorescence spectra are consistent with the denaturation of LaL. As well, samples incubated in 1.0 M or less denaturant retained activity while those at the higher concentrations did not. In all essence, it can be presumed that LaL is completely denatured at 6.0 M guanidine and therefore, lending credence to the calculated value of ε_{den} = 29280 M^{-1} cm^{-1} (see 2.2.10.3).

The results concerned in the determination of the molar extinction coefficient at 280 nm for LaL (ε_{280}) are given in Table 2.5. The absorbance of solutions containing identical concentrations of LaL were obtained under both denaturing and native conditions. As was described previously (see 2.2.10.3, Eq. 4), the ε_{280} of native LaL can be determined using ε_{den} and the absorbance values of the denatured and native samples (Gill & von Hippel, 1989). Under the experimental conditions performed, the ε_{280} for LaL was

found to be 31712 ± 210 (± 0.7%) M^{-1}-cm^{-1}. The confidence in the precision of this value is demonstrated in the low resultant standard deviation (although the error does not take into account the small observed variability (approximately 0.3-3%) in A_{280} values). For example, an A_{280} measurement of 0.400 of a LaL solution would equate to an upper and lower (based on the obtained error) protein content of 0.226 and 0.223 mg·mL^{-1}. This level of precision is sufficient for essentially any application.

Table 2.5. Calculation of the molar extinction coefficient for LaL

Stock Protein Solution Prepared From	Volume of Stock [a] (μL)	Condition [b]	A_{280} [c]	$\varepsilon_{280,\ native}$ [d] (Molar Extinction Coefficient, M^{-1}-cm^{-1})
1. *E. coli* TG-1 pHDM10	50	denatured	0.302 ± 0.004	
		native (MQW)	0.327 ± 0.010	31704
		native (KPB)	0.325 ± 0.007	31510
	100	denatured	0.610 ± 0.003	
		native (MQW)	0.665 ± 0.009	31920
		native (KPB)	0.663 ± 0.005	31824
				avg. 31740 ± 177
2. *E. coli* BL21 pLR102	50	denatured	0.351 ± 0.004	
		native (MQW)	0.379 ± 0.004	31616
		native (KPB)	0.384 ± 0.012	32033
	100	denatured	0.692 ± 0.013	
		native (MQW)	0.749 ± 0.002	31692
		native (KPB)	0.742 ± 0.011	31396
				avg. 31684 ± 264
				overall average 31712 ± 210 (0.7%)

[a] The protein aliquot added to solutions (final vol. of 1.00 mL) under the desired conditions. [b] Conditions are: denatured, stock in 6.014 M Gdn·HCl, 20 mM KPB, pH 6.5 for 1 hr at 25 °C; native (MQW), stock in MQW; native (KPB), stock in 50 mM KPB, pH 7.0. [c] Absorbance at 280 nm. Values represent the average ± SD for three independently prepared samples. [d] Calculated from $\varepsilon_{native} = (\varepsilon_{denatured})(A_{native})/(A_{denatured})$ where $\varepsilon_{denatured}$ for LaL was calculated to be 29280 M^{-1} cm^{-1}.

Recently, a ε_{280} value of 31000 M^{-1}-cm^{-1} was reported based on titration of the unique cysteine of LaL with excess 4,4'-dithiodipyridine (Soumillion et al., 1995). This method requires complete accessibility to all the Cys residues (Grasseti & Murray, 1967).

The value reported was obtained under conditions of 1 M guanidine. As discussed above, under our experimental conditions, this concentration of denaturant does not appear to result in denaturation of LaL. Therefore, the lower value of 31000 $M^{-1}cm^{-1}$ reported may be an underestimate if complete denaturation was falsely assumed and the Cys are not entirely accessible.

The ε_{280} values determined for LaL purified from either TG-1/pHDM10 or BL21/pLR102 are in excellent agreement with each other (0.17% difference in their average values) affirming their analogous composition. No trend in the ε_{280} can be noted for native samples determined in water or 50 mM phosphate buffer. This is a favourable observation in that the 280 nm absorbance of LaL does not appear to be affected by phosphate buffer (at least to a concentration of 50 mM). This is noteworthy since all of the protein determinations by A_{280} measurements towards this thesis were performed on LaL samples in either water or phosphate buffer (2.5-50 mM) and the values will therefore not have been influenced by buffer concentration. It is also useful because accurate stock LaL solutions can be prepared free of salt in water (e.g. following exhaustive dialysis) for subsequent dilution into any desired buffer. The absorbance spectrum of a protein can be altered by a buffer since the chemical and physical properties of a buffer can affect the tertiary structure of a protein or the buffer can interact directly with amino acids affecting their absorption properties (Stoscheck, 1990). It is therefore suggested that the experimentally determined ε_{280} for LaL be used only for protein determination of samples in phosphate buffer of 50 mM or less. The affect of other buffers on the 280 nm absorbance of LaL has not been determined and protein determinations involving other buffer systems may be erroneous and should be avoided.

2.3.4. Comparison of Protein Quantitation Methods

Any study involving a protein or enzyme requires a reliable, reproducible and accurate means of determining the concentration of the protein in solution. Primarily because of its convenience, total protein was estimated in the early stages of this work by the dye binding (Coomassie Brilliant Blue G-250) method of Bradford (1976). The absorbance maxima of the dye in an acidic solution shifts from 465 to 595 nm after addition of a protein due to stabilization of the anionic form of the dye by its interaction with specific residues with the protein (Compton & Jones, 1985). The method suffers in that variations in color yields are observed from protein-to-protein and from interfering factors such as buffers, detergents and pH (Stoscheck, 1990).

Bradford analysis tended to overestimate quantities of purified LaL when bovine serum albumin was employed as the protein standard. The linear response of colour development for LaL was approximately 1.5 times greater than that of BSA (for example, 3 ug of LaL would give the same colour response that would result from 4-5 μg of BSA). In addition, the presence of ammonium sulfate in the final samples of LaL obtained from the Phenyl-Superose chromatography appeared to increase the colour response.

It was very important to establish a method(s) that would allow accurate determinations of LaL concentration, especially important for the preparation of LaL samples intended for the growth of crystals. The quantity of protein obtained from purification of LaL from *E. coli* TG-1/pHDM10 permitted this determination by measuring the mass of the protein powder obtained from lyophilization of a sample. An obvious tactic would be to exhaustively dialyse the sample against water to eliminate salts from the sample prior to lyophilization and mass measurement. However, when dialysed against water, the lyophilized protein powder could not be re-solubilized. It was determined that lyophilization of LaL previously dialysed extensively against a dilute phosphate buffer (2.5, 5 and 10 mM were explored) produced a powder that was readily solubilized. In one case, dialysis was performed in 10 mM Tris, pH 7.0 which also gave a lyophilized powder that could be re-solubilized.

As such, the total protein obtained from purifications of LaL from TG-1/pHDM10 and BL21/pLR102 was quantitated by measuring the mass of the protein following its extensive dialysis against dilute buffer, lyophilization and correction for buffer salts as described in section 2.2.10.2. A sample calculation that would allow the calculation of the resulting protein and buffer concentrations resulting from dissolving a portion of the lyophilzed powder utilizing this methodology is also given in section 2.2.10.2.

If stored for prolonged periods, it is expected that moisture could add to the mass of the lyophilized powder. For this reason and also to verify the accuracy of the previous method, the molar extinction coefficient for LaL was determined (see sections 2.2.10.3 and 2.3.3). Following the determination of ε_{280}, four purifications of LaL were performed allowing a comparison between the two protein quantitation methods. The results are given in Table 2.6. As can be seen, it is encouraging that the total protein quantitated for these purifications by the two methods are in excellent agreement.

Table 2.6. Comparison of methods to quantitate LaL

Purification[†]	Total Protein (in mg) Quantitated by: (1) mass following dialysis, lyophilization and buffer salt correction	(2) A_{280}	% Difference in (1) and (2)
Pur#20	145.04	146.46	1.0
Pur#21	42.01	42.52	1.2
^{13}CMetPur#4	46.41	45.77	1.4
CF_3Pur#6	11.40	11.31	0.8

[†] Pur#20 and #21 was LaL purified from *E. coli* BL21(λDE3)/pLR102 while CF3Pur#6 was low level trifluoromethionine labelled LaL (see Chapter 4) and ^{13}CMetPur#4 was [*methyl*-^{13}C]methionine labelled LaL purified from *E. coli* B834(λDE3)/pLR102 (see Chapter 3).

2.3.5. Distinctive Properties of LaL Relevant to its Manipulation

2.3.5.1. Maintaining Lysozyme Preparations

Lambda lysozyme exhibited markedly strong stability in solution when prepared at relatively high concentrations (1-5 mg/mL). In one instance, a solution of LaL (≈ 15 mg/mL in 5 mM KPB, pH 7.0, 10% D_2O) was prepared for ^{1}H NMR studies. A small aliquot of this sample was filter sterilized and set aside at ambient temperature. The sample was periodically assayed over the course of 104 days which indicated that essentially no loss in activity occurred. The period monitored was ended when the sample became contaminated with microbial growth.

For kinetic experiments, it was necessary to have access to solutions of LaL at concentrations of 2-10 μg/mL. When solutions were prepared simply in 50 mM KPB, pH 7.0 and maintained on ice, enzyme activity was progressively lost over the course of 45 min, with nearly complete loss after 90 min. Inclusion of 20% glycerol retarded but did not prevent loss of bacteriolytic activity. It was found that if solutions of LaL were prepared in 0.5 M KCl, 100 mM KPB (pH 7.0), 20% glycerol and kept on ice, complete enzyme activity of these dilute solutions was retained over a 3 hr period. This buffer system was not required for solutions of LaL prepared at concentrations of greater than 1 mg/mL.

As was mentioned earlier (2.3.1.1), dilute solutions of LaL (0.05-0.1 mg/mL) obtained from Phenyl-Superose chromatography (which also contained approximately 1 M ammonium sulfate) were found to lose activity when stored at 4 °C or if frozen. These fractions could be maintained indefinitely when glycerol was added to 20%. One such Phenyl-Superose fraction made to 20% glycerol obtained from *E. coli* TG-1/pHDM10 (at a

concentration of 2.0 mg/mL) has been stored at -20 °C for 4 years and has maintained activity. The effect of maintaining more concentrated solutions (without the addition of stabilizing additives) of LaL for prolonged periods at 4 °C was not examined.

Storage of LaL (at -20 °C) in the form of the protein powder obtained following dialysis and lyophilization was also successful in preserving activity. One such preparation has been successfully stored in this manner for over 2 years. It was also possible to prepare concentrated stock solutions of the LaL powder (2-10 mg/mL were routinely prepared) in either KPB or simply MQW which could be repeatedly frozen (in liquid N_2, then kept at -20 °C) and thawed for use without problem.

2.3.5.2. Difficulties Encountered

Over the course of this study, it was found that particular standard protein manipulation techniques posed some obstacles when attempted with LaL. It appears that the source of these difficulties are in part, due to certain physical properties possessed by LaL.

Numerous products are commercially available for the intended use of concentrating protein samples. Our laboratory was equipped with the following that were used at various times in attempts to concentrate LaL samples; Amicon Diaflo® ultrafiltration membranes PM 10 and YM 10 (10000 MWCO), Amicon Centricon®-3 concentrator (3000 MWCO) and Millipore UltraFree Microconcentrator (10000 MWCO, low binding cellulose). The membranes in these products which selectively permit the passage of certain solutes depending on the molecular weight cut off of the membrane, are usually ascribed with a feature of low protein binding. However, when used with LaL solutions, there was always an observed loss of protein. This loss did not result because of LaL passing through the membranes. Although very minute activity could be detected in the "flow through" solutions with the various products, the loss in protein could not be totally accounted for by this reason. The retention of LaL on Phenyl-Superose suggests that LaL is a hydrophobic protein and therefore, could possibly adsorb strongly to the membranes. This is illustrated in the following example. It was desired at one point to prepare a solution of LaL completely free of salt for crystallization purposes. The LaL solution (with a total of 8.91 mg LaL, determined by A_{280}) was exchanged with MQW using a Centricon-3 concentrator. Following the exchange, 6.82 mg of protein was recovered in the retentate solution while the flow through contained only 10 μg of protein. As a result, 2 mg of protein could not be accounted for, presumably as a result of adsorption to the membrane.

LaL was also observed to exhibit abnormal retention on molecular-sieve chromatography resins. This was observed for both Sephadex® G (Pharmacia) and Bio-Gel P (BioRad) type resins. This type of abnormal elution from Sephadex and Bio-Gel resins has also been described previously for HEWL (Whitaker, 1963; Bonilla, 1970). The T4 lysozyme was also found to migrate at an apparent MW of 14 kDa over Sephadex G-75 which is slower than expected for its MW of 18.6 kDa (Jensen and Kleppe, 1972). Sephadex is a gel prepared from the crosslinking of dextran. It is plausible that LaL demonstrated an association with the carbohydrate matrix, thereby increasing its retention to that expected for its migration based solely on the principle of gel filtration. In addition, Sephadex resins contain a small degree of carboxyl groups which could also interact electrostatically with LaL (*pI* ≈ 9.51-9.80, see 2.3.2.2 *iii*)). This may offer an explanation as to the inability to resolve LaL (17.8 kDa) from a 12 kDa contaminating protein using Sephadex G-50 (1500-30000 fractionation range) during the establishment of a purification protocol from TG-1/pLR7. Sephadex G-25 (1000-5000 fractionation range) however, could be used as an effective desalting column when the mobile phase was 50 mM KPB, pH 7.0.

Chromatography of LaL over Superdex 75 HR 10/10 (Pharmacia, 3000-70000 fractionation range), a high performance pre-packed gel filtration column, was also unusual. The matrix of this column is also produced from dextran. When eluted at 0.5 mL/min, the expected retention time for LaL on this column is 24 min when compared to the elution of protein standards using 50 mM KPB, pH 7.0 as the eluent. However, in this buffer, LaL eluted much later than expected (> 60 min). When the elution buffer was 150 mM KPB, pH 7.0, the retention time for LaL was 34 min suggesting that the increase in ionic strength of the buffer reduced interactions of LaL with the matrix. As such, neither Sephadex nor Superdex 75 would be suitable for molecular weight measurements for LaL.

The Bio-Gel matrices are based on crosslinked polyacrylamide and would acquire some hydrophobic nature. When Bio-Gel P-2, P-4 and P-10 (fractionation ranges of 100-1800, 800-4000 and 1500-20000 respectively) were used as potential desalting columns using MQW as the mobile phase, LaL essentially migrated with the salt and not in the void volume as would be expected. The hydrophobic nature of LaL once again appeared to retard its migration over these resins. Increasing the ionic strength of the eluent would be expected to only augment the hydrophobic association and would also be self-defeating if desalting was the end result. Protein adsorption to acrylamide based gel filtration resins and the enhancement of adsorption in the presence of salts has been described previously (Kang et al., 1992).

Methods to remove all salts from LaL were explored primarily to generate concentrated samples of LaL for crystallization. Unfortunately, even dialysis presented some dilemmas. It was determined through NMR experimentation that the Spectra/Por® 3 grade of dialysis tubing (MWCO 3500) routinely used to dialyse LaL samples, was contributing contaminating material(s) to the dialysis retentate solutions. This is illustrated in Fig. 2.32. In an experiment, the simple dialysis of 5 mM KPB, pH 7.0 was performed and samples were obtained from both the retentate (the buffer sample contained in the dialysis tubing) and the dialysate. Approximately equal volumes of each sample (≈ 20 mL) were lyophilized and ^{1}H NMR spectra of each were obtained. The two spectra were virtually identical (Fig. 2.32 A and B) suggesting that the material coming from the tubing (which added a mass of 0.9 mg to the lyophilized retentate sample salts) equilibrated between the retentate and the dialysate solutions and has therefore, a molecular weight of near 3500 or less. The contaminating material was not seen in the spectrum obtained from a sample of 100 mM KPB, pH 7.0 (10 mL) that was simply lyophilized without any contact with the tubing (Fig. 2.32 C). Pre-treatment of the tubing with 1% $NaHCO_3$, 0.05% EDTA, 50% ethanol reduced but did not eliminate these materials. Consultation with a representative from the manufacturer (Spectrum Medical Ind.) did not provide any indication as to the possible origin of these materials, apart from a list of potential inorganic impurities.

It was also found that the ^{1}H NMR spectra of these contaminants were slightly different when the retentate solution also contained protein. This was realized during the NMR studies of [*methyl*-^{13}C]methionine labelled LaL. The ^{1}H NMR spectra of a ≈ 0.1 mM solution of this LaL sample, 1 mM wild type LaL and of a 2 mM solution of HEWL (which were each obtained following dialysis from 5 mM KPB, pH 7.0 and lyophilization) are shown in Fig. 2.33. All peaks visible (except for the water and reference signals at 4.6 and 0.0 ppm) in Fig. 2.33 A are believed to arise from the tubing contaminants (the resonances from LaL could be seen on expansion of the spectra). In the more concentrated enzyme sample of LaL, these contaminants are still readily visible (Fig. 2.33 B). One of the contaminant resonances (at 3.7 ppm) is also clearly visible in the HEWL spectrum (Fig. 2.33 C). The chemical shifts of the most intense signals arising from the contaminants and observed from the [^{1}H-^{13}C] HMQC spectra of the LaL sample (of which the ^{1}H spectrum is shown in Fig. 2.33 A) are; i) the singlet at 3.72 ppm (^{13}C, 32.35 ppm) and ii) the doublet at 1.19 and 1.17 ppm (^{13}C for both, 18.26 ppm).

When the above protein solutions were chromatographed over Mono-S, the unabsorbed material was collected and was found to give the same spectrum as in Fig. 2.33 A. Therefore, these materials are assumed to have overall negative charge character

Figure 2.32. NMR spectra demonstrating contamination resulting from dialysis tubing.

Equal volumes (≈ 20 mL) of retentate and dialysate solutions were obtained from the dialysis of 5 mM KPB, pH 7.0 against itself (overnight at 4 °C). The solutions were lyophilized to dryness and 250 MHz ^{1}H NMR spectra were acquired from the resulting material. Each spectra was taken in D_2O with reference to HOD (4.6 ppm).

(A) Dialysate solution (i.e. the solution obtained from the outside of the dialysis tubing)

(B) Retentate solution (i.e. the solution obtained from the inside of the dialysis tubing)

(C) Spectra of the resulting material from the lyophilization of 100 mM KPB, pH 7.0 (10 mL)

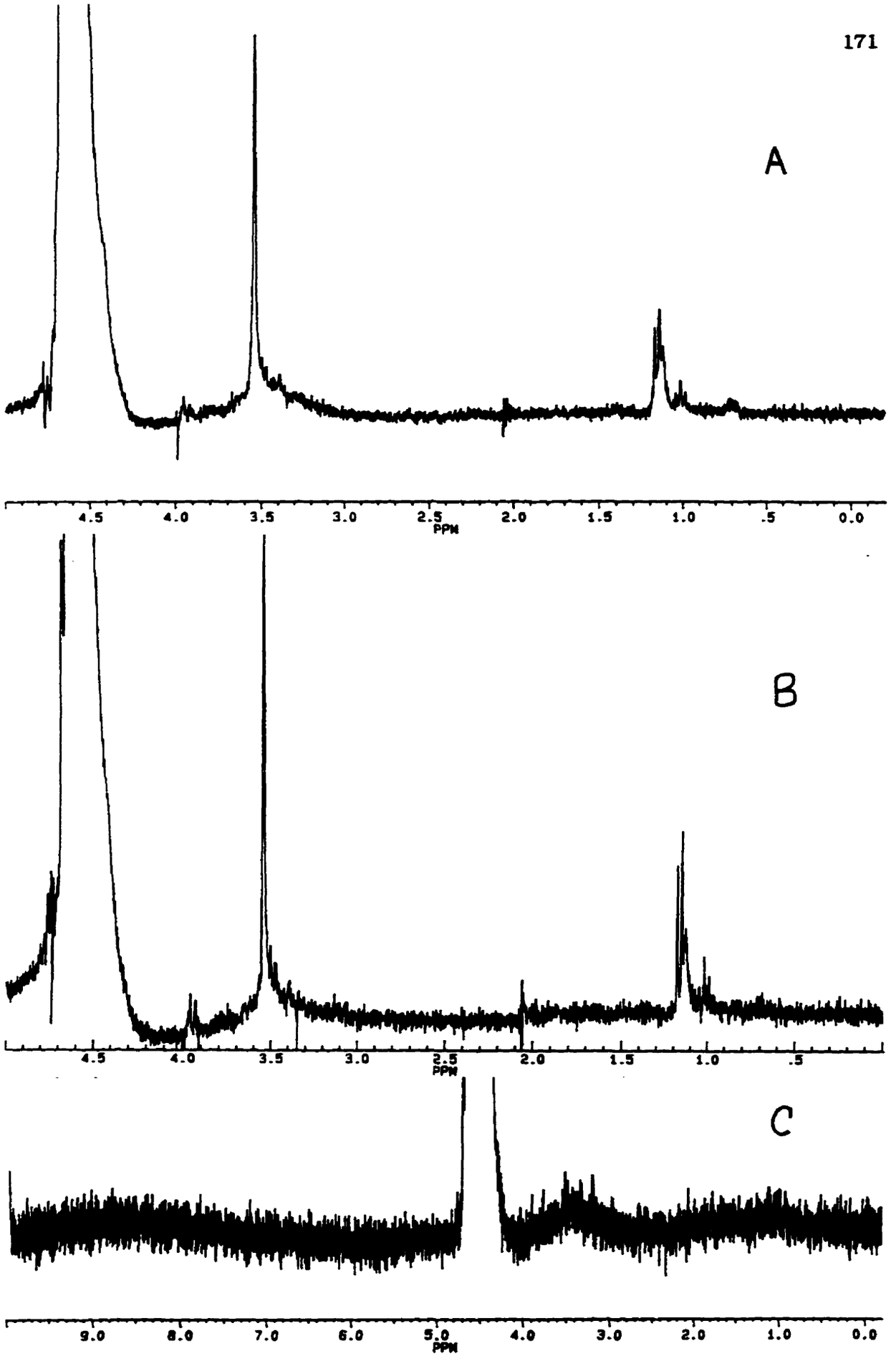
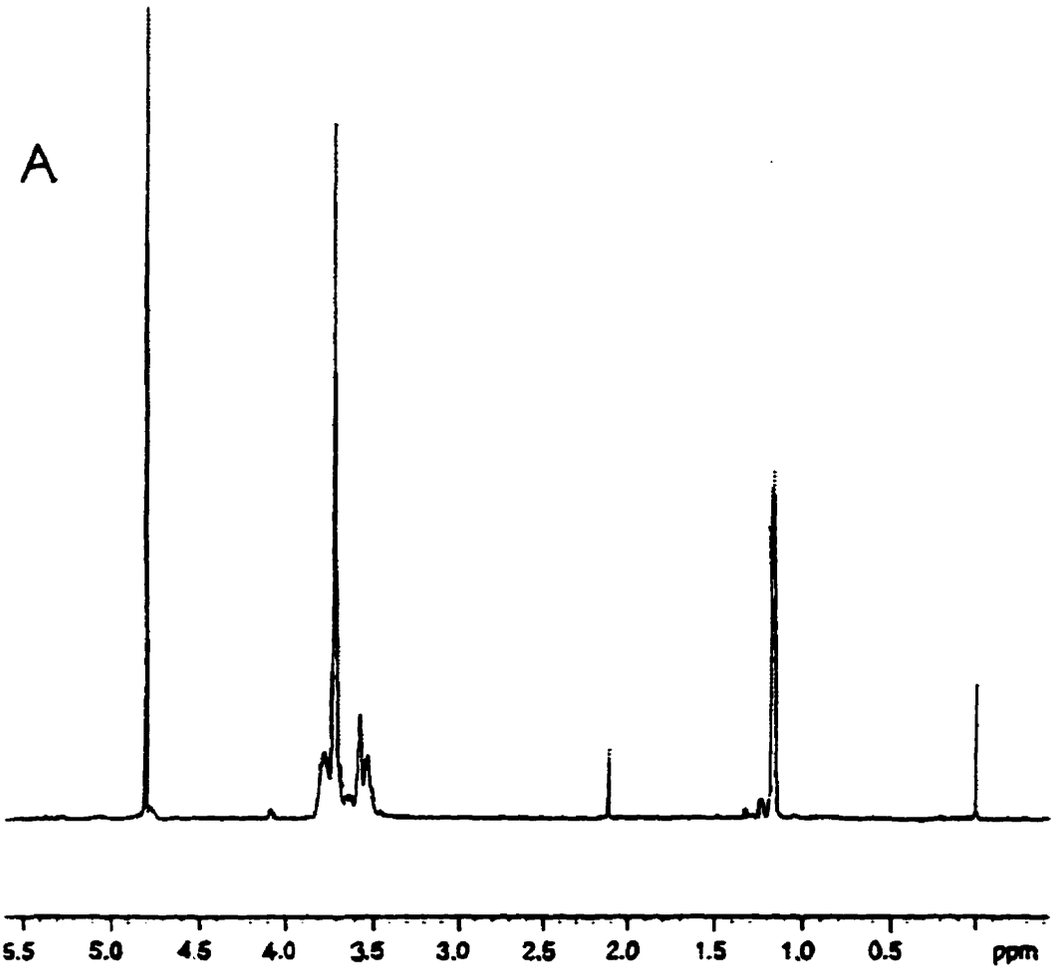

A
4.5
4.0
3.5
3.0
2.5
PPM
2.0
1.5
1.0
.5
0.0
B
4.5
4.0
3.5
3.0
2.5
PPM
2.0
1.5
1.0
.5
C
9.0
8.0
7.0
6.0
5.0
PPM
4.0
3.0
2.0
1.0
0.0

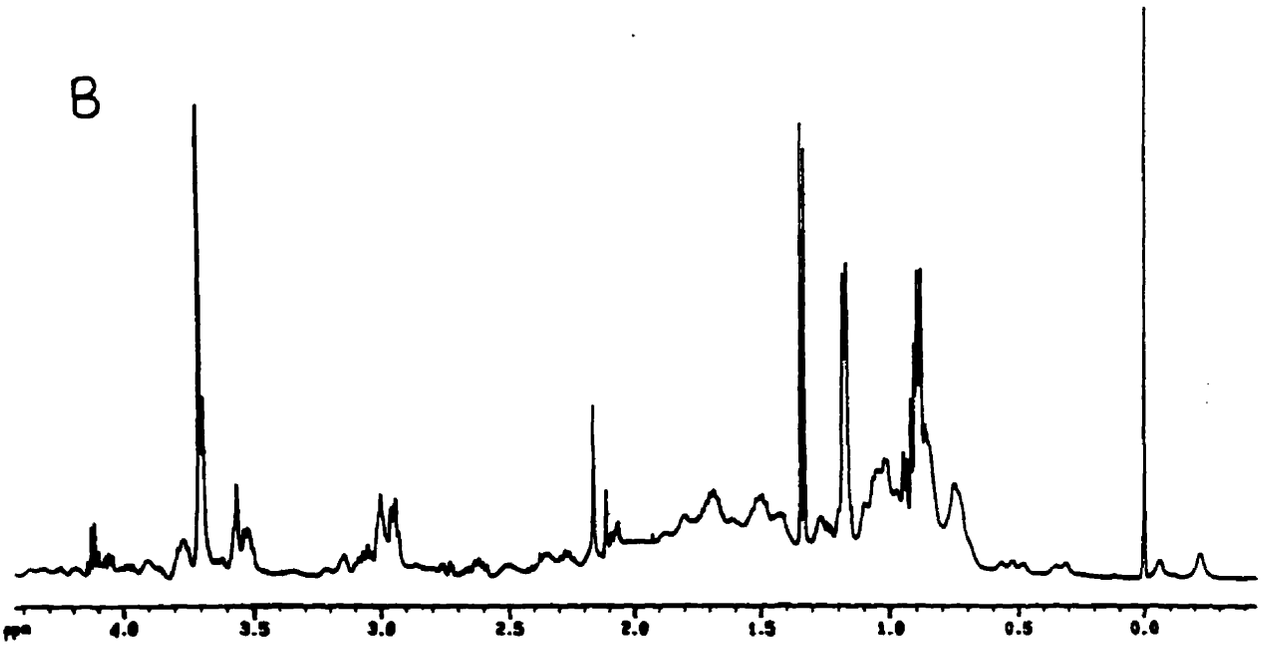
Figure 2.33. NMR spectra obtained for LaL and HEWL following dialysis.

(A) 500 MHz ^{1}H NMR spectrum of a sample obtained from dialysis and then lyophilization of [*methyl*-^{13}C]methionine labelled LaL against 5 mM KPB, pH 7.0. The protein and buffer concentrations are ≈ 0.1 mM and 54 mM respectively from calculations of the mass of lyophilized protein/salts obtained (2.2.10.2). The spectrum was taken in D_2O with supression of the HOD signal and reference to TSP (0.00 ppm).

(B) 500 MHz ^{1}H NMR spectrum of *wild type* LaL prepared as described in (A). The protein and buffer concentrations are ≈ 1 mM and 30 mM respectively and the spectrum was acquired as in (A).

(C) 250 MHz ^{1}H NMR spectrum of HEWL prepared as described above in (A). The NMR sample contained ≈ 2 mM HEWL and 35 mM KPB, pH 7.0 and the spectra was taken in D_2O with reference to HOD (4.6 ppm).


A
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm


B
ppm
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

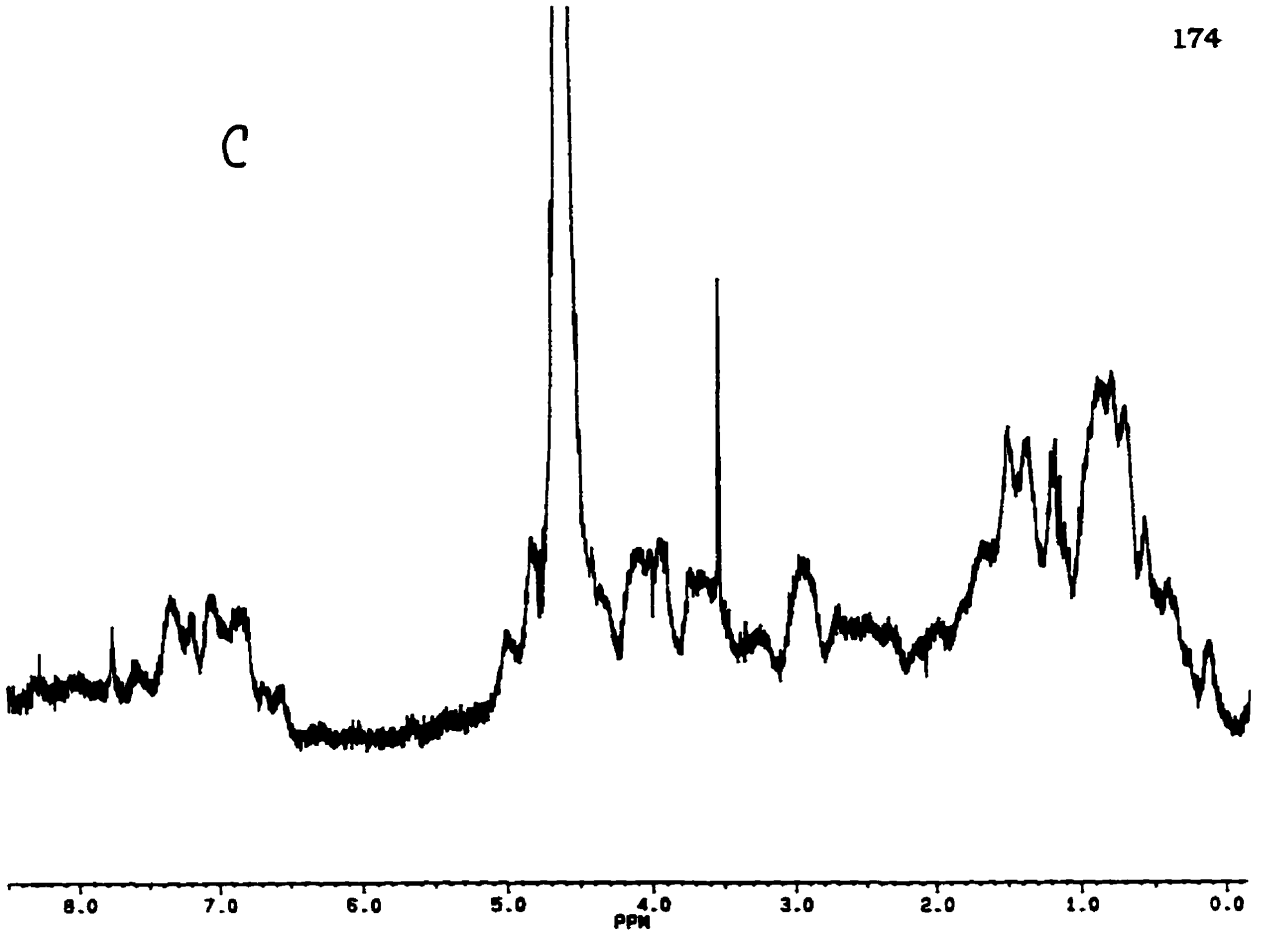


Figure 2.33 (C).

since they do not bind to the cation exchange column. The basic nature of LaL and HEWL may actually promote their binding to these tubing contaminants.

It was found that switching to the higher grade Spectra/Por® 7 tubing substantially reduced, but did not completely eliminate the contamination problem. In addition, the greater the amount of protein dialysed per amount of tubing used, lessened the proportion of contaminant(s) to protein (compare the spectrum for Fig 2.33 B, which LaL is more concentrated than in A). Contamination from dialysis tubing is apparently a common problem (Dr. T. Viswanatha, personal discussions) causing difficulties in both absorption and fluorescence spectroscopy. Lyophilization following dialysis against dilute buffer appears to be the only feasible method of securing low salt preparations of LaL that can be subsequently dissolved at high concentrations, an essential pre-requisite for crystallography and NMR studies. Since alternative methods towards this purpose were not apparent, it is unfortunate that this contamination problem appears to be unavoidable. Different brands of tubing have not been examined but it is feared that similar problems would arise.

2.3.6. Crystallization Efforts

The Crystal Screen crystallization kit provided a convenient method to possibly identify initial conditions for the crystallization of LaL. The set of 50 reagents (listed in Appendix B) encompass a variety of salt, buffer and precipitant compositions. A filter sterilized solution of LaL in MQW was mixed with each of the 50 reagents (which are also sterile) such that the final concentration of LaL was 6.06 mg/mL. Mixing of the protein with several of the reagents (No. 20, 30, 31, 49 and 50) resulted in immediate precipitation or a "cloudy" solution, while with others (No. 15 and 40), precipitation ensued within 10-20 min. Except for No. 40, each of the precipitating causing reagents contained either ammonium (15, 20, 30, 31) or lithium (49, 50) sulfate as a salt additive in combination with either PEG 4000 or PEG 8000 as the precipitant. Crystallization samples were then incubated at both 23 °C and 4 °C. Over a period of 2 months, many of the other samples also produced cloudy solutions. This was generally observed after 4-10 days and the precipitation either remained the same or became somewhat heavier.

The following observations were made:

1) Samples in which the solution appeared cloudy or precipitation was noted:

- at 23 °C with reagents 3, 4, 7, 8, 33, 36, 37, 47, 48
- at both 23 and 4 °C with reagents 1, 12, 19, 20, 24, 28, 30, 31, 38, 40, 45, 49, 50

2) Samples that remained essentially clear

- at 4 ° C with reagents 3, 4, 7, 8, 33, 36, 37, 47, 48
- at both 23 and 4 °C with reagents 2, 5, 6, 9-11, 13-18, 21-23, 25-27, 29, 32, 34, 35, 39, 41-44, 46.

Precipitation in some of the 23 °C samples may have resulted from evaporation of the reagent in the reservoir since samples with reagents 1, 3, 19, 22, 30, 34, 36, and 48 (kept at 23 °C) eventually dried out.

None of the samples appeared to result with protein crystal growth, or at least crystals could not be identified. It is possible that some of the amorphous material noted as a precipitate may have contained microcrystals. Amorphous and microcrystalline precipitates can often be distinguished on the basis of their birefringence. A microcrystalline precipitate may "glow" if examined between crossed polarizing lenses (Carter, 1992). However, at the time, such a microscope was not readily available. Three of the samples (with reagents 10, 22 and 32) gave which appeared to resemble a ring of crystals which formed on the outer edge of the sample drop. Closer examination and probing of the drops with a sharp needle revealed that what was observed resulted from a thin film or crust on the surface of the drops.

Due to the dedication of time and resources required for intensive crystallization studies, no further efforts to crystallize LaL were performed in our laboratory.

In July 1993, we entered into a collaboration with Dr. Fusao Takasagawa from the Dept. of Chemistry, University of Kansas (Lawrence, KS). On three occasions, a total of 75 mg of LaL purified from *E. coli* TG-1/pHDM10 was sent to his laboratory. Crystal growth was attempted primarily using the following bath: 12-20% PEGME 2000 (polyethylene glycol methylether, MW 2000), 100 mM Tris (pH 8.5), 10 mM NiCl, 100 mM NaCl. These conditions appeared promising at first but it was determined that only salt crystals were obtained. The collaboration was mutually terminated in May 1995 when Dr. Takasagawa could no longer dedicate the required resources in his laboratory towards this collaboration. We are extremely grateful for his considerable efforts.

A second collaboration was initiated in September 1995 with Dr. Albert Berghuis from the Dept. of Biochemistry, McMaster University (Hamilton, ON). At the time of preparation of this thesis, a total of 120 mg of LaL purified from *E. coli* BL21/pLR102 has been provided to Dr. Berghuis. The results have become very promising as conditions have been found that have resulted in obtaining crystals of LaL, some of which have diffracted to a resolution between 2.3 and 3.0 Å. Dr. Berghuis and Adelaine Leung, a graduate student with Dr. Berghuis and the primary investigator involved with the crystallization, have begun searching for a heavy atom derivative.

2.3.7. Summary and Future Work

We have incorporated the bacteriophage lambda *R* gene onto four plasmids for the inducible expression of LaL in *E. coli* and were able to purify the enzyme to apparent homogeneity from three of these systems. The plasmids pLR7 and pHDM10 were constructed with the isolation of λ DNA fragments onto the expression plasmid pTTQ18. An approximate 30-fold increase in purified LaL was obtained from induction of *E. coli* TG-1 harboring pHDM10 (6-7 mg/L) than was for *E. coli* TG-1 harboring pLR7 (0.2-0.3 mg/L). The size of the λ DNA fragment, most notably the DNA present 5' to the *R* gene, was considerably reduced in pHDM10 which appears to be partially responsible for the increased expression of LaL directed from this plasmid. We have also employed the polymerase chain reaction to specifically amplify the *R* gene from pHDM10 generating a DNA fragment flanked with unique restriction sites for the optimized isolation of the DNA onto the expression plasmid pET-22b. Induction of *E. coli* BL21(λDE3) harboring the construct pLR102 permitted purification of LaL at levels of 25-30 mg per litre of culture. Therefore, an extremely effective plasmid directing the inducible expression of LaL has been prepared.

Purification of LaL was achieved following sequential cation-exchange (S-Sepharose, Mono-S) and hydrophobic (Phenyl-Superose) chromatography. Purification of LaL using the protocol developed from *E. coli* TG-1/pHDM10 and BL21/pLR102 was established from N-terminal sequencing and ESMS analysis of the isolated proteins. LaL can be effectively and conveniently stored as a powder following its lyophilization from a dilute buffer. Accurate and concurring methods to quantitate LaL have been described which has included the determination of the molar extinction coefficient for this enzyme. The purified enzyme demonstrated excellent stability at higher concentrations and the enzyme activity of dilute preparations of LaL could be maintained in an appropriate buffer. Certain difficulties were encountered upon application of commonly employed protein manipulation techniques with LaL.

The availability of an effective expression system and purification protocol sufficiently justifies the preparation and characterization of site specific mutants of LaL. However, it may be prudent to delay these studies until the crystal structure of LaL is solved and rationale choices for mutants can be made based on the tertiary structure. We eagerly await the anticipated structure and strides towards this end that have been provided from our collaboration with Dr. Berghuis are encouraging. In order to possibly

accelerate the process of obtaining a heavy-atom derivative, the preparation of LaL containing selenomethionine is currently in progress. Selenomethionine containing proteins have demonstrated a practical approach in solving the phase problem associated with protein crystallography (Hendrickson et al., 1990). Future interests could also centre on the determination of the solution structure of LaL by multi-dimensional NMR analysis, albeit, another formidable undertaking. The high level overexpression of LaL uniformly labelled with ^{15}N and ^{13}C, a prerequisite for such a study, is now possible with our current system.

CHAPTER 3

Turbidimetric Assay, Interactions of Saccharides and Peptides with LaL and Synthesis and Evaluation of Alternative Substrates

3.1. INTRODUCTION

The natural substrate for lambda lysozyme is the *E. coli* peptidoglycan. However, very little is known about the distinctive structural attributes of the peptidoglycan which are necessary for binding to and ultimately required for LaL activity. In order to explore the substrate and ligand specificities of LaL, we have investigated the binding and potential substrate propensity of several saccharides including a detailed study of chitooligosaccharides with LaL. In addition, studies on the importance of peptide binding to LaL were initiated. In Chapter 1, the intricacies associated with the cleavage and binding modes of the simple chitooligosaccharide substrates have been presented. This introduction will bring attention to the unique specificities of lysozymes from different origins for "peptidoglycan-like" compounds that are pertinent to binding and important for catalysis and some of the approaches that have been applied to measure lysozyme activity with various substrates. Unlike most enzymes and their well-defined substrates, lysozymes are much smaller than the structurally complex peptidoglycan and it can be easily appreciated that a "lock and key fit" can not adequately describe the recognition of substrate to enzyme. A greater appreciation and interpretation of our results can be made with the knowledge that has accumulated for other lysozyme systems.

3.1.1. Methods to Demonstrate Lysozyme Activity

3.1.1.1. Bacteriolytic Assays

Bacteriolytic assays are based on monitoring the lysis of bacterial cells and are the most widely employed of all the methods used for lysozyme activity measurements. Bacteriolytic activity is commonly demonstrated by way of the turbidimetric assay. The premise of the turbidimetric assay (for all lysozymes studied by this method) lies in the ability of the lysozyme to degrade the peptidoglycan of bacteria suspended in a hypotonic buffer. This action will weaken the cell wall which will ultimately result in the osmotic lysis of the bacterial cells. The bacteriolytic activity is observed through measurement of

the rate of decrease in absorbance (clearing) of a suspension of the bacterial cells following the addition of an enzyme aliquot.

The use of a turbidimetric assay to measure HEWL activity was first reported by Shugar (1952) using a suspension of *Micrococcus lysodeikticus*. As this Gram-positive bacterium lacks an outer membrane, lysozyme present with the bacterial cells will have direct access to the peptidoglycan. However, in the case of *E. coli* and other Gram-negative bacteria, if lysozyme is added exogenously to a suspension of cells, the peptidoglycan is safeguarded from lysozyme degradation due to the presence of the outer membrane. Therefore, in order for Gram-negative bacteria to be susceptible to lysis by exogenously added lysozyme, the outer membrane must be sufficiently compromised or "sensitized" such that the protein molecules can gain access through the outer membrane and to the peptidoglycan.

A variety of the early techniques which render Gram-negative bacteria susceptible to lysis by lysoyzme have been described (Noller & Hartsell, 1961). The initial report on the purification of T4L utilized a turbidimetric assay using *E. coli* cells that were frozen for about 1 week and then lyophilized (Tsugita et al., 1968). Although this method was described with the occasional inability to generate cells sufficiently sensitive to the action of T4L (Tsugita et al., 1968), it is nonetheless customarily chosen in more recent investigations to describe the activity of this lysozyme and of its mutants (Anand et al., 1988; Dao-Pin et al., 1991; Zhang et al., 1993) and has also been applied to measuring LaL activity (Bienkowska-Szewczyk & Taylor, 1980). Sensitization of *E. coli* cells has also been achieved by means of EDTA treatment (McMacken et al., 1970). EDTA has been shown to weaken the outer membrane of *E. coli* (Birdsell & Cota-Robles, 1967; Watt & Clarke, 1994) presumably by removal of divalent cations which are believed to occupy membrane stabilizing binding sites (Hancock et al., 1994). More often however, *E. coli* cells having a superior sensitivity to lysis have been prepared by their treatment with chloroform. This treatment apparently solubilizes or disrupts segments of the outer membrane generating gaps of sufficient size to permit passage of the enzyme molecules. The original procedures describing such chloroform treatment (Sekiguchi & Cohen, 1964; Black & Hogness, 1969a; Koteswara Rao & Burma, 1971) have been adopted with slight modifications by others to improve the sensitivity of the cells (Jensen & Kleppe, 1972; Knight et al., 1987; Caldentey & Bamford, 1992; Ziermann et al., 1994). One group has chosen autoclaving as a pre-treatment of the bacterial cells (Vasala et al., 1995) although the effectiveness of this method is questionable.

The plate assay is another method based on observing lysis of the bacterial cells. A bacterial lawn is prepared on solid medium in a petri dish and the lawn is exposed to

chloroform vapours (Streisinger et al., 1961). Formation of a zone of clearing (or halo) due to cell lysis is the indication of activity. The method has been applied in which an enzyme solution is spotted directly onto the lawn (Hitoshi et al., 1978; Caldentey & Bamford, 1992) and as a rapid method for screening lysozyme activity expressed from phage (Alber & Matthews, 1987; Rennell et al., 1991). Recently, a novel approach has been reported in which lytic activity can be detected employing SDS-PAGE (Leclerc & Asselin, 1989; Audy & Asselin, 1989). In this method, the acrylamide gel is prepared containing a suspension of the cells and activity is visualized as a clear band in the gel. This method and the plate assay are useful for only qualitative descriptions of activity.

3.1.1.2. Insoluble Peptidoglycan Substrates

Although useful as an assay, detection of lytic activity does not indicate the specific bonds of the peptidoglycan that are cleaved as lysis can be promoted from the action of a glycosidase (N-acetylglucosamidase or N-acetylmuramidase), N-acetylmuramyl-L-alanine-amidase, or an endopeptidase (i.e. peptidoglycan hydrolases). More specific assays have emerged using insoluble, isolated peptidoglycan as the substrate. Use of peptidoglycan as a substrate can permit identification of released reducing groups, from the action of a glycosidase, or detection of released amino- or carboxy-terminal from the action of an amidase or endopeptidase respectively (Ghuysen et al., 1966). These methods necessitate isolation of the peptidoglycan and general procedures have been described (Sharon & Jeanloz, 1964; Wheat, 1966; Rogers et al., 1980). More elaborate methods have been developed for the preparation of *E. coli* peptidoglycan free of contaminating material (Hoyle & Beveridge, 1984; Glauner, 1988).

Essentially, the peptidoglycan is incubated with the peptidoglycan hydrolase which will liberate soluble fragments. The residual insoluble peptidoglycan is removed (by filtration or centrifugation) and the solubilized fragments can then be separated chromatographically or the specific functional groups mentioned above can be detected. For lysozyme, activity can be assayed by measuring the production of reducing sugars (Park & Johnson, 1949; Ward, 1973) or appearance of *N*-acetylamino sugars (Ghuysen et al., 1966). The use of insoluble peptidoglycan as a substrate has also been extended with the specific labelling of the peptidoglycan. Radiolabelled peptidoglycan can be prepared from bacteria grown in the presence of a radiolabelled precursor (eg. [^{3}H]DAP, *L*-[^{14}C]Ala, [acetyl-^{3}H or -^{14}C]GlcNAc) and following its enzymatic digestion, the solubilized fragments can be measured for radioactivity (for examples see: Inouye et al., 1973; Höltje et al., 1975; Jensen et al., 1976; Ursinus & Höltje, 1994). Reports on the post-isolation labelling of the peptidoglycan peptide with fluorescent compounds and their application as

lysozyme substrates have also appeared (Mintz et al., 1975; Maeda, 1980; Moser et al., 1988).

The assays based on the detection of cell lysis or a specific peptidoglycan degrading activity are accompanied by an inherent heterogeneous composition of the substrate. The "peptidoglycan substrate" (i.e. either isolated or part of the cell wall) will be lacking in uniformity with respect to the glycan chain lengths, peptide substitution and peptide crosslinking (these variations have been discussed in section 1.2). As such, different "segments" of the peptidoglycan macrostructure will differ in their chemical composition and polymeric arrangement and can therefore exhibit varying degrees of susceptibility to enzymatic digestion (specific examples are brought to light in section 3.1.5). These complications are heightened when whole cells are the substrates and other gross features of the cell wall are present with the peptidoglycan. For example, the lipopolysaccharides present in the outer membrane of Gram-negative bacteria have been shown to interact with and inhibit HEWL (Ohno & Morrison, 1989a,b; Ohno et al., 1991). Strominger and Tipper (1974) have suggested that the negatively charged teichoic acids of the Gram-positive cell wall can bind basic proteins, a common feature of lysozymes, and provide some resistance to peptidoglycan degradation. In addition, the *O*-acetylation of peptidoglycan at the C6-hydroxyl group of muramic acid residues in both Gram-negative and Gram-positive species confers resistance to the hydrolytic action of several muramidases including HEWL (Clarke & Dupont, 1992; Clarke, 1993). In spite of these complications, intact cells and peptidoglycan are routinely utilized for the assay of lysozymes.

3.1.1.3. Soluble Chemically-Defined Substrates

As quickly as it was realized that the glycosidic bond between MurNAc and GlcNAc was hydrolysed by HEWL (Salton & Ghuysen, 1960), much interest was demonstrated in preparation of small molecular weight, soluble substrates. Substrates in which only a limited number of bonds (preferably only one) can be acted upon are of great interest to determine the specificity of lysozymes and to evaluate the enzyme's kinetic parameters. Although a paucity of such compounds presently exists, especially for application to phage lysozymes, some advances in this area have been made. The types of small molecular weight, soluble substrates that have been investigated with lysozymes will be divided into two classes.

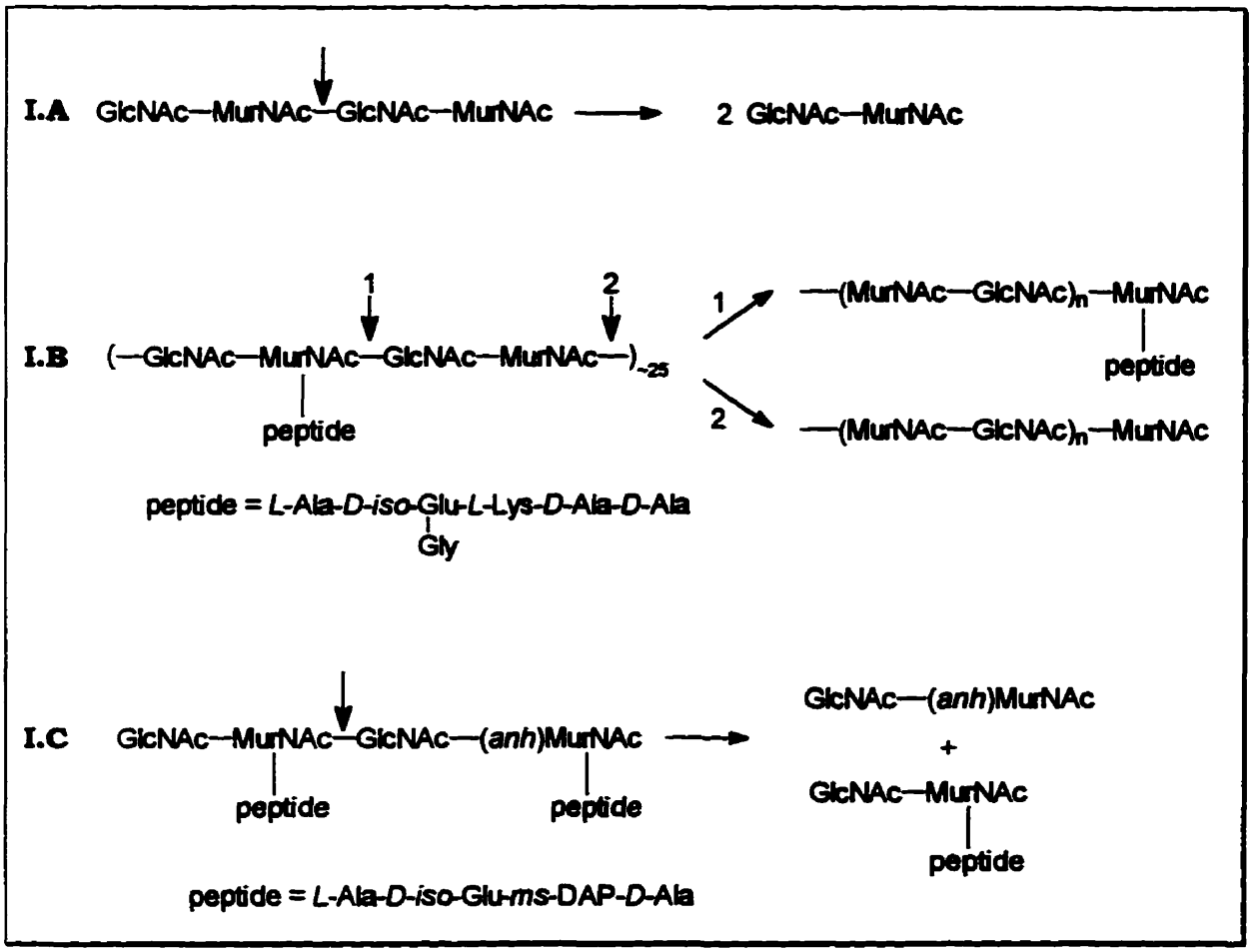


Figure 3.1. Class I soluble substrates for lysozymes. The substrates can be cleaved by lysozyme at the positions indicated by the arrows giving the indicated products. For **I.B**, cleavage at position 1 or 2 is dependent on the specificity of the particular lysozyme. See text for other details.

3.1.1.3.1. Class I: Substrates Derived from Peptidoglycan

Class I (Fig. 3.1) includes oligosaccharides and glycopeptides that have been obtained from peptidoglycan. The tetrasaccharide **I.A** (Fig. 3.1) has been isolated from the digestion of peptidoglycan from *M. lysodeikticus* by HEWL (Sharon, 1967). Only 50% of the MurNAc residues of the peptidoglycan from this bacteria are substituted with peptide (refer to section 1.2). Because of this composition, the peptide-free tetrasaccharide **I.A** can be obtained. The substrate **I.B** is a soluble, uncrosslinked peptidoglycan that is secreted into the culture medium when *M. lysodeikticus* cells are grown in the presence of penicillin (Mirelman et al., 1974). The average chain length of **I.B** is approximately 50 disaccharide units in which, again, only half are substituted with the hexapeptide (Fig. 3.1). The molecular weight of **I.B** is approximately 38 kDa and is therefore, not a small molecular weight substrate but is included in Class I since it is a soluble compound.

The substrates **I.A** and **I.B** have been primarily exploited for two purposes. In the late 1960's and early 1970's when interest in lysozyme research had expanded to include lysozymes from different origins, the specificities of the lysozymes were customarily examined with these substrates (specific reference to the use of these compounds is discussed in section 3.1.2). At that time, lysozyme activity was broadly characterized as an ability to lyse bacterial cells or to solubilize peptidoglycan. The compounds **I.A** and **I.B** were used to unambiguously demonstrate muramidase activity. For example, cleavage of **I.A** into the disaccharide GlcNAc-MurNAc or the release of compounds with a terminal MurNAc residue from **I.B** would indicate hydrolysis of these substrates between MurNAc and GlcNAc. Secondly, the compounds were used to probe the importance of the presence of a MurNAc residue or peptide substituent in the substrate for the lysozyme to hydrolyze the particular saccharide.

No detailed kinetic parameters have been determined for the compounds **I.A** and **I.B**. Lysozyme activities on these compounds have been only qualitatively described as either strong, weak or non-existent (see section 3.1.2). However, the tetrasaccharide **I.C** (Fig. 3.1), designated as the muropeptide CB and is obtained from the LaL digests of *E. coli* peptidoglycan (Taylor et al., 1975), has found use as a substrate for T4 lysozyme. The compound **I.C** (or CB) is completely hydrolysed by T4L and a K_m value of 1.6 mM was determined (Bienkowska & Taylor, 1979). The compound is an equally good substrate for HEWL, but no kinetic data was given. It is surprising that the muropeptide **I.C** has not received more wide-spread application as a substrate for T4L, especially for the characterization of the many mutant forms of this enzyme. Only recently has this compound again been used as a substrate for T4L (Kuroki et al., 1995). However, in this report, the (*anh*)MurNAc was reduced to a muramicitol residue and the reduced substrate was found to be hydrolysed at a rate of 0.868 $M^{-1}\cdot min^{-1}$ by the *wt* T4 lysoyzme.

3.1.1.3.2. Class II: Substrates Derived from Chitin

The Class II substrates include chitin and oligosaccharides derived from GlcNAc (Fig. 3.2). The chitinase activity of HEWL was first demonstrated by Berger and Weiser (1957). Chitin (**II.A**, Fig. 3.2) is a homopolymer of β(1→4)-linked GlcNAc. Chitin can form an array of intra- and inter-chain hydrogen bonds which produces a highly insoluble polymer (Blackwell, 1988) and is therefore impractical as a lysozyme substrate. However, chitin can be made soluble by glycolation with ethylene oxide or ethylene chlorohydrin. The resulting glycol chitin (**II.B**, Fig. 3.2) serves as an effective soluble substrate for some c type lysozymes in which the standard assay involves measurement of increased reducing

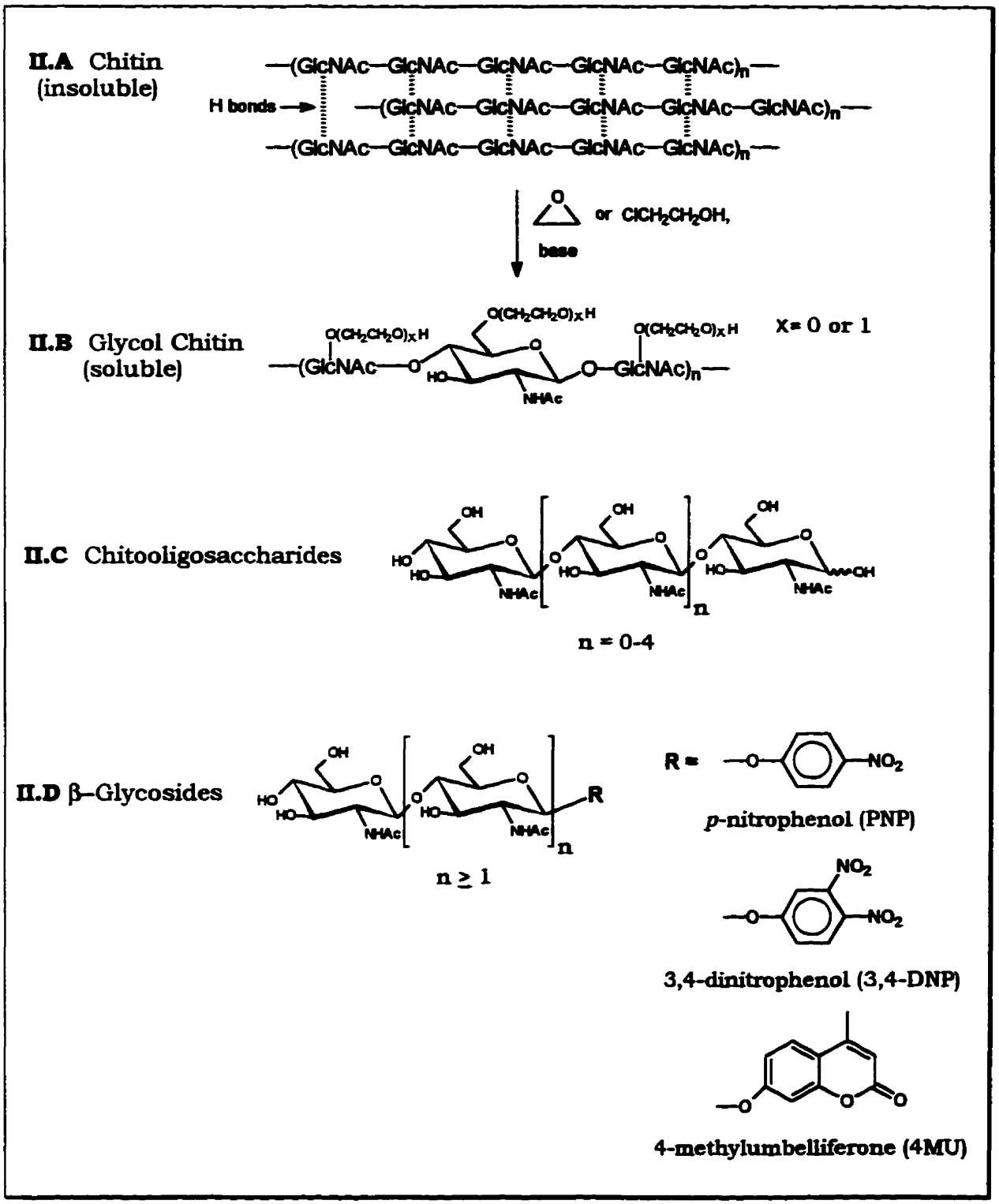


Figure 3.2. Class II soluble substrates for lysozymes.

groups released from the glycolated chitin (Yamada & Imoto, 1981). The use of glycol chitin has been widely adopted to scrupulously measure activities of chemically modified or mutant forms of HEWL (Yamada et al., 1981; Yamada et al., 1982; Kuroki et al., 1986; Malcolm et al., 1989; Kumagai et al., 1992b; Inoue et al., 1992a; Maenaka et al., 1994; Matsumura & Kirsch, 1996a,b) and human lysozyme (Kikuchi et al., 1988; Kuroki et al., 1989). However, in these and most cases, bacteriolytic activities are also determined in tandem by turbidimetric analysis. It is interesting that the glycolation at the C6-hydroxyl group of GlcNAc does not inhibit the chitinase activity of HEWL considering that acetylation at this position on muramic acid appears to inhibit the bacteriolytic activity of HEWL.

The next group of Class II substrates are the small molecular weight oligomers of GlcNAc (chitosaccharides, chitooligosaccharides, $(GlcNAc)_n$; **II.C**, Fig. 3.2). The different chitosaccharides (from dimer to hexamer) are commercially available but can be prepared from the partial acid hydrolysis of chitin and fractionation of the products by chromatography over charcoal or by gel filtration (Rupley, 1964; Raftery et al., 1969; Banerjee et al., 1973; Van Eikeren & McLaughlin, 1977). Chitosaccharide substrates are limited to the *c* type lysozymes and even in these cases, their use is beset by a number of difficulties: (i) $(GlcNAc)_{n<4}$ are cleaved very slowly, require large amounts of substrate and enzyme and therefore, only the penta- and hexasaccharide are effective substrates (Rupley & Gates, 1967); (ii) apart from the hexamer (which is uniquely cleaved between the fourth and fifth residue) the cleavage patterns are complex; and (iii) hydrolysis of $(GlcNAc)_n$ is accompanied with glycosyl transfer reactions in which $(GlcNAc)_n$ acts as an acceptor instead of water. These topics have been addressed previously in section 1.9.1.

Chitosaccharides have received limited application and there exists a dearth of information on their use for quantitative kinetic measurements. Although the HEWL catalyzed reaction with $(GlcNAc)_6$ has been investigated in some detail (Banerjee et al., 1973) and a report on the activity of several egg white lysozymes on $(GlcNAc)_5$ has appeared (Fukamizo et al., 1983), the chitosaccharides have not gained general acceptance as an assay method. Like the Class I substrates **I.A** and **I.B**, chitosaccharides have been primarily examined as to whether or not they can serve as substrates for lysozymes from different origins (see 3.1.2). As will be discussed, activity with chitosaccharides has been identified only for the *c* type lysozymes. HPLC methods, in which the hydrolysis products are separated and quantified, have been reported to be the most efficient and sensitive means of monitoring the activity with chitosaccharides (Van

Eikeren & McLaughlin, 1977; Usui et al., 1987) although the observance of the increase in reducing sugar has also been applied (Banerjee et al., 1973).

The methodology associated with the detection of reducing groups or HPLC analyses using chitosaccharides are tedious and has prompted the development of substrates for potential use in more convenient assays. The search for chromogenic substrates has led to the synthesis of β-arylglycosides of $(GlcNAc)_n$. Of these, the *p*-nitrophenylglycosides have received the greatest attention (see group **II.D**, R = PNP, Fig. 3.2). Osawa (1966) was the first to synthesize $(GlcNAc)_2$-PNP (**II.D**, n=1) and demonstrated that HEWL catalyzed release of PNP from this compound (which can be measured at 400 nm) but not from GlcNAc-PNP. However the hydrolysis of $(GlcNAc)_2$-PNP was extremely slow (only ≈ 25% of the glycoside was liberated after 100 hr incubation) and it was also shown that the pathway for hydrolysis of $(GlcNAc)_2$-PNP was complex and involved transglycosylation (Rand-Meir et al., 1969). The need for enhanced sensitivity has led to the chemical (Osawa & Nakazawa, 1966; Ballardie et al., 1977; Nanjo et al., 1988a) and chemoenzymatic (Usui et al., 1988) synthesis of the higher oligomeric β-PNP-glycosides (**II.D**, n=2-3). Table 3.1 lists values relating to the production of nitrophenol from these compounds.

Table 3.1. Values of k_{cat}/K_m for the hydrolysis of nitrophenyl-glycosides by HEWL.

Substrate	k_{cat}/K_m (M^{-1} s^{-1})
$(GlcNAc)_6$	18000 [a, b]
$(GlcNAc)_2$-PNP	0.0012 [c]
$(GlcNAc)_3$-PNP	0.18 [d], 0.02 [a]
$(GlcNAc)_4$-PNP	0.90 [d]
$(GlcNAc)_2$-3,4-DNP	0.064 [d]
$(GlcNAc)_3$-3,4-DNP	7.9 [d]
$(GlcNAc)_4$-3,4-DNP	205 [d]

[a] Imoto et al. (1972); [b] Banerjee et al. (1973); [c] Lowe (1967); [d] Ballardie et al. (1977).

The most efficient of the PNP-glycosides, $(GlcNAc)_4$-PNP, is still hydrolysed 2×10^4 times more slowly than $(GlcNAc)_6$. In order to obtain better substrates, Ballardie et al. (1977) have also prepared and investigated the reactions of 3,4-dinitrophenyl-glycosides

(see group **II.D**, R = 3,4-DNP, Fig. 3.2) with HEWL (Table 3.1). The 3,4-DNP is a better leaving group than PNP and these substrates were found to be substantially more effective than the PNP-glycosides. The dependence of pH on k_{cat} and k_{cat}/K_m for the hydrolysis of $(GlcNAc)_4$-3,4-PNP catalyzed by HEWL (Ballardie et al., 1977) was very similar to that reported for $(GlcNAc)_6$ (Banerjee et al., 1973) suggesting that these substrates are hydrolysed by the same mechanism. Furthermore, comparable kinetic constants for the hydrolysis of $(GlcNAc)_4$-3,4-PNP were found for the c type lysozymes from human and duck egg white. The observation that k_{cat}/K_m for the hydrolysis of $(GlcNAc)_6$ is approximately 90 times greater than for $(GlcNAc)_4$-3,4-PNP indicates the importance of binding of the terminal $(GlcNAc)_2$ residue of the former to the E and F sites in HEWL. Even though 3,4-DNP (i.e. from hydrolysis of $(GlcNAc)_4$-3,4-PNP) is expected to be a better leaving group than $(GlcNAc)_2$ (i.e. from hydrolysis of $(GlcNAc)_6$), the binding of the two terminal GlcNAc residues in the E/F sites of HEWL is thought to account for the greater reactivity of $(GlcNAc)_6$ (Ballardie et al., 1977).

Recently, the preparation of $(GlcNAc)_5$-PNP (Nanjo et al., 1988a; Usui et al., 1988) has been described and this substrate has been investigated with HEWL and human urinary lysozyme (Nanjo et al., 1988b). A partial presentation of some of the results of this work are shown in Figure 3.3. It can been seen that the cleavage pattern for each of $(GlcNAc)_n$-PNP (n=3-5) is complex and the amount of PNP liberated is small for each substrate with both lysozymes. Although the kinetic parameters of the cleavage of $(GlcNAc)_5$-PNP were not reported, the rate of production of the respective products increases from the trimer to the pentamer (products were separated and quantitated by HPLC analysis; note the different time scales for the activity curves in Fig. 3.3). For each substrate and lysozyme (except the human lysozyme and $(GlcNAc)_3$-PNP) the predominant hydrolysis product is GlcNAc-PNP. Chain elongation increases the rate of GlcNAc-PNP formation. The assay was extended to a colorimetric determination of PNP liberated from the predominating GlcNAc-PNP product by a coupled reaction involving β-N-acetylhexosaminidase (NAHase). The coupled assay (for the reaction of HEWL with $(GlcNAc)_5$-PNP) increased the sensitivity approximately 100 times as compared to that without NAHase. Enhanced sensitivity was also observed for the coupled assay with $(GlcNAc)_5$-PNP as compared to $(GlcNAc)_4$-PNP for both lysozymes. Some reports on the application of $(GlcNAc)_5$-PNP and this methodology have appeared (Kumagai et al., 1993; Maenaka et al., 1994; Yamada et al., 1994).

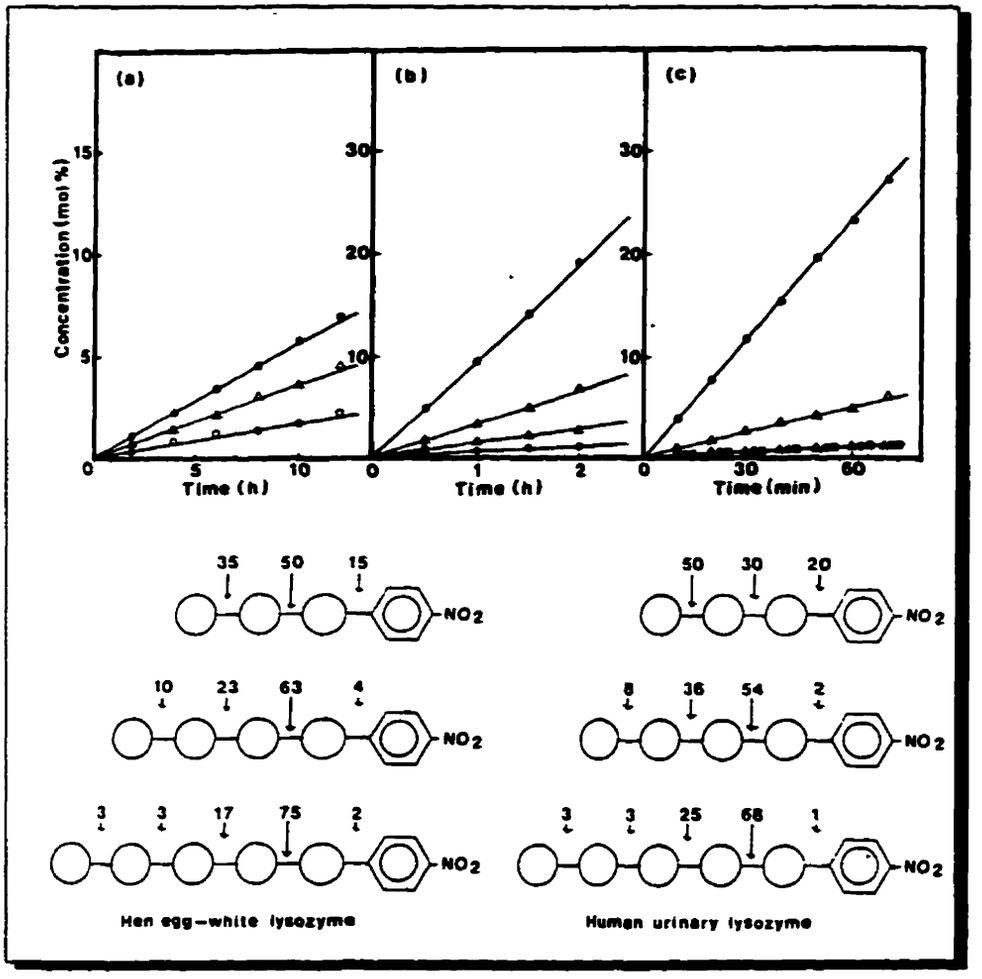


Figure 3.3. Reaction of $(GlcNAc)_n$-PNP with HEWL and human urinary lysozyme. Top: Time course of products formed from $(GlcNAc)_n$-PNP by HEWL. (a) n = 3; (b) n = 4; (c) n = 5. ○, PNP; ●, GlcNAc-PNP; △, $(GlcNAc)_2$-PNP; ▲, $(GlcNAc)_3$-PNP; □, $(GlcNAc)_4$-PNP. Bottom: Frequency of the enzyme catalyzed cleavage of glycosidic linkages of $(GlcNAc)_n$-PNP (n = 3-5). Reproduced from Nanjo et al. (1988b).

The final type of Class II lysozyme substrates are the 4-methylumbelliferyl-glycosides of GlcNAc (see group **II.D**, R = 4MU, Fig. 3.2). The lysozyme catalyzed release of 4MU from these substrates can be monitored fluorimetrically. Investigations of the reactions of $(GlcNAc)_{n\leq 3}$-4MU (Delmotte et al., 1975; Yang & Hamaguchi, 1980) and $(GlcNAc)_4$-4MU (Fukada et al., 1985) with lysoyzme have appeared. As with the PNP-glycosides, it is evident that the 4MU-glycosides have not gained general acceptance as standard lysozyme substrates. It is of some interest that the 4MU-glycosides have not received greater attention. Although bulkier than the PNP group, the 4MU apparently binds favorably to the active site of HEWL. The association constant for $(GlcNAc)_3$-4MU

(5.6×10^4 M^{-1}; Delmotte et al., 1975) is not much less than that for $(GlcNAc)_3$ or $(GlcNAc)_4$ ($\approx 1\times10^5$ M^{-1}) but is approximately 45 times than that of $(GlcNAc)_3$-PNP (1.25×10^3 M^{-1}; Osawa & Nakazawa, 1966). In addition, the rate of 4MU liberation is nearly 1000 times that for PNP from the corresponding glycoside of $(GlcNAc)_3$ (Delmotte et al., 1975). It may be that due to practical considerations, the colorimetric determination of PNP is favoured over the more cumbersome fluorometric method for routine assays.

3.1.2. Substrate-Induced Evolution

Investigations of lysozymes from various origins have indicated that the specificities of lysozymes are not the same and that the enzymes have evolved with much different substrate requirements. The variation in the composition of the cell walls from bacteria (not only relating to the peptidoglycan but also to other cell wall components) and the relatively simple cell walls of fungi or insects composed of chitin, may have provided an incentive for the lysozymes to develop either a more broad or more specific substrate specificity depending on the nature of the substrate(s) they might encounter; i.e. a substrate-induced evolution (Schindler et al., 1977b).

The substrate specificities of representative lysozymes on three different substrate types are presented in Table 3.2. It is important to mention that substrate type I (Table 3.2) does not designate a general type of peptide substituted peptidoglycan. Specifically, Schindler et al. (1977b) investigated the lysozyme activity with the soluble uncrosslinked peptidoglycan secreted by *M. lysodeikticus* described previously (i.e. the Class I substrate **I.B**, Fig. 3.1). The discussion which follows will primarily emphasize the results presented in Table 3.2; other relevant studies comparing the substrate specificities of several lysozymes (Jollès et al., 1968; Jollès et al., 1974; Sharon et al., 1974; Fukamizo et al., 1983) or specifically those on T4L (Mirelman et al., 1975; Jensen et al., 1976) and GEWL (Arnheim et al., 1973b) have also been reported.

The *c* type lysozymes (HEWL and human) appear to act equally well on the peptide-substituted (type I) or unsubstituted (type II) peptidoglycan substrates as well on the chitooligosaccharide substrates (type III). GEWL (a *g* type lysozyme) acts upon the linear peptidoglycan substrates I and II and does so slightly better than does HEWL. However, HEWL's bacteriolytic activity against *M. lysodeikticus* was claimed to be more effective than that of GEWL (Schindler et al., 1977b). Furthermore, the goose enzyme acts on the chitosaccharides only very weakly and, in contrast to the *c* type lysozymes, does not promote reactions involving transglycosylation (Fukamizo et al., 1983). The specificity

Table 3.2. **Substrate specificity requirements of lysozymes. Hydrolysis of the substrate is indicated as well hydrolysed (+++) to not hydrolysed (-).**

Lysozyme	Type	Substrate Type I $(GlcNAc\text{-}MurNAc)_n$ peptide	II $(GlcNAc\text{-}MurNAc)_2$	III $(GlcNAc)_n$	Presumed main natural substrate (type)
HEW	*c*	++	+++	++	I, II, III
human	*c*	++	+++	++	I, II, (III)
GEW	*g*	+++	+++	+/-	I, II, (III)
T4	*p*	+++	-	-	I
Papaya	plant	+/-	+/-	++	III
Streptomyces erythraeus	*ch*	+		-	I, (II ?)
Chalarposis	*ch*	++	-	-	I

Data taken from Schindler et al. (1977b) except for *S. erythraeus* (from Morita et al., 1978) and *Chalarposis* (from Tipper & Strominger, 1966; Tipper et al., 1967; Hash, 1974; Fouche & Hash, 1978).

requirement of the phage T4L (a *p* type lysozyme) is much more stringent and acts only upon the peptide-substituted peptidgoglycan (type I substrate). The fungal and bacterial lysozymes (*ch* type lysozymes) appear similar to the phage lysozyme. They do not act upon GlcNAc oligomers but do act on *M. lysodeikticus* cell walls (the bacterial enzyme acting upon type I, type II or both as indicated in Table 3.2; the exact specificity as to type I or II has not been described). More interestingly is that both the fungal and bacterial lysozyme (*ch* type) are also capable of hydrolyzing the cell walls of *Staphylococcus aureus* which contain 6-*O*-acetylmuramic acid (see references given in Table 3.2).

The specificities of the different lysozymes can now be ascribed to their possible biological roles. Plants are susceptible to fungal invasion and the enhanced chitinase over peptidoglycan hydrolase activity of the papaya lysozyme would be more suited to provide protection against pathogenic fungi which often contain chitin in their cell walls (Schindler et al., 1977b). High concentrations of lysozyme are present in the egg-whites of many bird eggs and are thought to protect the embryo from bacterial infection and therefore, the specificity is for the bacterial cell wall (the egg and developing embryo do not produce immunoglobulins until ≈ 1 week before hatching). Depending on the bird species, the egg could hatch near water, in the soil, or on land and will therefore be exposed to the local flora of bacteria. The eggs of birds hatched near water, such as goose,

are more commonly exposed to Gram-negative bacteria. Shindler et al. (1977b) cite[3.1] that GEWL is more effective against *E. coli* than is HEWL which is in accord with the postulated relationship between environment and lysozyme substrate selection.

In order to fulfill their roles as lytic agents during host lysis, bacteriophage lysozymes must be active against peptide-substituted peptidoglycan. The specificity of phage lysozymes appear to be uniquely adapted to the peptidoglycan of their bacterial hosts (Table 3.3). Most of the Gram-negative phage lysozymes are active against their host or other Gram-negative bacteria and not against Gram-positive bacteria. The lysozymes from phage T4, P22 and 9NA are active against *M. lysodeikticus* but much less so than against their bacterial host. It was reported that activity of T4L toward *E. coli* cell walls is 2500 times greater than against the cell walls of *M. lysodeikticus* (Grütter & Matthews, 1982). The PRD1 lysozyme displays a broad specificity for several Gram-negative bacteria,

Table 3.3. Bacterial substrate specificity of Gram-negative and Gram-positive phage lysozymes.

Lysozyme[†]	Primary Host	Active Against	Not Active Against	Reference
Gr (-) Phage				
T4	EC	EC >>> ML		Grütter & Matthews (1982)
Vi II	EC, ST	EC, ST	ML	Taylor et al. (1970); Taylor & Gorazdowska (1974)
T7	EC	EC	ML	Inouye et al. (1973) Cheng et al. (1994)
9NA	ST	ST, EC >>> ML		Verma & Siddiqui (1986)
PRD1 (P15)	EC, ST, PA	EC, PP > Ec > PA, ST	ML, BS, LL	Caldentey et al. (1994)
P22	ST	ST, EC >> ML		Koteswara Rao & Burma (1971)
Gr (+) Phage				
PL-1	LC	LC	LSs, ML, BS	Hayashida et al. (1987)
LL-H lysin	LD	LD, LA, PD	LC, BS, ML, SP, EC	Vasala et al. (1995)
ϕ29	BS	ML		Saedi et al. (1987)

[†] All lysozymes have demonstrated or are assumed to be N-acetylmuramylhydrolases except T7 which is an N-acetylmuramyl-L-alanine amidase. Abbreviations:
Gram-negative bacteria: **EC**, *Escherichia coli*; **Ec**, *Erwinia carotovora*; **PA**, *Pseudomonas aeruginosa*; **PP**, *Pseudomonas phaseolicola*; **ST**, *Salmonella typhimurium*.
Gram-positive bacteria: **BS**, *Bacillus subtilis*; **LA**, *Lactobacillus acidophilus*; **LC**, *Lactobacillus casei*; **LD**, *Lactobacillus delbrueckii*; **LL**, *Lactococcus lactis*; **ML**, *Micrococcus lysodeikticus*; **PD**, *Pediococcus damnosus*; **SP**, *Streptococcus pneumaniae*; **LSs**; several *Lactobacilli* and *Streptococci* species.

[3.1] A specified reference was not given. Whether GEWL is more effective than HEWL with *E. coli* is debatable considering that supporting (Arnheim et al., 1973b) and contradicting (Matsumura & Kirsch, 1996) reports on GEWL's superior activity have appeared.

however, with varying efficiency (even amongst different species of *Pseudomonas*) but is not active towards Gram-positive bacteria. The lysozyme from the Gram-positive phage LL-H is active towards some Gram-positive bacteria but not to others, and the disparity exists even amongst *Lactobacilli* species (i.e. active against LD but not LC, Table 3.3). It is evident that phage lysozymes have evolved with specificity towards only very closely related peptidoglycans. There must be sufficient differences in the structures of the peptidoglycans (or cell walls) from even closely related bacterial species which can influence the activity of lysozymes. Furthermore, the activities of LaL, T4, P22 and 9NA lysozymes towards *E. coli* are 200 (Bienkowska-Szewczyk & Taylor, 1980), 250 (Grütter & Matthews, 1982), 233 (Kotesara & Burma, 1971) and 270 (Verma & Siddiqui, 1986) times that of HEWL respectively.

3.1.3. Comparative Inhibitory Behaviour

Not only are the substrate specificities of lysozymes distinct, but so is their disposition towards inhibition. Table 3.4 compares the inhibition of the bacteriolytic activity of representative lysozyme types. Chitooligosaccharides and the tetrasaccharide $(GlcNAc\text{-}MurNAc)_2$ are effective inhibitors of only HEWL (and other *c* type lysozymes; Jollès et al., 1968). Increase in oligomeric size is accompanied with enhanced inhibition of HEWL with both $(GlcNAc)_n$ and the MurNAc-containing compounds (Table 3.4) suggesting stronger binding for the larger saccharides. Although the *Chalarposis* lysozyme is not inhibited by chitosaccharides they are however inhibited by the non-N-acetylated analogues of chitin, the chitosan oligosaccharides (oligomers of β(1→4)-glucosamine) (Fouche & Hash, 1978). The *Streptomyces erythraeus* lysozyme is also inhibited by

Table 3.4. Inhibitory properties of lysozymes.

	Inhibitor		Specific Data for HEWL	
Lysozyme	$(GlcNAc\text{-}MurNAc)_n$	$(GlcNAc)_n$	Inhibitor	IC_{50} (mM)
HEWL	yes	yes	GlcNAc	7
GEWL	n.a.†	very weak‡	$(GlcNAc)_2$	0.6
T4L	no	no	$(GlcNAc)_4$	0.15
Chalarposis	n.a.†	no	(GlcNAc-MurNAc)	12
			$(GlcNAc\text{-}MurNAc)_2$	3

References: HEWL (Sharon, 1967; Jollès et al., 1968); GEWL (Jollès et al., 1968; Arnheim, 1974); T4L (Mirelman et al., 1975); *Chalarposis* (Hash, 1974; Fouche & Hash, 1978).
‡ measurable inhibition by $(GlcNAc)_4$ not detected until 4.5 mM. † data not available.

chitosan saccharides (Morita et al., 1978).

It is quite interesting that T4L and GEWL are not inhibited by chitosaccharides, or weakly in the case of GEWL, considering that these enzymes have crystal structures showing $(GlcNAc)_3$ bound in their active sites when crystallized in the presence of this trisaccharide (Anderson et al., 1981; Weaver et al., 1995). In addition, T4L has been purified by affinity chromatography over chitin and the enzyme has been shown to bind strongly to chitin (Jensen & Kleppe, 1972). Competitive inhibition essentially describes competition between the substrate and the inhibitor for the enzyme. It appears that although chitosaccharides bind to T4L and GEWL, they do so with an affinity that is much less than that for the substrate (i.e. peptidoglycan) and therefore, the bacteriolytic activity is not inhibited.

The inhibition of T4L is dramatically altered if the saccharide inhibitors contain peptide-substituted muramic acid residues (Table 3.5). Whereas saccharides of the form $(GlcNAc)_n$ containing only sugar residues do not inhibit T4L (for n = 2, 3, 4, and 6; Mirelman et al., 1975), a single muramic acid residue substituted with peptide can act as an effective inhibitor (see compounds **1** and **2**, Table 3.5). The position in the peptide occupied by the diamino acid appears to be well tolerated, as those inhibitors containing either *L*-Lys, *m*s-DAP or ε-N-acetyl-*L*-Lys demonstrate comparable inhibition (compare compounds **1**, **2**, **4** and **5**). However, substitution of Gly at the α-carboxyl of the *D*-Glu in compounds **6** and **7** appears to abolish the inhibitory capacity of these muropeptides. In addition, composition of the *C*-termini of the peptides with respect to the number of *D*-Ala residues (i.e. 0, 1 or 2) does not strongly influence their inhibitory properties (see compounds **1**, **2**, **3** and **4**). The peptide-crosslinked muropeptide **8** demonstrated the strongest inhibition (Table 3.5).

3.1.4. Supporting Evidence for Additional Lysozyme/Substrate Interactions

There is currently a paucity of structural information implicating interactions between lysozymes and portions of the peptidoglycan other than the carbohydrate residues. The inhibition of T4L by various muropeptides (Table 3.5) provides indirect evidence of the importance of enzyme interactions with the peptide-substituents. However, more recent structural data has firmly established the direct participation of peptide binding to T4L. Mutation of Thr26 to Glu in the active site of T4L produced an enzyme that cleaved *E. coli* peptidoglycan yet the product remained covalently bound to the mutated residue (Kuroki et al., 1993). The structure of the adduct was determined which revealed that the side chain of Glu26 was covalently attached to C1 of the muramic

Table 3.5. Inhibition of T4L by various muropeptides. Data from Jensen et al. (1976).

Compound	No.	K_i (mM)†
MurNAc		
R		
R = -*L*-Ala-*D*-*iso*-Glu-*ms*-DAP	1	0.31
-*L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala-*D*-Ala	2	0.10
GlcNAc-MurNAc		
R		
R = -*L*-Ala-*D*-*iso*-Glu-*ms*-DAP	3	0.12
-*L*-Ala-*D*-*iso*-Glu-*ms*-DAP-*D*-Ala	4	0.22
-*L*-Ala-*D*-*iso*-Glu-*ms*-DAP-*D*-Ala (N-Ac)	5	0.20
-*L*-Ala-*D*-*iso*-Glu(Gly)-*L*-Lys-*D*-Ala	6	N.I.
-*L*-Ala-*D*-*iso*-Glu(Gly)-*L*-Lys(N-Ac)-*D*-Ala	7	N.I.
-*L*-Ala-*D*-*iso*-Glu-*ms*-DAP(*L*-Ala-*D*-*iso*-Glu-*ms*-DAP-*D*-Ala)-*D*-Ala	8	0.05
MurNAc-GlcNAc		

† Measured using the soluble peptidoglycan substrate **I.B** (refer to Fig. 3.1).

residue of the muropeptide GlcNAc-MurNAc-*L*-Ala-*D*-*iso*-Glu-*ms*-DAP-*D*-Ala (Fig. 3.4).

The structure of the adduct was not only informative in demonstrating that the site D MurNAc residue was distorted (Fig 1.26 B, p. 74) but also provided the first indications as to the specific interactions with the peptide portion of the peptidoglycan. The peptide was found to extend across a surface groove in the C-terminal domain of the enzyme formed between two α helices (residues 108-113 and 126-124, Fig 3.4 A, B) which confirmed previous model building studies (Grütter & Matthews, 1982). Hydrogen bonding interactions of the peptide with both main chain and side chain atoms of the enzyme are extensive (Fig. 3.4 C). In reference to our work centred on the incorporation of trifluoromethionine into LaL (Chapter 4), note the presence of M106 in the peptide binding region of T4L (Fig. 3.4 B).

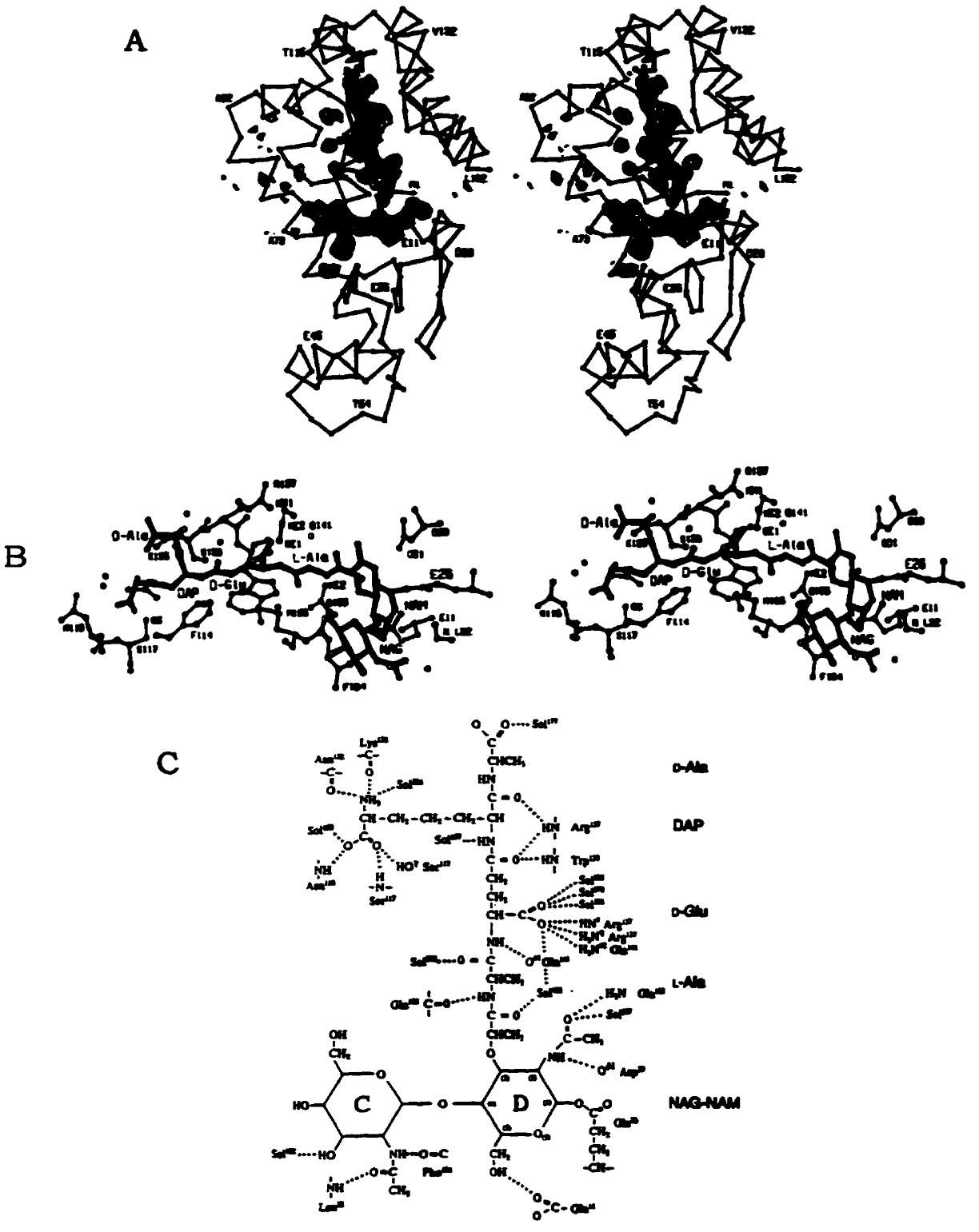


Figure 3.4. Structure of the mutant T4L/peptidoglycan complex. Reproduced from Kuroki et al. (1993). The disaccharide-tetrapeptide is covalently linked to Glu26 of the enzyme.
(A) Stereoview of the adduct and the protein backbone. The peptide extends upwards from the centre of the molecule.
(B) Stereoview including enzyme side chains; rotated 90° as shown in A.
(C) Schematic of the possible hydrogen-bonding interactions in the complex.

There is further evidence implicating the importance of other cell wall components to the activity of phage lysozymes. The lysozyme CPL-1 (*ch* type) is produced from bacteriophage Cp-1 which infects *Streptococcus pneumoniae*. The cell wall of this Gram-positive bacteria contains teichoic acids, of which choline is a structural component (García et al., 1988a). All the known endogenous lytic enzymes (i.e. peptidoglycan hydrolases) of *S. pneumoniae* can only hydrolyze cell walls that contain choline in their teichoic acids (Sanz & García, 1990). There is strong evidence that the CPL-1 lysozyme has specific interactions with choline. The activity of CPL-1 lysozyme depends on the presence of choline in the cell wall (García et al., 1988b; Romera et al., 1990) and is inhibited by choline (Sanz & García, 1990). In an elegant study, the two domains comprising the CPL-1 lysozyme were isolated and it was shown that the thermal stability of the C-terminal domain (and not the N-terminal domain) was increased in the presence of choline (Sanz et al., 1993). It appears that the C-terminal domain is directly involved in the binding of this lysozyme to the choline component of the cell wall of its bacterial host and this binding is required for the activity of this lysozyme.

The *ch* type lysozyme from the related pneumococcal phage Cp-7 contains an N-terminal domain similar to that of CPL-1 lysozyme but a completely different C-terminal domain and the Cp-7 lysozyme shows no dependency on choline (García et al., 1990). It was proposed that the pneumoccal phage lysozymes have arisen by the fusion of two genes; one coding for the N-terminal domain and one for the C-terminal domain (Sanz & García, 1990; Diaz et al., 1991; Sanz et al., 1993). The N-terminal domain confers the catalytic activity and the C-terminal domain is responsible for choline binding. Furthermore, the C-terminal domains of CPL-1 lysozyme and the LYTA amidase (the major lysin of *S. pneumoniae*, i.e. the bacterial host) show homology. Therefore, it appears that the gene for the C-terminal domain of the CPL-1 lysozyme could have possibly been acquired from its bacterial host. This type of genetic organization could perhaps be a common theme for phage lysozymes. As was discussed above, the C-terminal domain of T4L is involved in peptide binding. Therefore, T4L may have evolved from the fusion of an ancestral gene coding for catalytic activity (also located in the N-terminal domain of T4L) and a gene acquired from *E. coli* that provided the specificity for the *E. coli* peptidoglycan.

Even in the case of HEWL, whose substrate specificities are much less discriminating than those of phage lysozymes, is there supporting evidence that HEWL can also bind the peptide portion of the peptidoglycan. It was reported that HEWL binds $(GlcNAc\text{-}MurNAc)_2$ more strongly when it is substituted with peptide (Chipman et al., 1967). The activity of HEWL against a peptidoglycan substrate in which all the peptide

substituents were enzymatically removed was substantially less than the "intact" peptidoglycan (Strominger & Tipper, 1974). The substrate was polydispersed but with an average length of approximately 10 (GlcNAc-MurNAc) disaccharide units. It was suggested that a peptide-substituted peptidoglycan may be more apt to present or is better suited to fit into the cleft of HEWL (Strominger & Tipper, 1974). Therefore, there may be an assumed role of the peptide in directing the large molecular-weight-peptidoglycan substrate to the active site.

Solution studies on the binding of the dyes bromophenol blue and red appear to support this conclusion. It was shown that these dyes inhibited the activity of HEWL towards isolated cell walls but not against solely polysaccharide substrates (Krishnamoorthy & Prabhananda, 1982a,b). Subsequent crystallographic investigations have revealed that bromophenol red bound just outside subsite F while bromophenol blue bound close to subsites A and B (Madhusudan & Vijayan, 1992). With reference to the Phillips model, both sites B and F would be occupied by muramic acid and therefore, the attached peptide would lie just outside of these sites. This would account for the ability of the dyes to inhibit peptidoglycan hydrolysis. It was suggested that the peptides attached to the MurNAc residues interact with HEWL and provide interactions that are necessary for appropriately positioning the enzyme onto the peptidoglycan substrate (Madhusudan & Vijayan, 1992).

3.1.5. Observations on Activity with use of Different Substrates

Some striking differences are observed when lysozyme activity is measured using different substrates. For example, chemical modification of Lys residues in HEWL results in an enzyme with substantial reduction in bacteriolytic activity but no inhibition on oligosaccharide substrates (Imoto et al., 1972). Recently, several mutants of human lysozyme were prepared to remove the disulfide between the cysteine residues at postion 77 and 95 (Yamada et al., 1994). The enzymatic activities of the mutants were measured both by the lytic activity against *M. lysodeikticus* and liberation of PNP from $(GlcNAc)_5$-PNP. In some of the mutants, the relative activities against $(GlcNAc)_5$-PNP (as compared to the activity of the *wt* enzyme with $(GlcNAc)_5$-PNP) increased while the relative lytic activities of the same mutants (when compared to the *wt* lytic activity) decreased. For example, the relative activity of one mutant using $(GlcNAc)_5$-PNP was 292% while the lytic activity was 52% to that of the *wt* activities for these substrates respectively. With another mutant, the situation was reversed (lytic activity increased 205% and activity with $(GlcNAc)_5$-PNP increased to only 155%) and in other mutants, the change in relative activity was

comparable for both substrates. These findings clearly suggest that there exists differences between the *wt* and and mutant lysozymes in their recognition of the two substrates or that an altered catalytic mechanism may operate depending on the substrate used.

Recently, a very timely report has appeared which accentuates the possible relationship between substrate structure and enzyme activity. In an investigation to probe the essentiality of the carboxylate of Asp52 in HEWL, Matsumura and Kirsch (1996) have provided strong evidence towards the involvement of natural substrate-assisted hydrolysis. A general schematic of this model is shown in Fig. 3.5 (A). The model proposes that some of the catalytic functionality in lysozymes that are lacking a catalytic carboxylate (i.e. such as Asp52 in HEWL) can be provided in substrates that possess a carboxylate which can be appropriately positioned to assume the role of the "missing" enzyme carboxylate. For example, a lysozyme lacking the catalytic carboxylate (type II enzyme) may have fair activity on a substrate that contains an appropriate carboxylate (type II substrate) but displays poor activity on the same substrate that lacks this group (type I substrate) (Fig. 3.5 A).

To probe the validity of this model, a mutant HEWL was prepared replacing the aspartate with Ala. The activity of the D52A mutant was compared to that of the *wt* enzyme on two broad substrate types. In addition, the activity of GEWL, which differs from the *wt* HEWL in that GEWL lacks a natural counterpart to Asp52, was also compared.

The relative activity of the enzymes lacking the asparate (i.e. D52A HEWL and GEWL) to the enzyme containing the asparate (i.e. *wt* HEWL) on glycol chitin were 2.4% and 0.4% for the mutant and goose enzyme respectively. However, the relative rate of hydrolysis of carboxymethyl chitin of D52A HEWL and GEWL were 5.9% and 1.2% to that of *wt* HEWL, an approximate 3-fold increase than for glycol chitin. Model building shows that when a 6-*O*-carboxymethylated GlcNAc residue is bound to the E site of HEWL, the carboxylate of the 6-*O*-carboxymethyl group can be rotated into the location that is normally occupied by Asp52 (Matsumura & Kirsch, 1996) (Fig. 3.5 B). Since it is expected that the carboxylate is the only difference in glycol chitin and carboxymethyl chitin, the substrate assistance model may reasonably explain the approximate 3-fold increase in activity of the latter substrate with D52A HEWL and GEWL.

The wild type HEWL catalyses the clearing of *M. lysodeikticus* at a rate that maintains linearity with respect to time (Matsumura & Kirsch, 1996). Previously, the

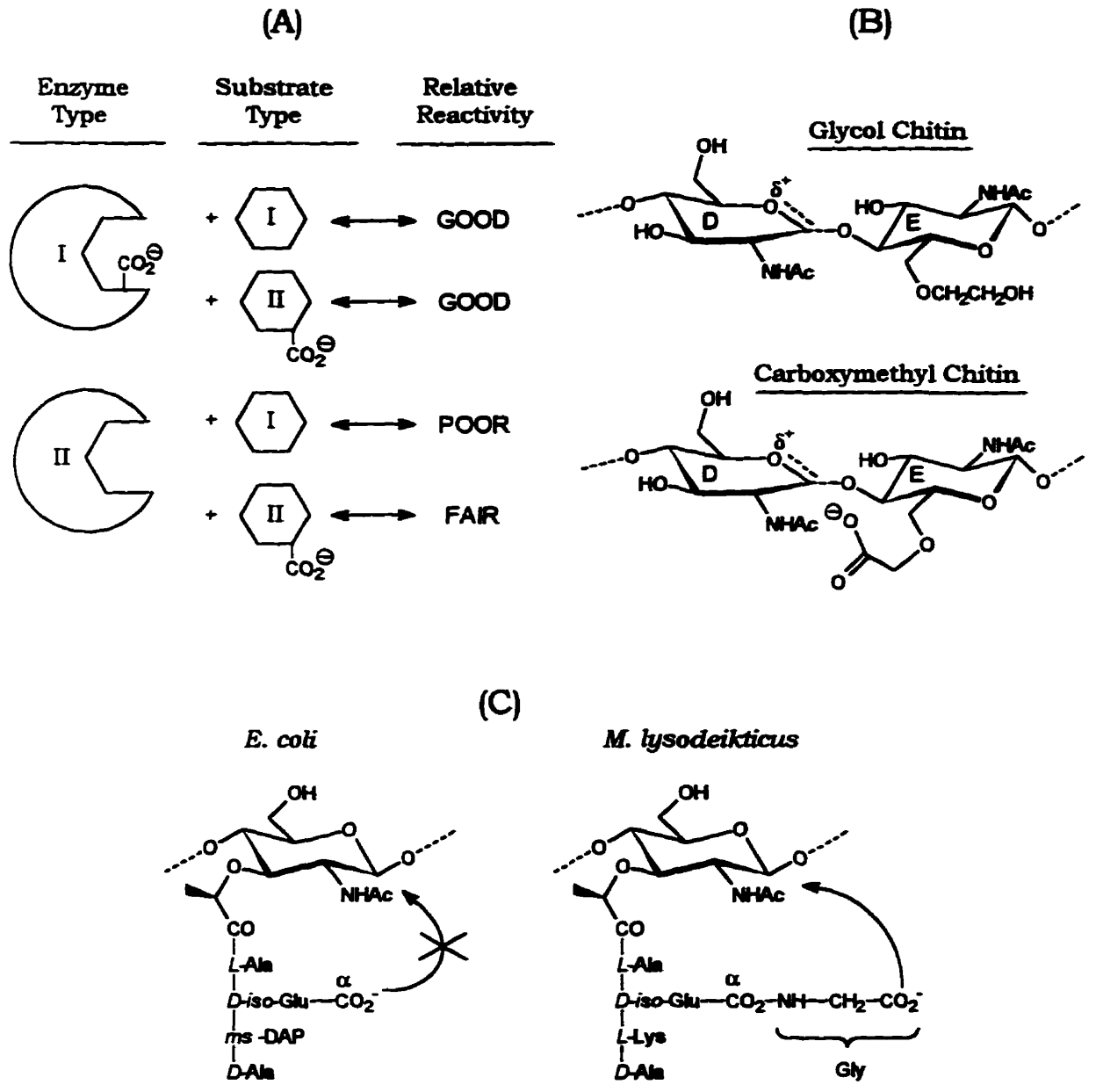


Figure 3.5. Models describing substrate-assisted catalysis. Adapted from Matsumura & Kirsch (1996).

(A) General model comparing relative activities between lysozymes containing (type I) or lacking (type II) a catalytic carboxylate and substrates lacking (type I) or containing (type II) an assisting carboxylate.

(B) Possible transition states for the cleavage of glycol chitin and carboxymethyl chitin. The GlcNAc residues are assumed to occupy the D and E subsites as indicated.

(C) Comparison of the primary structures of *E. coli* and *M. lysodeikticus* substituted muramic acids indicating that only the latter can participate in assisted catalysis by involvement of the Gly residue.

mutant D52N HEWL was shown to exhibit biphasic kinetics on *M. lysodeikticus* in which an intial fast phase occurred at 5% the relative rate of *wt* HEWL that was followed by a slow phase of clearing at 0.5% of the *wt* rate (Malcolm et al., 1989). Similar kinetics have been observed for the D53E human lysoyzme (Muraki et al., 1991). It was suggested that the 5% initial activity shown by D52N HEWL resulted from the cleavage of a small class of "hyperlabile linkages" in the peptidoglycan. The rate decays to the smaller residual rate of 0.5% once these more susceptible linkages have been exhausted (Malcolm et al., 1989).

The D52A HEWL and GEWL also exhibited biphasic kinetics. The fast phase and slow phase velocities for D52A HEWL are equal to 22.1 and 4.1%, respectively, of the *wt* HEWL rate. For GEWL, the fast and slow phases are equal to 204 and 142% of the *wt* HEWL rate[3.2] . However, for both the D52A HEWL and GEWL, the fraction of clearing that occurred during the fast phase was 2.7% of the total clearing by the respective enzymes. Therefore, it was suggested that the fast phase of the reaction corresponded to cleavage of a particular subset of peptide-substituted muramic acids in the peptidoglycan of *M. lysodeikticus* having a carboxylate group that can assist in catalysis (Matsumura & Kirsch, 1996).

In *M. lysodeikticus*, approximately 5% of the MurNAc residues are substituted with the free pentapeptide shown in Fig. 3.5 C. It was possible to model the peptide (with planar peptide bonds) such that the carboxylate of the Gly could be positioned near the C1 glycosidic bond (Matsumura & Kirsch, 1996). In this fashion, the Gly carboxylate can assist in catalysis in the D52A HEWL and GEWL which lack the enzyme aspartate (Fig. 3.5 C) as in the case for the carboxymethyl chitin discussed earlier. The other peptide substituted MurNAc residues in *M. lysodeikticus* are involved in crosslinking and have restricted conformational freedom to correctly position the Gly carboxylate (Matsumura & Kirsch, 1996). It was not possible to model the *E. coli* peptidoglycan, which lacks the Gly group, such that a carboxylate could assume a catalytic position (Fig. 3.5 C). In accord with the lack of substrate assistance and the absence of hyperlabile linkages, the clearing of *E. coli* cells by D52A HEWL and GEWL did not exhibit biphasic kinetics.

The results of the work presented above strongly indicate that an enzyme carboxylate is not *essential* for catalysis; however, the presence of such a carboxylate, either provided by the enzyme or the substrate, will apparently increase the catalytic efficiency. Equally noteworthy is the relationship that has been established between variations in the substrate structures and enzyme activity.

[3.2] The superior activity of GEWL than HEWL contradicts the claim made by Schindler et al. (1977b) that HEWL is more effective than GEWL on *M. lysodeikticus*.

3.1.6. Concluding Statements

Although lysozymes are one of the most intensively studied enzymes, the review presented in this introduction has brought attention to the fact that measuring lysozyme activity is as intricate as the substrates used. The most widely adopted assay is that involving the measurement of the lytic activity on bacterial substrates. In the case of Gram-negative phage lysozymes, sensitization of the bacterial substrates is required to render them susceptible to the action of lysozyme. The choice, reproducibility and effectiveness of sensitization methodology are likely to introduce additional factors that could influence the observed activity of phage lysozymes on such a bacterial substrate. It is alarming that more elaborate descriptions on the preparation and evaluation of Gram-negative bacterial substrates have not appeared in the literature.

It has also been made apparent that both gross and subtle differences in the composition of ligands and substrates can affect their binding to and turnover by lysozymes. These general features appear to have arisen by selective pressure on the lysozyme to evolve with specific ligand and substrate specificities. Comparisons of ligand and substrate specificities can be informative to suggest evolutionary relatedness between lysozymes.

The lack of uniformity in most lysozyme substrates, most notably the peptidoglycan derived substrates, has provided strong incentive for the preparation of lysozyme substrates that are chemically defined and possess a limited number of bonds that can be acted upon. Although some success has been achieved for the *c* type lysozymes and the substrates derived from chitosaccharides, the apparent peptide-binding specificity of the *g* and *p* type lysozymes has limited the available substrates for these lsyzoymes. Well defined, soluble substrates containing peptide-substituted MurNAc residues will be indispensable for research on these lysozymes and the specific need for such substrates has been expressed (Mirelman et al., 1975; Fouche et al., 1978; Sanz & García, 1990; Hardy & Poteete, 1991; Kuroki et al., 1993).

3.2. EXPERIMENTAL

3.2.1. Materials

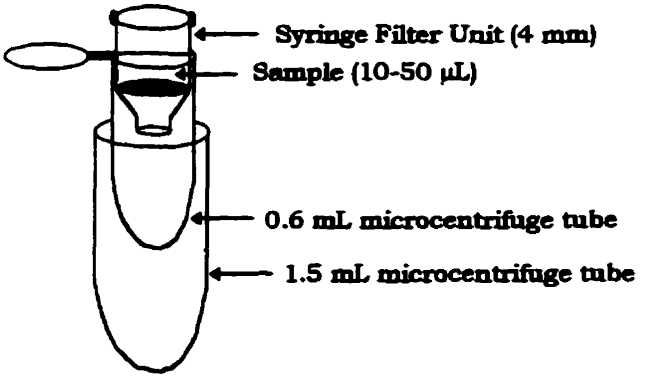
All materials used were of the highest grade commercially availiable and were used without further purification except when noted otherwise. Reagents purchased from Aldrich (Milwaukee, WI) included acetyl chloride (AcCl), N-acetyl-*D,L*-homoserine thiolactone, anisole, benzotriazole-1-yloxytris(dimethylamino)phosphonium hexafluorate (BOP), 4-dimethylaminopyridine (DMAP), dimethyl sulfide, 1-hydroxybenzotriazole monohydrate (HOBt·H_2O), ^{13}C-methyl iodide, *L*-[*methyl*-^{13}C]methionine (99% ^{13}C), 4-methylmorpholine, phosphomolybdic acid solution (20% in ethanol), sodium hydride (80% dispersion in oil), trifluoroacetic acid (reagent grade and protein sequencing grade), trifluoromethanesulfonic acid (TFMSA) and trimethylsilyl trifluoromethanesulfonate ($(CH_3)_3SiOSO_2CF_3$). Materials obtained from Sigma (St. Louis, MO) included *N*-acetyl-*D*-glucosamine (GlcNAc), the adjuvant peptide (*N*-acetylmuramyl-*L*-alanyl-*D*-isoglutamine), *O*-benzotriazolyl-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU), N-t-BOC-*D*-alanine (BOC-*D*-Ala-OH), N-t-BOC-*L*-alanine (BOC-*L*-Ala-OH), N-t-BOC-*D*-glutamic acid α-benzyl ester (BOC-*D*-Glu(αOBzl)-OH), D-(+)-cellotetraose, α-cyano-4-hydroxycinnamic acid, Gramicidin S, Kemptide, maltotetraose, freeze-dried *Micrococcus lysodeikticus*, *p*-nitrophenyl-2-acetamido-2-deoxy-β-*D*-glucopyranose (β-PNP-GlcNAc), *p*-nitrophenyl-β-*D*-*N,N'*-diacetylchitobioside, *p*-nitrophenyl-β-*D*-*N,N',N''*-triacetylchitotrioside, *p*-nitrophenyl-β-*D*-glucose, *p*-nitrophenyl-α-*D*-glucose, *p*-nitrophenyl-β-*D*-galactose and *o*-nitrophenyl-β-*D*-galactose. Acetonitrile (HPLC grade), benzaldehyde (AnalaR), glacial acetic acid (HOAc), potassium cyanide (KCN), pyridine, sodium metal, sulfuric acid and zinc chloride ($ZnCl_2$, AnalaR) are products of BDH (Toronto, ON) and were of Assurance grade except when noted. Acetic anhydride (Ac_2O) and anhydrous magnesium sulfate ($MgSO_4$) were obtained from Fisher Scientific (Don Mills, ON) and vanillin is a product of J. T. Baker Chemical Co. (Phillipsburg, NJ). NovaBiochem (La Jolla, CA) was the source for N-*t*-BOC-*L*-lysine ε-9-fluorenylmethoxycarbonyl (BOC-*L*-Lys(Fmoc)-OH). Piperidine and cesium carbonate (Cs_2CO_3) were obtained from Fluka (Ronkonkoma, NY) while 4-methylumbelliferyl-N-acetyl-chitotriose ($(GlcNAc)_3$-4MU) is a product of Calbiochem (La Jolla, CA). The chitooligosaccharides $(GlcNAc)_n$, (n=2-6, Fine grades), *p*-nitrophenyl-penta-N-acetyl-β-*D*-chitopentaoside ($(GlcNAc)_5$-PNP) and hen egg white lysozyme were purchased from Siekagaku America Inc. (Rockville, MD). *N-N'*-Diisopropylcarbodiimide is a product of the Chemical Dynamics Corporation (Plainfield, NJ) and (S)-(-)-2-bromo-propanoic acid was obtained from Janssen Chimica (Geel, Belgium). *L*-Alanine methyl ester hydrochloride was a kind gift from Dr. G. Lajoie.

3.2.2. Analytical Techniques

i) Optical spectroscopy was performed on a Varian Cary 3 UV-Vis Spectrometer and operated utilizing the Cary 13 software (Varian Associates, 1990 Release1). Chromatograms obtained from the Cary 3 UV-Vis Spectrometer were presented in this thesis by converting the Cary data files into ASCII format (packed data, comma delimited) using the Cary13 software and then imported into the Microcal Origin v2.90 software.

ii) High-Performance Liquid Chromatography (HPLC) was performed on a Waters 625 LC unit equipped with a Waters 994 photodiode detector and a 5200 Printer Plotter. The detector was typically set to monitor a wavelength range of 190-400 nm when chromatography of *p*-nitrophenyl-containing glycosides was performed so that appropriate spectra of the individual peaks could be analysed.

When reference is made to centrifugation of samples prior to HPLC, this was performed with a BioFuge A table top microcentrifuge. Samples were centrifuged at 13000 rpm for 5 min. Small volume (< 50 μL) filtering of samples was performed with the apparatus outlined below:



The desired volume of sample requiring filtration was pipetted into a syringe filter unit which was placed into a 0.6 mL microcentrifuge tube which was then subsequently placed into a 1.5 mL microcentrifuge tube. The entire unit was then centrifuged using the BioFuge A table top centrifuge at 2000 rpm for 5-10 min and the filtered sample recovered from the 0.6 mL microcentrifuge tube (NOTE: centrifugation at speeds >5000 rpm will cause breakage of the filtration membrane). This filtration methodology was found to be very practical and efficient with 100% recovery of sample.

Reverse phase C18 chromatography was performed using a Delta-Pak™ (300 Å, 5 μM) 2 mm × 150 mm column or μBondapak™ (125 Å, 10 μM) Prep-Pak™ cartridge (25 mm × 10 cm, 50 mL) or Radial-Pak™ cartridge (8 mm × 10 cm, 5 mL), each a product of Waters.

In most cases, a μBondapak™ Guard-Pak™ cartridge (25 mm × 10 mm, 5 mL) was used in conjunction with the Prep-Pak™ cartridge. The precolumn consisted of a Guard-Pak holder using either μBondapak™ C18 or Nova-Pak® C18 precolumn inserts (Waters). Analytical separation conditions for purification of compounds were first developed using the Radial-Pak™ cartridge and then scaled up and further optimized using the Prep-Pak™ cartridge for preparative scale purification.

Gel permeation HPLC using the Showdex® OHpak Q-801 (8 mm × 50 cm) column (Waters) at 55 °C was achieved by heating of the column with a column heater kindly provided by Dr. T. Viswanatha.

All solvents used for HPLC were filtered (0.22 or 0.45 μM) and degassed prior to use. When included as an additive, TFA (sequencing or HPLC grade) was added to the desired concentration to the solvents following their degassing.

iii) Capillary Zone Electrophoresis (CZE or CE) was performed with a BioFocus® 3000 Automated Capillary Electrophoresis System (BioRad) operated with the BioFocus® 3000 software (v3.00 and v3.10). Data were analyzed using the BioFocus® Spectra software. In some cases, the BioFocus BFF (braided-format datafile) datafiles were converted into ASCII format (at a 10 Hz conversion rate) using the conversion option provided with the software and then imported into the Microcal Origin v2.90 software for presentation.

All materials used for CZE are products of BioRad. BioFocus® Capillary cartridges used included the 24 cm × 25 μM I.D. coated (cat# 148-3031) and the 50 cm × 50 μM coated (cat# 148-3033) cartridges. Low pH CZE was performed using the Phosphate Buffer (0.1 M, pH 2.5 with polymer modifier) as the running buffer while high pH CZE was performed using the Basic Protein Analysis Buffer (0.3 M sodium borate, pH 8.5 with polymer modifier) as the running buffer. Unless otherwise stated, samples were prepared in running buffer that had been diluted with MQW to one-tenth of the running buffer concentration (0.01 M for the phosphate and 0.03 M for the borate buffer respectively). All buffers and samples were filtered (0.22 μM or 0.45 μM) prior to use and samples and buffers were maintained in the carousel compartment at 10 °C. Samples were introduced into the cartridges by pressure injection at 5 psi. For example, a 20 psi × sec injection therefore corresponds to a 5 psi × 4 sec injection. All electrophoresis was performed at constant voltage. Further details concerning individual CZE runs (voltage and polarity, cartridge temperature, pressure injection) are given in the respective figure legends. Cartridges were washed between runs using a cycle of MQW, Capillary Wash Solution, MQW and then the running buffer. The performance of the cartridges was routinely

monitored using the Peptide Calibrator (cat# 148-2012) as described in the Peptide Analysis Kit instruction manual.

iv) Nuclear Magnetic Resonance Spectroscopy (NMR) was performed on either a Bruker AC200 (^{1}H, 200 MHz; ^{13}C 50.3 MHz), AM250 (^{1}H, 250 MHz; ^{13}C 62.9 MHz) or AMX500 (^{1}H, 500 MHz; ^{13}C, 125.7 MHz) instrument and processed using standard Bruker software. Chemical shifts (δ) are reported in parts per million (ppm) and were referenced to the deuterated solvent used to obtain the spectra unless otherwise stated and are: acetone-d_6, ^{1}H 2.04 (5), ^{13}C 29.8 (7); acetonitrile-d_3, ^{1}H 1.93 (5), ^{13}C 1.3 (7); $CDCl_3$, ^{1}H 7.24, ^{13}C, 77.0 (3); D_2O, ^{1}H 4.63; DMSO-d_6, ^{1}H 2.49 (5), ^{13}C 39.5 (7); methanol-d_4, ^{1}H 3.30 (5), ^{13}C, 49.0 (7). The number in parenthesis indicates the multiplicity of the solvent peak.

v) Mass Spectrometry. Matrix Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI MS) was performed on a VG Analytical time-of-flight (TOF) mass spectrometer (Fisons Instruments) and spectra were analyzed using the OPUS software. All samples for MALDI MS were prepared by mixing equal volumes of the analyte solution with a matrix solution consisting of MQW: CH_3CN (2:1, containing 0.2% TFA) saturated with α-cyano-4-hydroxycinnamic acid. Samples were loaded onto the targets on the sample stage, allowed to crystallize by drying, and then mounted into the mass spectrometer and the mass spectra were obtained by summing the data from pulses (typically 10-30) of a 337 nm nitrogen laser. Spectra were acquired using positive ionization and the coarse and fine laser energy levels were adjusted to optimize the resolution and were typically set at 2 (energy coarse) and 120-150 (energy fine). Matrix suppression was used to attenuate the matrix signals. Gramicidin S (1141.7 Da) and Kemptide (771.9 Da) were used as external calibrants and their spectra were obtained under similar acquisition parameters as those used for the sample of interest.

Electrospray mass spectrometry (ESMS) was described earlier (section 2.2.11). For the non-protein small molecular weight samples, the operating parameters for data aquisition were essentially the same as for protein samples, with the exception that i) the range was typically set to ± 200-300 Da to the mass of the sample, ii) the cone voltage was lowered to between 10 and 25 volts and iii) the LM resolution was increased to between 18 and 21. The data were processed to give a centred spectrum and the auto-component function was employed to obtain the molecular weights.

Protein purification techniques were described earlier in section 2.2.9.1.

3.2.3. Assays Involving Substrate Cells (Turbidimetric Assay)

3.2.3.1. Preparation of Bacterial Substrate Cells

The activity of LaL was assayed turbidimetrically using *E. coli* cells that were sensitized with $CHCl_3$ and EDTA. Because inconsistencies were observed when the substrate cells were prepared as previously described (Black & Hogness, 1969a), modifications to this original methodology were deemed necessary. The procedure outlined below is a culmination of many attempts to establish conditions that would result with the substrate cells having certain desirable properties.

An overnight culture (25 mL) of *E. coli* MG1655 (prototrophic) grown in M9 minimal medium (Appendix A) supplemented with glucose (0.3%) was subcultured into 1 L of the same medium (in a 4 L flask). The culture was incubated at 37 °C with shaking (180 rpm) and grown to an absorbance of 0.8 at 600 nm (typically requiring 6-8 hr). The cells were collected by centrifugation (9800 × g, 10-12 min, 2 °C) using Nalgene™ HDPE centrifuge bottles which are more resistant to organic solvents. Following centrifugation, the supernatants were decanted and the cell pellets were loosened with a rubber policeman.

Prior to centrifugation, $CHCl_3$ saturated substrate cell buffer (60 mM KPB, pH 7.0, 5 mM EDTA, Appendix A) was prepared (it is also necessary to cool some of the substrate cell buffer to 4 °C for later use). Buffer saturation was achieved by vigorously stirring a 20% (v/v) mixture of $CHCl_3$ and substrate cell buffer for 30 min at 25 °C, allowing the $CHCl_3$ to settle for 5 min and then decanting the buffer (top) layer into a sealed container. Gradually increasing aliquots (starting at 2-4 mL) of the saturated buffer was added to each of the loosened cell pellets (at 25 °C) while gently resuspending the cells with a rubber policeman. Attempts were made to achieve uniform suspension of the cells (i.e. no clumps) in the minimal amount of time. The resuspended cells were pooled into a single centrifuge bottle after a total volume of 50-75 mL of saturated buffer was added and the centrifuge bottle was sealed with its screw top lid.

At this stage, the response of the cells was determined by diluting an aliquot of the suspension (30-50 µL) with 50 mM KPB, pH 7.0 (700 µL) to an $OD_{600} \approx 0.7$. An aliquot of LaL (5-10 µg) was added which had been determined from previous experiments, to cause almost immediate clearing of a successfully prepared sample of substrate cells. If it appeared that the suspension was giving indications that the cells were responding to the treatment, no further $CHCl_3$ was added. If the suspension was not clearing as expected after 5-10 min, an additional amount of $CHCl_3$ (obtained from the original mixture used to saturate the substrate cell buffer) was added (2-3 mL). During the entire $CHCl_3$ treatment, the cells (with the lid on the centrifuge bottle) were gently and continuously shaken by

hand at 25 °C. After the additional $CHCl_3$ was added, the response of the cells was continually monitored. Treatment was considered sufficient when the diluted aliquot (with an initial absorbance ≈ 0.7) would clear and level to an absorbance of ≈ 0.2 after 1-2 min following the addition of LaL (5-10 μg). In the case of successfully prepared substrate cells, the treatment was usually extended for 8-15 min (with continuous shaking by hand) following the addition of the extra $CHCl_3$.

The next steps were performed as quickly as possible after the treatment above was deemed complete. The cell suspension was placed into an NaCl ice bath during which time as much of the $CHCl_3$ which had settled to the bottom of the centrifuge bottle was removed with a pasteur pipette. The cells were pelleted (9800 × g, 10 min, 2 °C) and the supernatant was decanted. Substrate cell buffer (without $CHCl_3$ saturation, at 4 °C) was added 2 or 3 times and simply washed over the pellet without disturbing the pellet. The cell pellet was then loosened and gently resuspended with the gradual addition of substrate cell buffer (lacking $CHCl_3$, 4 °C) such that a 1:20 dilution of this suspension resulted in an absorbance of 0.7-0.75 at 600 nm (approx. 25-40 mL is required, but varies with each preparation). Meticulous care must again be taken to insure that uniform suspension of the cells is achieved. The suspension was transferred to a conical bottom flask. While the flask was continually swirled on ice, aliquots (≈ 0.5 mL, using a pasteur pipette) of the suspension were dispensed into 1.5 mL eppendorf tubes and frozen in liquid N_2. The conical flask was used since there was always some "clumps" of material present that would gather at the bottom of such a flask and could therefore be more readily avoided upon removal of the aliquots. The frozen cells were then stored at -80 °C.

The following day, several aliquots of the substrate cells were assessed for their response characteristics (3.2.3.2). If the cells behaved acceptably as required for kinetic studies, they were reserved for this purpose. Often, the substrate cells were not useful for kinetic studies but were of adequate quality for simple activity measurements during purification of LaL.

3.2.3.2. Characterization of the Response of Substrate Cell Preparations to LaL Concentration

Successfully prepared substrate cells were characterized to determine the response they displayed to various concentrations of LaL. Because different preparations of substrate cells displayed different propensities to clearing by LaL, the characterization was somewhat unique for each.

A solution of LaL (typically from 2-10 ng/μL) was prepared in stabilizing buffer (0.5 M KCl, 100 mM KPB, pH 7.0, 20% glycerol; Appendix A). As was discussed previously

(section 2.3.5.1) this buffer system was necessary to maintain the activity of the enzyme at these dilute concentrations. A frozen aliquot of the substrate cells was allowed to thaw on ice and then constantly maintained on ice. Occasionally, the frozen aliquot was periodically warmed between hands to accelerate thawing, but this was kept to a minimum. An aliquot of the cells (40-45 μL) was diluted with 50 mM KPB, pH 7.0 to an absorbance between 0.72 and 0.75 at 600 nm and a total volume of 750 μL in a 1.0 mL cuvette. An appropriate amount of the LaL solution (1-5 μL) was added, the cuvette was inverted 5 times, and the decrease in turbidity of the suspension kept at 25 °C was followed by measuring the absorbance at 600 nm (usually up to 3 or 4 min). After an initial lag period (30-60 s) the turbidity of the solution decreased linearly with time over the first several minutes. Addition of too much LaL (see below) resulted in a non-linear response. The slope of the curve between 2 and 3 min was found using the slope command function available with the Cary13 (Ver. 2.0) software.

Initially, the amount of LaL which gave a decrease of ≈ 0.07-0.08/min between 2 and 3 min needed to be determined. This was taken as the maximal amount of LaL to be used with the particular substrate cell preparation, as addition of LaL in excess of this amount would give a non-linear response. The concentration of LaL (prepared in stabilizing buffer) was then adjusted so that a 5 μL aliquot of the solution would deliver this amount of LaL. For example, in one case, 20 ng of LaL gave a decrease of 0.075 au/min. Therefore, the LaL solution was diluted (in stabilizing buffer) to 4 ng/μL such that 5 μL of this solution represented 20 ng of protein. Then, during characterization of these particular cells, aliquots of the 4 ng/μL solution (1-5 μL, 4-20 ng) were added and the response determined as described above. If 50 ng of LaL was required for maximal response, aliquots (1-5 μL) of a 10 ng/μL solution of LaL (in stabilizing buffer) would be used to characterize the response of these substrate cells. As a blank, 2.5 μL of stabilizing buffer would be used. The blank run would be performed as the first and the last measurement and the slopes for the two runs averaged.

The total assay time needed to cover a suitable range of LaL concentrations was kept as short as possible (data for each assay was stored on the computer and slope measurements were made after completion of assays). This was done to minimize any changes in response of the substrate cells that were observed to occur with time after thawing and maintaining on ice. Typically, 7 individual assays could be performed in 30 min (if each assay was monitored for 3 min) in which time, the substrate cells behaved uniformly and consistently maintaining their sensitivity.

3.2.3.3. Inhibition of Bacteriolytic Activity Assays

The inhibition by $(GlcNAc)_n$ (n = 1 - 6), maltotetraose, cellotetraose, peptides and the adjuvant peptide was determined as described in section 3.2.3.2 with the following additions. The test compound was prepared in 50 mM KPB, pH 7.0 such that mixing of this solution (700-710 μL) with an aliquot of substrate cells (40-50 μL) would give the desired concentration of the compound and an initial OD_{600} of the cells between 0.72 and 0.75. The effect of each test compound and concentration was determined independently using a freshly thawed aliquot of substrate cells. The amount of LaL used would be that amount which gave a response of at least 0.05 au/min or typically between 0.06-0.08 au/min decrease between 2 and 3 min for the substrate cells in the absence of any inhibitor. The LaL solution would be prepared in stabilizing buffer at a concentration (usually 2-10 ng/μL) such that 3-4 μL of the solution would deliver the response described. For each compound or test concentration, three alternating runs were performed in the absence or presence of inhibitor in the assay buffer. Again, the total assay time for each inhibitor was kept below 30 min to minimize any changes in response of the substrate cells. As alternating runs in absence or presence of compound was performed, this also tended to compensate for any time-dependent aging effect of the substrate cells. As a blank, the same volume of stabilizing buffer as enzyme aliquot was used. The blanks were performed as the first and last run.

Residual activity was calculated by comparison of the average slope between 2 and 3 min between runs with inhibitor to those without taking into account the enzyme blank runs. A sample calculation including propagation of errors is presented in Appendix C (section C.1).

3.2.3.4. Effect of Buffer and Ionic Strength on the Turbidimetric Assay

In each of the sytems described below, a freshly thawed aliquot of substrate cells from the same substrate cell preparation was used. A solution of LaL was prepared in stabilizing buffer at 4 ng/μL. In each case, the final buffer concentration is that obtained after dilution of the substrate cell aliquot (50 μL) with the different buffers (700 μL) to an initial OD_{600} between 0.72-0.75.

1) Potassium Phosphate. A 1.0 M stock solution of KPB, pH 7.0 was appropriately diluted with MQW such that the final concentration of KPB in the assay ranged between 25 and 100 mM.

2) Tris. A 0.5 M stock solution of Tris, pH 7.0 (which was prepared by mixing 0.5 M Tris HCl and 0.5 M Tris base to pH 7.0) was appropriately diluted with MQW such that the final concentration of Tris buffer in the assay ranged between 25 and 150 mM.
3) Potassium Chloride. A 1.0 M KCl, 53.6 mM KPB, pH 7.0 stock solution was appropriately diluted with 53.6 mM KPB, pH 7.0 such that the final concentration of KCl in the assay ranged between 0 and 20 mM in 50 mM KPB, pH 7.0.
4) Ammonium Sulfate. A 1.0 M $(NH_4)_2SO_4$, 53.6 mM KPB, pH 7.0 stock solution was appropriately diluted with 53.6 mM KPB, pH 7.0 such that the final concentration of $(NH_4)_2SO_4$ in the assay ranged between 0 and 10 mM in 50 mM KPB, pH 7.0.

In each assay, an aliquot of the enzyme solution (2.5 μL, 10 ng protein) was added, the cuvette was inverted 5 times, and the decrease in turbidity of the suspension kept at 25 °C was followed by measuring the absorbance at 600 nm for 4 min. The activity was obtained from the linear portion of the curves. For the assays performed in (1) and (2) above, the slope was taken between 0 and 1 min for 25 mM Tris and 25 mM KPB and between 1 and 2 min for 50 mM Tris and 40 mM KPB. In all other cases, the slope between 2 and 3 min was obtained.

3.2.4. Biophysical Studies on the Interactions of $(GlcNAc)_n$ with LaL

3.2.4.1. Fluorescence

Fluorescence spectra were recorded with a SLM Instruments Inc. spectrophotometer, model LM-480 interfaced with a computer fitted with software for spectra processing in the laboratory of Dr. J. Lepock (Dept. of Physics, University of Waterloo). All solutions were prepared in 50 mM KPB, pH 7.0 and measurements were performed at ambient temperature without temperature control. Fluorescence spectra were obtained for LaL at concentrations of 1.1-2.5 uM in the absence or presence of $(GlcNAc)_n$ (n = 2-6, fine grade) or maltotetraose (concentrations of enzyme and saccharides are indicated in Table 3.8). A stock solution of LaL (2 mg/mL in 50 mM KPB, pH 7.0, purified from pHDM10/TG-1, maintained on ice) was added to the desired saccharide solution (maintained at ambient temperature), allowed to equilibrate for 2 min and emission spectra (over 300-600 nm) were recorded with excitation at 280 nm. A spectra of the free enzyme was always obtained for each set of spectra obtained for the different saccharides. Appropriate blanks, containing only the same concentrations of saccharide in buffer but no enzyme, were also obtained and revealed that no emission was observed over 300-535 nm (only the buffer exhibited a small emission maxima at 565 nm with a

relative intensity of ≈ 0.05). Spectra were also recorded with HEWL at 2.5 μM both in the absence and presence of selected saccharides. Emission maxima and fluorescence intensities were measured directly from the plotted spectra following a 1st order smoothing of the raw data. From repeated (n = 5) measurements of the same sample of free LaL, errors in emission maxima were found to be less than ± 1 nm while relative intensities were found to have an associated error of approx. ± 0.02.

Measurements were also made as above with 2.2 μM $(GlcNAc)_3$-4MU in the presence or absence of varied concentrations of LaL or HEWL with an excitation wavelength of 330 nm.

3.2.4.2. Differential Scanning Calorimetry (DSC)

DSC was performed with a Microcal-2 MC-2 calorimeter (Microcal, Inc., Northhampton, MA) interfaced to a DEC Pro 380 computer in the laboratory of Dr. J. Lepock (Dept. of Physics, University of Waterloo). DSC scans of samples were obtained under a constant external pressure of 50 lbs/in² using 1.21 ml sample cells. Sample and reference (50 mM KPB, pH 7.0) solutions were degassed on ice under mild vacuum and gentle stirring for approximately 5 minutes immediately before addition to the calorimeter cells. Samples were allowed to equilibrate to 5 °C (approx. 10 min) and unless otherwise stated, were scanned from 5 to 105 °C at a scan rate of 1 °C/min, the sample was then cooled back *in situ* to 5 °C, and a rescan from 5 to 105 °C was immediately run to check for reversibility. Data were analyzed with the Microcal Origin v2.90 software and transition temperatures (T_m) were obtained following curve smoothing and baseline correction.

All calorimetric measurements were conducted with LaL purified from *E. coli* TG-1/pHDM10 that had been lyophilized following dialysis against 5 mM KPB, pH 7.0. Calorimetric measurements were carried out in 50 mM KPB, pH 7.0 and at a protein concentration of 3.5 mg/mL unless otherwise stated. Samples at the desired protein and buffer concentration were prepared by careful weighing of the protein/buffer powder that was dried *in vacuo* prior to weighing. Because the lyophilized powder contained buffer salts, they would contribute to the final buffer concentration. Therefore, the final buffer concentration was achieved by the addition of a calculated amount of 1.00 M KPB, pH 7.0 as described in Appendix C (section C.2).

The effect of $(GlcNAc)_5$, $(GlcNAc)_3$ and maltotetraose were examined by their inclusion with protein samples at a concentration of 5 mM.

3.2.4.3. NMR Studies of [methyl-^{13}C]Methionine Labelled LaL

3.2.4.3.1. Synthesis of [methyl-^{13}C]methionine

General synthetic techniques and definition of TLC_2 are described in 3.2.6.1.

N-Acetyl-*DL*-[*methyl*-^{13}C]Methionine Methyl Ester (1)

The procedure is based on general methodology previously developed (Houston, 1992). Distilled anhydrous methanol (20 mL) was cooled on ice and sodium metal (0.2 g, 8.7 mmol) was added in portions under argon. The suspension was stirred on ice until the sodium had dissolved. N-Acetyl-*DL*-homocysteine thiolactone (1.01 g, 6.34 mmol) was added at once and the solution was stirred for 40 min. To the solution was added ^{13}C-methyl iodide (1.00 g, 7.00 mmol) dropwise over 2 min. The reaction was allowed to warm to room temperature and stirred for an additional 2 hr. The solvent was then evaporated *in vacuo* and the residue was taken up in CH_2Cl_2 (100 mL) and washed successively with 2% $NaHCO_3$ (40 mL) and brine (40 mL). The organic layer was dried over $MgSO_4$, the solvent removed *in vacuo* and the product dried under vacuum to yield 0.981 g (4.76 mmol, 75%) of a white solid.

R_f (TLC_2 (see 3.2.6.1), EtOAc, phosphomolybdic acid visualization) 0.50; ^{1}H NMR (250 MHz, $CDCl_3$) δ 6.92 (d, 1H, $J_{NH,\alpha}$ = 7.7 Hz, **NH**Ac), 4.69 (dt, 1H, J = 5.1 Hz, 7.7 Hz, **Hα**), 3.75 (s, 3H, $CO_2C\mathbf{H_3}$), 2.53 (m, 2H, **Hγ** × 2), 2.21-1.93 (m, 2H, **Hβ** × 2), 2.09 (d, 3H, J = 138.4 Hz, $S^{13}C\mathbf{H_3}$), 2.04 (s, 3H, $NHCOC\mathbf{H_3}$); ^{13}C NMR (62.9 MHz, $CDCl_3$) δ 172.3, 170.0 ($NH\mathbf{C}OCH_3$, $\mathbf{C}O_2CH_3$), 52.1, 51.3 (**Cα**, $CO_2\mathbf{C}H_3$), 31.3, 29.8 (**Cβ**, **Cγ**), 22.6 ($NHCO\mathbf{C}H_3$), 15.1 ($S^{13}\mathbf{C}H_3$).

***DL*-[*methyl*-^{13}C]Methionine Hydrochloride (2)**

N-Acetyl-*DL*-[*methyl*-^{13}C]methionine methyl ester (0.85 g, 4.12 mmol) was added to 30 mL of degassed 2 N HCl and the resulting suspension further degassed on a high vaccuum pump for 5 min. The mixture was refluxed under argon for 15 hr. The solvent was evaporated *in vacuo* and the residue was dissolved in water (20 mL) and evaporated twice more followed by lyophilization. The solid was dissolved in 0.1 N HCl (3.0 mL), centrifuged, and applied to a Lobar Lichroprep® RP-18 column and eluted with 0.1 N HCl (2 × 1.5 mL runs were performed). A flow rate of ≈ 3 mL/min was used and fractions (4.5 mL) collected were monitored with ninhydrin visualization. Fractions giving a purple colour with ninhydrin were pooled, separating the product from an earlier eluting contaminant which gave a yellow colour with ninhydrin. The pooled fractions were lyophilized yielding 0.49 g (2.63 mmol, 64%) of a white solid.

R_f (TLC_2 (see 3.2.6.1), EtOAc, ninhydrin visualization) 0.12; ^{1}H NMR (200 MHz, D_2O) δ 3.99 (t, 1H, J = 6.6 Hz, **H**α), 2.55 (m, 2H, **H**γ × 2), 2.24-1.98 (m, 2H, **H**β × 2), 1.99 (d, 3H, J = 139.4 Hz, S^{13}**C**H$_3$); ^{13}C NMR (50.3 MHz, D_2O) δ 172.6 (**C**O$_2$H), 52.5 (**C**α), 29.3, 28.8 (**C**β, **C**γ), 14.1 (S^{13}**C**H$_3$).

3.2.4.3.2. Requirement of Methionine Supplementation for the Growth of E. coli DH93/pHDM10

In order to incorporate [*methyl*-^{13}C]methionine into LaL, the expression plasmid pHDM10 was isolated from *E. coli* TG-1/pHDM10 and used to transform *E. coli* DH93 employing standard methodology (section 2.2.4). The strain DH93 is auxotrophic for methionine being deficient for the two forms of methionine synthetase in *E. coli* (*metE*$^-$ and *metH*). The DH93 strain was constructed by Dr. E. Daub using *E. coli* MG1655 and bacteriophage P1 transduction using GW2537 (*metH*174::Tn*5*, from Graham Walker) and RK4349 (*metE*::Tn*10*, from Robert Kadner) as donors.

The amount of methionine supplementation required to sustain the growth of *E. coli* DH93/pHDM10 was determined as follows. An overnight culture of *E. coli* DH93/pHDM10 (5 mL) was prepared in M9 minimal medium (Appendix A) supplemented with glucose (0.4%), ampicillin (40 μg/mL) and *L*-Met (1.0 mM). The cells were collected, washed once with M9 media (10 mL) and suspended into M9 media (5 mL) to remove residual methionine carried over from the starter culture. The cells were then subcultured (1:100) into fresh M9 media (15 mL) supplemented with glucose and ampicillin as above as well as with various concentrations of *L*-Met ranging from 0.01-1.0 mM. The cultures were incubated at 37° C with shaking and aliqouts (1.0 mL) of each culture were removed at various times and the absorbance at 600 nm was recorded. During the time required to obtain absorbance measurements for all the cultures (≈ 10-15 min) the cells were maintained at 37 °C without shaking. As a control, the growth of *E. coli* MG1655 (prototrophic) in M9 media supplemented with glucose (0.4%) was also monitored as described above.

3.2.4.3.3. Preparation of [methyl-^{13}C]Methionine Labelled LaL

i) Purification from E. coli DH93/pHDM10. From the study performed above, it was determined that a *L*-methionine supplementation of 0.2 mM was sufficient to sustain normal growth of *E. coli* DH93/pHDM10. For this reason, supplementation with *DL*-[*methyl*-^{13}C]methionine hydrochloride was performed at 0.4 mM.

An overnight culture of *E. coli* DH93/pHDM10 grown in M9 medium supplemented with glucose (0.4%), ampicillin (40 μg/mL) and *DL*-[methyl-^{13}C]methionine hydrochloride (0.4 mM) was subcultured (1:50) into 4 × 1.5 L volumes (using 4 L flasks) of the same medium. Cultures were grown at 37 °C with shaking (180 rpm) to an OD_{600} of 0.5-0.7 (≈ 6 hr). To each flask was added solid IPTG (180 mg, to 0.5 mM) and the cultures were incubated an additional 4.5 hr to a final absorbance of 1.1-1.3 at 600 nm. Cells (2.3 g wet weight/L) collected by centrifugation (10000 × g) were suspended in 50 mM KPB, pH 7.0 (60 mL) and disrupted by passage through a French Press as described previously in section 2.2.9.4.

Purification of labelled lysozyme was carried out as described in section 2.2.9.4 with the exception that only single chromatographic runs over both the Mono-S HR10/10 and Phenyl-Superose HR 10/10 columns were performed. The Phenyl-Superose fraction was dialysed (Spectra/Por® 7, 3500 MWCO) extensively against 5 mM KPB, pH 7.0 and lyophilzed to dryness. Protein was quantitated by mass of the lyophilized powder after correcting for mass arising from the buffer salts (section 2.2.10.2). In this case, the dialysis solution (10.7450 g) yielded 11.95 mg of protein/salts in which 0.841 ± 0.017 mg of salt was obtained per gram of dialysis buffer. Therefore, the protein yield

= 11.95 mg - [(10.7450 g)(0.841 mg/g)]

= 2.91 mg from 6 L of culture.

As such, 0.49 mg of labelled LaL was obtained per litre of culture.

ii) Purification from E. coli B834(λDE3)/pLR102. Plasmid pLR102 was used to transform *E. coli* B834(λDE3) employing standard methodology (section 2.2.4). A similar study as described above (3.2.4.3.2) was performed with this strain and also revealed that a *L*-methionine supplement of 0.2 mM sustained its normal growth. An overnight culture of *E. coli* B834 harboring pLR102 grown in M9 minimal medium supplemented with glucose (0.4%), ampicillin (40 μg/mL) and 0.2 mM *L*-[*methyl*-^{13}C]methionine (Aldrich, 99% ^{13}C) was subcultured (1:50) into 1.5 L of the same medium. The culture was grown at 37 °C with shaking (180 rpm) to an absorbance of 0.46 at 600 nm (4.5 hr). Solid IPTG was added (268 mg, to 0.75 mM) and cells were grown an additional 6 hr to a final absorbance of 1.5 at 600 nm. Cells were harvested by centrifugation (10000 × g) and were suspended in 50 mM KPB, pH 7.0 (30 mL) and disrupted by French press as described previously (2.2.9.4).

Purification of labelled LaL was performed with the established protocol described in section 2.2.9.5 by sequential chromatography over S-Sepharose Fast Flow, Mono-S and Phenyl-Superose. The conditions for each of the column purification procedures were the

same as those given in the legends of Figs. 2.15 and 2.24 (S-Sepharose Fast Flow), Figs. 2.16 and 2.25 (Mono-S) and Figs. 2.17, 2.18 and 2.26 (Phenyl Superose) with the following exceptions:

1) The dialysed S-Sepharose Fast Flow fraction (44 mL) was chromatographed over the Mono-S HR 10/10 column in 2 individual runs collecting both a major and minor fraction as described previously (2.2.9.5). The minor fractions were pooled (≈ 15 mL) and diluted to 45 mL with MQW and then reapplied to the Mono-S column pooling the major peak with the major peaks collected from the first two runs (the chromatograms are virtually identical to those shown in Fig. 2.25 A and B).

2) The pooled Mono-S fraction was chromatographed over the Phenyl-Superose HR 10/10 column in a single run (using the conditions given in Figs. 2.17 and 2.26 A). The active fraction was then reapplied in a single run to the column (using the conditions given in Figs. 2.18 and 2.26 B) to concentrate the sample.

Following chromatography, the sample was dialysed extensively and lyophilized to dryness as described in section 2.2.10.2. Protein was quantitiated by mass following buffer salt correction (giving 46.41 mg) and by A_{280} measurements using the ε_{280} of 31712 M^{-1} cm^{-1} (giving 45.77 mg), giving an average yield from the two quantities of 30.7 mg per L of culture. A sample of the purified labelled protein was desalted by reverse phase chromatography and subjected to ESMS using the parameters given in section 2.2.11.

3.2.4.3.4. NMR Spectroscopy

Spectra were recorded on a Bruker AMX500 spectrometer at 298 K using a standard 5 mM NMR tube. A published pulse sequence (Bax et al., 1990) was used to record [^{1}H-^{13}C] heteronuclear multiple-quantum correlation (HMQC) spectra. The NMR data were processed using Bruker UXNMR software. NMR experiments and sample preparation were as follows: the 11.95 mg protein/buffer salt powder purified from *E. coli* DH93/pHDM10 (of which 2.91 mg is ^{13}C labelled LaL, see 3.2.4.3.3 (*i*)) described above was dissolved in 0.42 mL of D_2O containing 0.5 mM 2,2,3,3-tetradeutero-3-trimethylsilyl-propionic acid sodium salt (TSP) as an internal reference. The protein concentration is calculated to be 0.39 mM and the resulting buffer is 128 mM KPB, pH 7.0 (i.e. 10.7450 mL/0.42 mL × 5 mM = 128 mM, see section 2.2.10.2 for the rationale for this calculation). A spectra was recorded of this sample. To the same sample was then added solid $(GlcNAc)_5$ first to 2 mM, and subsequently to 5 and 10 mM (with the assumption that the volume of the sample did not change), and spectra were recorded for each. Upon each addition of sugar, the NMR tube was gently agitated to obtain complete dissolution. The

labelled active enzyme was recovered and purified from the $(GlcNAc)_5$ by chromatography over Mono-S HR 10/10 following the conditions used during purification of LaL (see the legend for Fig. 2.16). The active fraction was dialysed extensively against 5 mM KPB, pH 7.0 and lyophilized to dryness. In this case, the dialysis solution (7.1090 g) yielded 7.76 mg of protein/salts in which 0.864 ± 0.009 mg of salt was obtained per gram of dialysis buffer. Therefore, the protein yield

$$= 7.76 \text{ mg} - [\,(7.1090 \text{ g})(0.864 \text{ mg/g})\,]$$
$$= 1.62 \text{ mg.}$$

The recovered protein was dissolved in 0.41 mL D_2O (containing 0.5 mM TSP) thereby giving a 0.22 mM labelled protein and 87 mM buffer solution. A spectra of the enzyme alone was again acquired, and then solid maltotetraose was added to first 5 and then 10 mM, and spectra were recorded for each.

3.2.5. Peptide Synthesis and Purification

3.2.5.1. Solid Phase Peptide Synthesis

i) Frequently used abbreviations in this section.

BOC:	*t*-butoxycarbonyl	IPA:	isopropanol
DCM:	dichloromethane (CH_2Cl_2)	HOBt:	1-hydroxybenzotriazole
DIPEA:	diisopropylethylamine	TFA:	trifluoroacetic acid
DMF:	dimethylformamide	TFMSA:	trifluoromethanesulfonic acid

ii) General. Except when cited otherwise, the procedures described below were performed essentially as described by Stewart & Young (1984) or as suggested by Drs. Anna Crivici and Gord Adamson. Any procedures involving solid phase peptide synthesis (SPPS) resins were performed in glassware pre-treated with trimethylchlorosilane. When mention is made to the drying of resins or peptide-substituted resins, the resin was dried *in vacuo* over P_2O_5 for at least 6 hr.

iii) Wash Cycles. Peptidyl-resins were subjected to specific wash cycles with desired solvents or solutions for the applications described below. In each case, the solvent added was of an amount that freely suspended the resin. The wash solvent in use was removed by filtration prior to treatment with the following solvent.

Cycle 1 - General Wash: the resin was washed in order with DMF, IPA and DCM. The cycle was repeated 3 times finishing with 2 additional washes with DCM.

Cycle 2 - BOC Deprotection Wash: a solution of 1:1 (v/v) TFA:DCM containing 1% anisole was prepared fresh before use. The resin was treated with the solution for first 2 min and

then a second wash was performed for 15 min. The resin was gently agitated during the treatment. CAUTION: CO_2 is generated requiring sufficient venting to relieve build up of pressure. The resin was then washed two times in order with DCM and DMF.
Cycle 3 - Neutralization Wash: a 10% (v/v) solution of DIPEA in either DCM or DMF was prepared fresh before use. The resin was washed 4 times with the solution and then 2 times with either DCM or DMF.
Cycle 4 - The resin was washed successively with each of the following: DMF, 1:1 (v/v) DMF:MQW, DMF, MeOH, DCM.
Cycle 5 - The resin was washed successively with each of the following: DMF, DCM, MeOH, DCM.

iv) Preparation of BOC-D-Ala-COO-Resins

Peptide synthesis was attempted using three different support resins: (1) hydroxymethyl resin (Peninsula Laboratories, 0.8 mequiv. -OH/g); (2) Merrifield Chloromethyl resin (Sigma, 1.3 mequiv. -Cl/g); (3) Phenylacetamidomethyl (PAM) Bromomethyl resin (INL Hukabel, 0.45 mequiv. -Br/g). The quantities of reagents to be used for introduction of BOC-*D*-Ala-OH was calculated based on the substitution values given above for each resin.

Resin (1). To DMF (20 mL) was added BOC-*D*-Ala-OH (1.514 g, 8.0 mmol), N-N'-diisopropylcarbodiimide (1.253 mL, 8.0 mmol), HOBt-H_2O (1.225 g, 8.0 mmol) and DMAP (0.098 g, 0.8 mmol). The solution was stirred at room temperature (15 min) and then transferred to dried hydroxymethyl resin (5.00 g, 4 mequiv. total substitution). The coupling reaction was shaken for 15 hr, the resin was washed with Cycle 1 and dried. Coupling efficiency was calculated as described below (*vi*).

Resin (2) and (3). BOC-*D*-Ala was introduced onto the resins by the procedure of Gisin (1973) using an excess of its cesium salt. A measured amount of BOC-*D*-Ala (2-3 g) was added to ethanol (6-7 mL/g BOC-*D*-Ala) and then MQW (2-3 mL/g BOC-*D*-Ala) was added. The pH of the solution was brought to approx. 7.1 with 2 M aqueous Cs_2CO_3 and the solution was evaporated *in vacuo*. After repeated evaporation to dryness with benzene and drying *in vacuo*, the cesium salt of BOC-*D*-Ala was obtained as a white powder. An amount of the salt giving 2.5 equivalents was added to dried resin (*2*) or (*3*) in DMF (7-8 mL/g resin) and the slurry was shaken for 10-15 hr at 50 °C. The resin was washed with Cycle 4 and dried. Coupling efficiency was calculated as described below (*vi*).

Following the introduction of BOC-*D*-Ala to each of the three respective resins, unreacted functional groups on the resin were capped. This was achieved by suspending the resin in DMF (6-7 mL/g) and treating with acetic anhydride (10 equiv.) and DIPEA (1 equiv.) for each original equivalent of functional group. The resin was shaken for 2-4 hr and then washed using Cycle 1 and dried.

v) Coupling of Amino Acids to Peptidyl-Resins

Both a tetra and pentapeptide of sequence *L*-Ala-*D*-iso-Glu-*L*-Lys-(*D*-Ala)$_n$ (n=1 or 2) were desired. BOC methodology was employed using the following commercially available amino acids; BOC-*D*-Ala-OH, BOC-*L*-Ala-OH, BOC-*D*-Glu(αOBzl)-OH, and BOC-*L*-Lys(Fmoc)-OH. Peptides were synthesized in order from their C-terminus onto the BOC-*D*-Ala-COO-Resins prepared as described above (*iv*).

The BOC protecting group of BOC-(AA)$_n$-COO-Resin was removed using wash Cycle 2 and then the resin was immediately neutralized using wash Cycle 3 (deprotection was verified by a positive Kaiser test, see (*vi*) below). Pre-activation of amino acids with coupling reagents was performed during the time of deprotection and was such that the time of completion of the pre-activation reactions were made to coincide with the completion of the neutralization wash as closely as possible.

v.1) BOP/HOBt Coupling. The pentapeptide was synthesized on both the hydroxymethyl and Merrifield resins (Resins (1) and (2), see (*iv*) above) employing BOP (benzotriazole-1-yloxytris(dimethylamino)phosphonium hexafluorate) and HOBt in the presence of 4-methylmorpholine (NMM) as the carboxyl activating agents. Three equiv. of each of the next protected amino acid, BOP, and HOBt and 4.5 equiv. of NMM were used for each equiv. (as determined below (*vi*)) of substituted peptidyl-resin. An example is given for the coupling of BOC-*D*-Ala-OH to H_2N-*D*-Ala-COO-Resin in which the total substitution on the resin was determined to be 1.65 mmol. To ≈ 10 mL DMF (typically 3-4 mL/g resin was used) was added in order BOC-*D*-Ala (0.936 g, 4.95 mmol), BOP (2.189 g, 4.95 mmol), HOBt·H_2O (0.758 g, 4.95 mmol) and NMM (0.726 mL, 6.60 mmol). The reaction was stirred for 15 min and then transferred to the BOC deprotected, neutralized resin. A small amount (2-3 mL) of DMF was used to rinse residual reagents into the resin from the reaction flask. The resin was shaken for 2-3 hours (the addition of the Lys and Glu derivatives required longer times) or until coupling was deemed complete by indication of a negative Kaiser test. The resin was then washed (Cycle 1) and dried.

v.2) DCC/HOBt Coupling. The tetrapeptide was synthesized on the PAM resin (Resin (3), see (*iv*) above) employing dicyclohexylcarbodiimide (DCC) and HOBt as the

carboxyl activating agents. Four equiv. of each of the next protected amino acid, DCC and HOBt were used for each equiv. (as determined below (*vi*)) of substituted peptidyl-resin. An example is given for the coupling of BOC-*D*-Glu(αOBzl)-OH to H_2N-*L*-Lys(Fmoc)-*D*-Ala-COO-Resin in which the total substitution on the resin was determined to be 0.315 mmol. To stirring DCM (2.5 mL) was added BOC-*D*-Glu(αOBzl)-OH (0.425 g, 1.26 mmol) and DMF was added dropwise (if necessary) until dissolution occurred. Next was added HOBt·H_2O (0.193 g, 1.26 mmol) and DMF again added where necessary and the reaction was stirred for 30-60 s. DCC (1.26 mmol) was introduced from a 0.5 M solution in DCM (2.52 mL). The mixture instantaneously clears after DCC addition and then a precipitate ensues following ≈ 10 s. The reaction was stirred for 15 min and then placed on ice for 10 min. The mixture was filtered transferring the filtrate to the BOC deprotected, neutralized resin rinsing the precipitate with DCM (2 mL). DMF (2 mL) was added to the resin and the resin was shaken until a negative Kaiser test was obtained (typically < 1 hr). The resin was washed (Cycle 5) and dried.

vi) Estimation and Monitoring of Coupling Efficiencies

The scale of peptide synthesis onto BOC-*D*-Ala-COO-Resin was initially based after estimation of the substitution (in mequiv./g resin) of BOC-*D*-Ala onto the resin. This efficiency was estimated by (i) the mass difference of the resin before and after attachment of BOC-*D*-Ala and (ii) the reaction with ninhydrin following the procedure of Sarin et al. (1981). The two methods gave substitution values typically within good agreement with each other.

A more accurate estimation of the substitution was also determined after the addition of BOC-*L*-Lys(Fmoc)-OH to the growing peptide (i.e. BOC-*L*-Lys(Fmoc)-*D*-Ala-COO-Resin or BOC-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala-COO-Resin). This was achieved by quantitating the amount of dibenzofulvene-piperidine adduct (Fields & Noble, 1990) formed upon treatment of the peptidyl-resin with piperidine, which removes the Fmoc protecting group. A small amount of accurately weighed dried BOC-*L*-Lys(Fmoc) containing resin (≈ 10 mg, performed in triplicate) was shaken in 20% piperidine/DCM (2.0 mL, 40 min). A blank (2.0 mL) of only 20% piperidine/DCM was also prepared. The solutions were filtered (through a glass wool plug to remove resin) and diluted to 50.0 mL (using a volumetric flask) with DCM. The absorbance at 301 nm was measured and an absorption coefficient (ε_{301}) of 7800 M^{-1} cm^{-1} for the dibenzofulvene-piperidine adduct was used to calculate the substitution per gram of resin.

Kaiser Test. Routine monitoring of BOC deprotection and amino acid couplings to NH_2-$(AA)_n$-Resin was performed qualitatively using a slight modification of the method of Kaiser et al. (1970). A small amount of resin (≈ 10-20 beads) obtained following BOC deprotection or amino acid coupling (see (*v*) above) was washed (Cycle 1) using a pasteur pipette containing a tight plug of glass wool. The resin was transferred to a small test tube and 1-2 drops of the following solutions were added in order:

1. Ninhydrin Solution: - dissolve 0.5 g ninhydrin in 10 mL n-butanol
2. Phenol Solution: - dissolve 10 g phenol in 2.5 mL n-butanol
3. KCN Solution: - dissolve 3.3 mg KCN in 5 mL MQW
   - dilute 0.2 mL with 9.8 mL pyridine for use

The resin/solutions mixture was heated at 100 °C for exactly 5 min and observed immediately. A positive test, revealing the presence of free amines from a successful BOC deprotection or incomplete amino acid coupling, is indicated by a light to dark blue colourization of the solution and resin. A negative test revealing no free amines is indicated by the lack of or minimal formation of the blue colour.

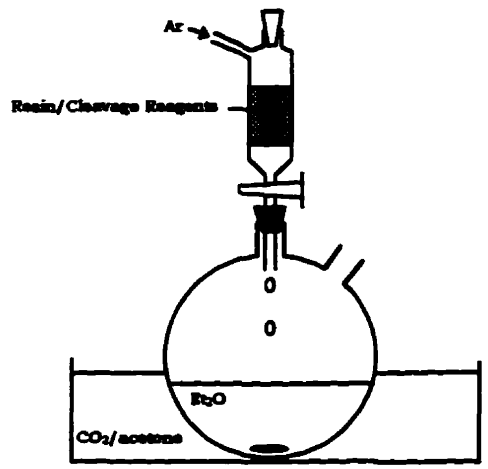
3.2.5.2. Resin Cleavage and Purification

i) Cleavage of Peptides from Resins

TFA used in cleavage reactions was prepared by distillation over P_2O_5. Approximately 0.5 g P_2O_5 was used for 100 mL TFA and the mixture was refluxed for 2 hr prior to collection of the distillate (bp = 72 °C).

Method 1. This method was used to cleave the Lys protected pentapeptide (*L*-Ala-*D*-iso-Glu-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala) from BOC-*L*-Ala-*D*-*iso*-Glu(αOBzl)-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala-COO-Resin for peptides synthesized on both the hydroxymethyl and Merrifield resins. The method was performed under the low TFMSA conditions essentially as described by Tam et al. (1986) which cleaves the peptide from the resin as well as removing the BOC and Bzl protecting groups. For each gram of peptidyl resin, 1 mL of TFMSA, 5 mL of TFA, 3 mL dimethly sulfide and 1 mL anisole were used. The peptidyl-resin was placed in a flask under Ar cooled on a NaCl/ice bath and the TFA (pre-cooled on ice) was added. To the resin was added the anisole and the dimethyl sulfide and the slurry was stirred for 10 min to cool the mixture. Next, the TFMSA was added slowly by dropwise additions and the mixture was stirred on the NaCl/ice bath for 30 min, and then for 4 hr at 25 °C.

The resin slurry was transferred into a filtering vessel as depicted below. The crude pentapeptide was precipitated by allowing the filtrate to drip slowly (≈ 2 drops/s) into a 15-fold volume of anhydrous diethyl ether cooled over CO_2/acetone. The resin was washed with a minimal amount of TFA. The ether precipitation was continued for 2-3 hr and then the precipitate was collected by filtration, washed with cold diethyl ether and lyophilized from acetic acid. The pinkish-red crude pentapeptide (8.5 g was obtained from the synthesis on the hydroxymethyl resin and 6.0 g was obtained from the synthesis on the Merrifield resin) was stored at -20 °C and purified as described below (*ii. 1*).



Schematic of the apparatus used to preciptate crude peptides from resin-cleavage reactions.

Method 2. This method was used to cleave the tetrapeptide (*L*-Ala-*D*-iso-Glu-*L*-Lys-*D*-Ala) from BOC-*L*-Ala-*D*-*iso*-Glu(αOBzl)-*L*-Lys(Fmoc)-*D*-Ala-COO-Resin synthesized on the PAM resin.

The Fmoc protecting group was removed by suspending the resin in 20% piperidine in DCM (≈ 50 mL/g resin) and shaking the resin for 40 min. The resin was washed (Cycle 1) and dried.

The cleavage method was performed under standard TFMSA conditions as described in *Strategies in Peptide Synthesis* (Applied Biosystems Manual, ©1990). For each gram of peptidyl-resin, 1 mL of TFMSA, 10 mL of TFA, 1 mL of thioanisole and 0.5 mL of ethanedithiol were used. The peptidyl-resin was placed in a flask under Ar cooled in a ice bath. The thioanisole and ethanedithiol were added, the mixture stirred for 10 min, and then the TFA (pre-cooled on ice) was added and let stir for 10 min. Next, the TFMSA was added slowly by dropwise additions, the mixture was stirred on ice for 5 min, and then for 1 hr at 25 °C.

The crude tetrapeptide was precipitated as described above (*Method 1*) except that the diethyl ether precipitation was continued overnight at -80 °C. The precipitate was collected by filtration and washed with cold diethyl ether. The peptide was recovered by washing the filter-funnel with acetic acid and lyophilizing the filtrate. The pinkish-red crude tetrapeptide (≈ 150 mg) was purified as described below (*ii.3*).

ii) Purification of Peptides

ii.1) L-Ala-D-iso-Glu-L-Lys(Fmoc)-D-Ala-D-Ala (**Peptide 1**)

Column 1. A portion of the crude pentapeptide (≈ 1 g) was dissolved with stirring into a minimal amount of acetic acid (1.5-2 mL). The solution was applied to a Lobar Lichroprep® RP-18 column equilibrated with 20% CH_3CN in MQW (containing 0.1% TFA). The column was eluted (4 mL/min) sequentially with equilibration buffer (120-150 mL) and then 40% CH_3CN in MQW (0.1% TFA) and fractions (9 mL) were collected. Fractions which quenched at 254 nm and were ninhydrin positive were pooled (typically fractions 9 to 21 of the 40% CH_3CN buffer) and the solvent evaporated to remove CH_3CN and then lyophilized. The column was regenerated (thoroughly washing with each of 95% CH_3CN, 50% CH_3CN, MQW and then with equilibration buffer) and more of the crude peptide was processed. The peptide purified from this column (the Lobar peptide) accounted for ≈ 8-10% of the mass of the crude material applied.

Column 2. The Lobar peptide (30 mg) was slowly added to a stirring solution of 60% CH_3CN in MQW (0.1% TFA, 0.7 mL). NOTE: it is important to add the peptide to the solvent for dissolution to occur. The solution was centrifuged and applied to the μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge) equilibrated with 80% Buffer A and 20% Buffer B and eluted at 6.0 mL/min using the linear gradient given below:

Time (min)	% Buffer A	% Buffer B
0	80	20
5	80	20
15	75	25
59	64	36
60	5	95
68	5	95
70	80	20

Buffer A: 0.1% TFA in MQW
Buffer B: 0.1% TFA in CH_3CN

The eluent was monitored at 210 and 265 nm and from 200-400 nm to obtain spectra of individual peaks. The Fmoc functionality absorbs at 265 nm (major) and 289 and 300 nm

(minor). The large peak eluting between ≈ 35 to 47 min was collected. A trailing shoulder was evident from the chromatograms and collection of the peak was terminated when the absorbance at 265 nm of the eluent had decreased to ≈ 1.5 au. The fraction was evaporated to remove CH_3CN and then lyophilized. The peptide purified from this column accounted for ≈ 80-85% of the mass of the Lobar purified peptide applied.

The HPLC purified peptide **1** was characterized by NMR, MALDI MS and CZE (see 3.3.3.1)

ii.2) L-Ala-D-iso-Glu-L-Lys-D-Ala-D-Ala (**Peptide 2**)

To DMF (1.2 mL) was added the HPLC purified peptide **1** (25.0 mg, 35.2 μmol) and piperidine (0.3 mL). The mixture was stirred for 30 min at room temperature and then the solvent was removed *in vacuo*. To the solid was added Buffer A (see below, 4 mL), the material was stirred for 30 min, centrifuged (BioFuge A, 13000 rpm), and the supernatant was filtered (0.22 μM).

Column 1. Aliquots of the filtrate (< 0.8 mL) were individually applied to a Mono-S HR 10/10 column equilibrated with 20 mM KPB, pH 3 (Buffer A). The column was eluted at 2.0 mL/min using the linear gradient given below with detection at 214 nm:

Time (min)	% Buffer A	% Buffer B
0	100	0
2	100	0
26	88	12
28	0	100
35	0	100
37	100	0

Buffer A: 20 mM KPB, pH 3
Buffer B: 1.0 M KCl (w/v) in Buffer A

Two partially resolved peaks were collected. A major peak eluting between ≈ 20 to 24 min and a minor peak eluting between ≈ 15 to 20 min. The elution profile was found to be dependent on the amount of material loaded. The intensity of the minor peak was reduced and even eliminated when smaller amounts of material were loaded onto the column. If the minor peak was collected and then a portion re-applied to the column, only the major peak would be observed. It was concluded that under the chromatographic conditions employed, the pentapeptide would elute in this two-peak profile. Therefore,

both the major and minor peaks were pooled from each chromatographic run and lyophilized giving a white solid.

Column 2. The entire solid from above was dissolved into 0.2% TFA in MQW (≈ 1.6 mL), centrifuged and applied in two individual runs (≈ 0.8 mL loaded for each) to the μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge) equilibrated with Buffer A and eluted at 5.0 mL/min using the linear gradient given below:

Time (min)	% Buffer A	% Buffer B
0	100	0
5	100	0
25	84	16
27	5	95
33	5	95
35	100	0

Buffer A: 0.2% TFA in MQW
Buffer B: 0.2% TFA in CH_3CN

The eluent was monitored at 214 nm and the fraction eluting between ≈ 20.5 to 23.0 min was collected from each of the two runs. The pooled fraction was evaporated to remove CH_3CN and then lyophilized. Typically, 15 mg (≈ 31 μmol) of purified, Fmoc deprotected peptide **2** is recovered from 25 mg (35.2 μmol) of the HPLC Fmoc protected peptide **1** following Fmoc removal and the two column chromatographic purification given above, representing approximately 88% recovery. Purified peptide **2** is a white solid when dry but becomes sticky and acquires a glassy appearance when exposed to the atmosphere. Peptide **2** was stored under Ar at -20 °C and characterized by NMR, MALDI MS and CZE (see 3.3.3.1).

ii.3) L-Ala-D-iso-Glu-L-Lys-D-Ala (**Peptide 3**)

The crude tetrapeptide (≈ 150 mg) was dissolved with stirring into a minimal amount of acetic acid (≈ 1.0 mL) and centrifuged. The sample was applied in two individual runs to the μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge) equilibrated with 0.1% TFA in MQW. The column was eluted at 5.0 mL/min using an immediate 1%/min linear gradient to 0.1% TFA in CH_3CN. The eluent was monitored at 214 nm and the void fraction eluting from ≈ 9.5 to 18 min was collected, the solvent was evaporated to remove CH_3CN and then lyophilized.

The pooled lyophilized material from the two runs above (95 mg) was dissolved with stirring into 0.25% TFA in MQW (2.0 mL) and centrifuged. Aliquots of the sample (0.4 mL) were applied in individual runs to the μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge) equilibrated with Buffer A and eluted at 5.0 mL/min using the linear gradient given below with detection at 214 nm:

Time (min)	% Buffer A	% Buffer B
0	100	0
10	100	0
30	80	20
33	5	95
45	5	95
47	100	0

Buffer A: 0.25% TFA in MQW
Buffer B: 0.25% TFA in CH_3CN

The peak eluting during the gradient from ≈ 20 to 26 min was collected and lyophilized from each run. NOTE: the retention time of peptide **3** was found to decrease marginally with increased load onto the column. Less than 20 mg of material from the first chromatography should be applied for optimal separation using the conditions given directly above. A total of 55 mg of peptide 3 was obtained and characterized by NMR, ESMS and CZE (see 3.3.3.1). The peptide was stored in aliquots as a 100.0 mM solution in MQW at -20 °C.

3.2.6. Synthesis of Substrates

3.2.6.1. General

Pyridine, methylene chloride (CH_2Cl_2) and 1,2-dichloroethane were dried over CaH_2 and distilled before use. Dimethylformamide (DMF) was distilled and dried over 4 Å molecular sieves. Methanol was dried by refluxing in the presence of Mg filings (5 g/L) and I_2 (0.5 g/L) for 4 hr and collected by distillation and stored over 4 Å molecular sieves. Anhydrous THF was prepared by distillation from sodium in the presence of benzophenone. Benzaldehyde was prepared by washing successively with 10% Na_2CO_3 (2×), saturated Na_2SO_3 (2×) and DDW, then dried over anhydrous $MgSO_4$ and distilled under vacuum at ≈ 60 °C. *p*-Nitrophenol was recrystallized from CH_2Cl_2.

TLC was performed on two types of aluminum-backed silica gel plates. When given as TLC_1, plates were 0.2 mm Merck Kieselgel 60 F_{254} (Merck; Darmstadt, Germany). When given as TLC_2, plates were 0.25 mm Alugram Sil G/UV254 (Macherey-Nabel; Dueren,

Germany). Vanillin dip for TLC development was prepared by mixing vanillin (2.6 g) with EtOH (90 mL), HOAc (2 mL) and H_2SO_4 (3.4 mL) and was prepared freshly every two weeks. Ninhydrin dip was prepared as a 0.2% solution in EtOH. Phosphomolybdic acid solution was used as a 4% solution in ethanol. Flash chromatography (Still et al., 1978) was performed using silica gel (70-230 mesh) obtained from either Merck or EM Science (Gibbstown, NJ).

Compounds **3-12** could be detected on TLC plates by their 254 nm quenching or through visualization with vanillin dip (ninhydrin and phosphomolybdic acid dips also develop these compounds but are not as effective as the vanillin dip). In addition, compounds **5-12** could also be visualized by prolonged heating of the TLC plate in which the compounds appeared as yellow spots.

Melting temperatures were obtained using a Mel-Temp melting point apparatus and are uncorrected. Elemental analysis was performed at M-H-W Laboratories (Phoenix, AZ).

3.2.6.2. Synthesis of p-Nitrophenyl 2-Acetamido-2-deoxy-β-D-Glucopyranoside (β-PNP-GlcNAc)

2-Acetamido-1,3,4,6-tetra-O-acetyl-2-deoxy-D-glucopyranose (3) (GlcNAc peracetate)

Freshly distilled acetic anhydride (320 ml) and pyridine (275 ml) were cooled to 0 °C under Ar and 2-acetamido-2-deoxy-D-glucose (50.5 g, 0.228 mol) was added over 15 min. The suspension was stirred for 10 min after which 4-dimethylaminopyridine (2.75g, 22.6 mmol) was added slowly over 10 min. The mixture was stirred on ice for 10 min, the ice bath was removed and the solution then stirred for 4 hr at room temperature. The clear solution was cooled to 0 °C and 120 ml of methanol was added over 40 min. The solvent was removed *in vacuo* and the resulting oil was dissolved in CH_2Cl_2 (550 ml) and washed successively with ice cold 2N HCl (150 ml), a mixture of sat. $NaHCO_3$ (150 ml) and brine (100 ml), and brine (200 ml). The organic layer was dried over anhydrous $MgSO_4$, the solvent was removed *in vacuo* and the product was dried under vacuum yielding 87.25 g (0.224 mol, 98%) of a white solid. R_f (TLC_1, $CHCl_3$:MeOH, 19:1) 0.28 (β-OAc), 0.33 (α-OAc); ^{1}H NMR (250 MHz, $CDCl_3$) δ 6.17 (1H, d, $J_{1\beta,2}$ = 3.5 Hz, **H-1β**), 5.97 (1H, d, $J_{NH,2}$ = 9.0 Hz, **NH**Ac), 5.08-5.35 (2H, m, **H-3** and **H-4**), 4.50 (1H, ddd, $J_{2,1\beta}$ = 3.5 Hz, $J_{2,NH} \approx J_{2,3} \approx$ 9-11 Hz, **H-2**), 4.26 (1H, dd, $J_{6b,6a}$ = 12.3 Hz, $J_{6b,5}$ = 4.2 Hz, **H-6b**), 3.95-4.16 (1H, m, **H-5**; 1H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{6a,5}$ = 2.2 Hz, **H-6a**), 1.94, 2.05, 2.07, 2.09, 2.19 (s x 5, 15 H, 4 x O**Ac** and NH**Ac**); ^{13}C NMR (62.9 MHz, $CDCl_3$) δ 171.3, 170.5, 170.0, 168.9, 168.5 (4 x $OCOCH_3$ and $NHCOCH_3$), 90.5 (**C-1**), 70.5 (**C-5**), 69.5, 67.5 (**C-3**, **C-4**), 61.4 (**C-6**), 50.8 (**C-2**), 22.8, 20.7, 20.5 (2), 20.4 (4 $OCOCH_3$ and $NHCOCH_3$).

NOTE: α denotes the anomeric posititon of the OAc. An anomeric ratio of α:β ≈ 15:1 was calculated from appropriate signal integrations. In the assignment given, H-1β refers to the H in the β position (i.e. α OAc). Reference values for ^{1}H and ^{13}C NMR chemical shifts reported previously (Nanjo et al., 1988a; Vårum et al., 1991; Agrawal, 1992) were used to aid in the interpretation of and assignment of the spectra.

2-Methyl-(3,4,6-tri-O-acetyl-1,2-dideoxy-α-D-glucopyrano)-[2,1-*d*]-2-oxazoline (4)

The compound was synthesized according to the method of Nakabayashi et al. (1986) with slight modification. To a two-neck flask equipped with a condenser was added GlcNAc peracetate **3** (10.00 g , 25.71 mmol) and anhydrous 1,2-dichloroethane (150 mL) under Ar. To the stirring solution was slowly added trimethylsilyl trifluoromethanesulfonate (5.5 mL, 28.46 mmol) by syringe and the reaction was heated to 45 °C. When TLC_1 (CHCl:MeOH, 19:1) showed complete reaction (18-24 hr), the reaction was allowed to cool to room temperature and was taken up in CH_2Cl_2 (550 mL). The solution was washed successively with sat. $NaHCO_3$ (120 ml) and brine (100 ml), the organic layer was dried over anhydrous $MgSO_4$, the solvent was removed *in vacuo* and the product dried under vacuum. The resulting heavy amber oil (8.55 g, 26.0 mmol, ≈ 100%) was of sufficiently high purity (>99%) according to TLC and the ^{1}H NMR spectrum to be used as a starting material in the subsequent step. R_f (TLC_1, $CHCl_3$:MeOH, 19:1) 0.44; ^{1}H NMR (250 MHz, $CDCl_3$) δ 5.97 (1H, d, $J_{1,2}$ = 7.4 Hz, **H-1**), 5.26 (1H, dd, $J_{3,2} \approx J_{3,4} \approx$ 2.4 Hz, **H-3**), 4.93 (1H, m, $J_{4,5}$ = 9.0 Hz, **H-4**), 4.11-4.19 (3H, m, $J_{6,5}$ = 4.3 Hz, **H-2**, **H-6** ×2), 3.61 (1H, dt, $J_{5,6}$ = 4.3 Hz, $J_{5,4}$ = 9.0 Hz, **H-5**), 2.07-2.13 (12H, m, 3 × OAc and N=C-**Me**). The ^{1}H NMR spectrum described is consistent with that reported for the synthesis of **4** using different methodology (Colon et al., 1991).

***p*-Nitrophenyl-2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-*D*-glucopyranoside (5) (β-PNP-GlcNAc peracetate)**

The oxazoline **4** (8.55 g, 26.0 mmol) was dissolved in CH_2Cl_2 (20 mL) and *p*-nitrophenol (24.3 g, 175 mmol) was added with vigorous stirring after which the solvent was removed *in vacuo*. The resulting mixture was connected to an aspirator and placed into an oil bath heated at 165 °C. After all the *p*-nitrophenol had melted (ca. 1 min), the melt was allowed to react under diminished pressure for 12 min with stirring. The mixture was cooled to room temperature under reduced pressure and then suspended in CH_2Cl_2 (600 mL) and washed repeatedly with sat. $NaHCO_3$ (200 mL) until the aqueous layer showed no yellow colour. The organic layer was then washed with brine (200 ml)

and dried over anhydrous $MgSO_4$. Decolorizing charcoal (1 g) was added, the solution was stirred for 15 min and then filtered over a bed of Celite™ and the solvent evaporated. The solid was triturated with ethyl acetate (40 mL), collected by filtration and then washed successively with EtOAc (50 ml) and hexane (50 ml). To remove remaining traces of *p*-nitrophenol, the solid was suspended in CH_2Cl_2 (800 ml) and washed with aqueous 0.1 N NaOH (3 × 100 ml) and then brine (100 mL). The organic layer was dried over anhydrous $MgSO_4$, concentrated *in vacuo* and then dried under vacuum yielding 7.62 g (16.27 mmol, 63%) of an off-white solid.

mp 229-234 °C (lit: 240 °C, Leaback & Walker, 1957); R_f (TLC_1, CH_2Cl_2:MeOH, 19:1) 0.36; ^1H NMR (200 MHz, DMSO-d_6) δ 8.22 (2H, d, $J_{m,o}$ = 9.2 Hz, ***m*-PNP-H** × 2), 8.09 (1H, d, $J_{NH,2}$ = 9.0 Hz, **NHAc**), 7.23 (2H, d, $J_{o,m}$ = 9.2 Hz, ***o*-PNP-H** × 2), 5.55 (1H, d, $J_{1,2}$ = 8.4 Hz, **H-1**), 5.23 (1H, dd, $J_{3,4} \approx J_{3,2} \approx$ 9.6-10.2 Hz, **H-3**), 4.94 (1H, dd, $J_{4,3} \approx J_{4,5} \approx$ 9.5-9.6 Hz, **H-4**), 3.98-4.23 (4H, m, **H-2**, **H-5**, **H-6** × 2), 1.77, 1.95, 2.00 (each a s; 3H, 3H and 6H respectively, O**Ac** × 3 and NH**Ac**).

***p*-Nitrophenyl-2-acetamido-2-deoxy-β-*D*-glucopyranoside (6) (β-PNP-GlcNAc)**

To a stirring solution of β-PNP-GlcNAc peracetate **5** (6.83 g, 14.6 mmol) in a mixture of anhydrous CH_2Cl_2 (300 ml) and MeOH (100 ml) under argon was added NaOMe (1.5 mmol; 1.5 ml of a 1 M solution prepared by the addition of 0.1 g of hexane washed Na metal to 4.3 mL of anhydrous MeOH). After 5 min, the product precipitated and was collected by filtration and washed with CH_2Cl_2 (100 ml). The mother liquor was concentrated under reduced pressure to approximately 100 ml and CH_2Cl_2 was added in portions (20 mL) to promote further precipitation of product which was collected and washed as above. The collected product was dried under vacuum giving 4.70 g (13.7 mmol, 94%) of a white to slightly beige solid.

mp 196-198 °C (dec.) (authentic sample (Sigma) 205-206 °C (dec.); lit: 204 °C, Leaback & Walker, 1957); R_f (TLC_1, CH_2Cl_2:MeOH, 4:1) 0.29; ^1H NMR (250 MHz, acetone-d_6 (1 g) and D_2O (8-10 drops) until dissolved) δ 8.13 (2H, d, $J_{o,m}$ = 9.5 Hz, ***m*-PNP-H** × 2), 7.18 (2H, d, $J_{m,o}$ = 9.5 Hz, ***o*-PNP-H** × 2), 5.28 (1H, d, $J_{1,2}$ = 8.4 Hz, **H-1**), 3.97 (1H, t, J = 9.3 Hz, **H-3** or **H-4**), 3.86 (1H, dd, $J_{6a,6b}$ = 11.9 Hz, $J_{6a,5}$ = 2.0 Hz, **H-6a**), 3.42-3.71 (4H, m, **H-2**, **H-3** or **H-4**, **H-5** and **H-6b**), 1.90 (3H, s, NH**Ac**); ^{13}C NMR (62.9 MHz, acetone-d_6 (1 g) and D_2O (8-10 drops)) δ 173.1 ($NHCOCH_3$), 162.9 (**PNP-O-C1**), 143.0 (**PNP-C4-NO_2**), 126.2 (**PNP-C3** × 2), 117.4 (**PNP-C2** × 2), 99.1 (**C-1**), 77.4 (**C-5**), 74.4 (**C-3**), 70.8 (**C-4**), 61.6 (**C-6**), 56.0 (**C-2**), 22.8 ($NHCOCH_3$).

Reference values for the ^{13}C NMR chemical shifts for β-GlcNAc reported previously (Agrawal, 1992) were used in the assignment of the sugar ring carbons.

3.2.6.3. Synthesis of β-PNP-MurNAc and Derivatives

p-Nitrophenyl-2-acetamido-4,6-O-benzylidene-2-deoxy-β-D-glucopyranoside (7) (Bzi-β-PNP-GlcNAc)

To β-PNP-GlcNAc **6** (2.00 g, 5.84 mmol) was added benzaldehyde (25 ml) and zinc chloride (1.00 g, 7.34 mmol) and the mixture was stirred under argon in darkness for 14 hr (TLC typically showed complete reaction after 7-10 hr). The starting materials were insoluble initially but as the reaction proceeded, total dissolution ensued. MQW (30 ml) and hexane (30 ml) were added and the suspension was stirred vigorously for 0.5 hr. The solid was collected and washed twice more with water and hexane as described. The solid was mixed with reagent grade acetone (≈ 30 mL), the solvent was removed *in vacuo* and the product was dried under vacuum giving 2.33 g (5.41 mmol, 93%) of a white powder. mp 233-237 °C (dec.) (lit: 245-247 °C (dec.), Jeanloz et al., 1968); R_f (TLC_1, CH_2Cl_2:MeOH, 9:1) 0.48; ^1H NMR (250 MHz, DMSO-d_6) δ 8.19 (2H, d, $J_{m,o}$ = 9.2 Hz, ***m*-PNP-H** × 2), 7.98 (1H, d, $J_{NH,2}$ = 8.3 Hz, **NH**Ac) 7.34-7.48 (5H, m, **Ph-H** × 5), 7.23 (2H, $J_{o,m}$ = 9.2 Hz, ***o*-PNP-H** × 2), 5.63 (1H, s, Ph-C-**H**), 5.49 (1H, d, $J_{HO,3}$ = 5.3 Hz, **3-OH**), 5.38 (1H, d, $J_{1,2}$ = 8.0 Hz, **H-1**), 4.23†, 3.70-3.88‡, 3.57* (1H†, dd, J = 9.7, 5.1 Hz; 4H‡, m; 1H*, t or dd, J = 8.5 Hz; **H-2, H-3, H-4, H-5, H-6** × 2), 1.80 (3H, s, NH**Ac**); ^{13}C NMR (62.9 MHz, DMSO-d_6) δ 169.5 ($NHCOCH_3$), 161.9 (**PNP-O-C1**), 142.0 (**PNP-C4-NO_2**), 137.6 (**Ph-C1**), 128.8 (**Ph-C4**), 128.0, 126.3, 125.7 (**Ph-C2** × 2, **Ph-C3** × 2, **PNP-C3** × 2), 116.6 (**PNP-C2** × 2), 100.7, 98.5 (Ph-**C**, **C-1**), 80.8 (**C-4**), 70.1, 67.6, 66.2 (**C-3, C-5, C-6**), 55.9 (**C-2**), 23.0 ($NHCOCH_3$).

p-Nitrophenyl-2-acetamido-4,6-O-benzylidene-2-deoxy-3-O-[*D*-1-carboxyethyl]-β-*D*-glucopyranoside (8) (Bzi-β-PNP-MurNAc)

To a two-neck flask equipped with a condenser was added NaH (261 mg, as a 80% dispersion in mineral oil, 8.7 mmol) which was washed with anhydrous THF to remove the mineral oil and anhydrous THF (85 ml) was added. The suspension was cooled to 0 °C on an ice bath, and then Bzi-β-PNP-GlcNAc **7** (1.50 g, 3.5 mmol) was added and the mixture was stirred on ice (10 min) and then allowed to warm to room temperature (the starting material does not dissolve). (*S*)-(-)-2-Bromo-propanoic acid (1.25 mL, 13.9 mmol) was added dropwise to a stirring mixture of anhydrous THF (40 mL) and NaH (523 mg, as a 80% dispersion, washed in THF, 17.4 mmol) cooled to 0 °C in a separate flask. The latter

mixture was added at once to the former, afterwhich the flask was immediately placed into a 70 °C water bath for 10 min (the reaction remains cloudy). The suspension was cooled to 0 °C and then ice water (prepared by freezing of MQW, 80 mL) was added which resulted in complete solubilization of the reactants. The pH of the solution was brought to 2 (pH paper) by the dropwise addition of 2N HCl (ca. 6 ml) and then additional ice water (500 mL) was added. The product precipitated and was collected by filtration and washed with ice water (800 mL). The collected solid was suspended in CH_2Cl_2 (100 ml) and MeOH was added until the solid completely dissolved. The solution was filtered through Whatman™ paper which removed a dark solid. The filtrate was dried over anhydrous $MgSO_4$, concentrated *in vacuo* and dried under vacuum yielding 1.49 g (2.97 mmol, 85%) of a white solid.

mp 230-232 °C (dec.); R_f (TLC_1, CH_2Cl_2:MeOH, 4:1) 0.54; ^1H NMR (250 MHz, DMSO-d_6) δ 12.5 (1H, br s, lactyl CO_2H), 8.18 (2H, d, $J_{m,o}$ = 9.0 Hz, ***m*-PNP** × 2), 7.90 (1H, d, $J_{NH,2}$ = 7.7 Hz, NHAc), 7.38-7.42 (5H, m, **Ph-H** × 5), 7.22 (2H, d, $J_{o,m}$ = 9.0 Hz, ***o*-PNP** × 2), 5.71 (1H, s, Ph-C-**H**), 5.50 (1H, d, $J_{1,2}$ = 7.5 Hz, **H-1**), 4.20-4.33†, 3.70-3.95‡ (2H†, m, 1H from lactyl **CH**-CH_3; 5H‡, m; **H-2**, **H-3**, **H-4**, **H-5**, **H-6** ×2), 1.80 (3H, s, NHAc), 1.27 (3H, d, $J_{Me,CH}$ = 6.7 Hz, lactyl CH-C**H**$_3$); ^{13}C NMR (62.9 MHz, DMSO-d_6), δ 173.9 (lactyl CO_2H), 169.6 ($NHCOCH_3$), 161.7 (**PNP-O-C1**), 142.1 (**PNP-C4-NO$_2$**), 137.5 (**Ph-C1**), 128.8 (**Ph-C4**), 128.1 (2C), 125.7 (4C) (**Ph-C2** × 2, **Ph-C3** × 2, **PNP-C3** × 2), 116.6 (**PNP-C2** × 2), 100.1, 98.2 (Ph-**C**, **C-1**), 80.9 (**C-4**), 77.5, 75.3 (**C-3**, lactyl **C**H-CH_3), 67.6, 65.6 (**C-5**, **C-6**), 54.5 (**C-2**), 22.9 ($NHCOCH_3$), 18.8 (lactyl CH-**C**H$_3$).

_p_-Nitrophenyl-2-acetamido-2-deoxy-3-O-[_D_-1-(methoxycarbonyl)ethyl]-β-_D_-glucopyranoside (9) (β-PNP-MurNAc-CO$_2$Me)

Anhydrous MeOH (10 mL) was cooled to 0 °C under Ar and freshly distilled acetyl chloride (0.5 mL) was added by syringe. Bzi-β-PNP-MurNAc **8** (100 mg, 0.199 mmol) was added at once and the reaction allowed to warm to room temperature and then stirred for 2 hr (NOTE: the starting material is partially insoluble initially but complete dissolution ensues during the reaction). Reagent grade MeOH (10 mL) was added, the solvent was evaporated and the process was repeated (3 ×). To the residue was added $CHCl_3$:MeOH (19:1, ≈ 1 mL) and then MeOH was added dropwise until the material dissolved (NOTE: a small amount of material remains insoluble which was found not to be the product). The sample was applied to a silica gel column (2.7 cm × 16 cm) and eluted by gravity flow with $CHCl_3$:MeOH (19:1) collecting 5 mL fractions. Fractions containing the product (17-27)

were pooled and the solvent evaporated. The product was dried under vacuum yielding 66 mg (0.154 mmol, 77%) of a white solid.

At this stage of purification, the ^{1}H NMR spectrum (see ^{1}H NMR-1 below) of compound **9** revealed no contaminating material. However, although analytical RP-C18 analysis (μBondapak™ C18 Radial-Pak® cartridge) suggested > 99% purity for **9** following the silica gel column, a hydrophobic contaminant was present in very small amounts. The contaminant was probably benzaldehyde dimethoxyacetal, the product of the methanolysis of the benzylidene group. Since **9** was to be evaluated as a potential substrate, it was feared that this contaminant could interfere and therefore **9** was purified further as follows.

Silca column purified **9** (no more than 50 mg) was added to a solution of $CHCl_3$:MeOH (3:7, ≈ 1 mL) and $CHCl_3$ was added dropwise (if necessary) with vortexing until dissolution occurred. The sample was centrifuged and applied to a μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge) using the linear gradient given below with detection at 210 and 300 nm:

Time	Flow Rate (mL/min)	% MQW	% CH_3CN
0	2.5	80	20
1	7	80	20
5	7	80	20
25	7	74	26
29	7	0	100
34	7	0	100
39	7	80	20
45	7	80	20

An increase in column pressure resulted upon sample loading which at 7 mL/min, exceded the column limit. Therefore, the sample was loaded at 2.5 mL/min. After 1 min at 2.5 mL/min, the pressure subsided and the column could then be safely operated at 7 mL/min. The eluent over 17 to 23 min was collected, evaporated to remove CH_3CN and then lyophilized giving a white solid. Between 85-90% of the mass of the material applied to the column was consistently recovered in this fraction, giving an overall recovered yield (encompassing the 2 column purification) of 65-70%.

The characterization given is for **9** purified by the two columns except for ^{1}H-NMR-1 (after silica gel only); R_f (TLC_2, CH_2Cl_2:MeOH, 9:1) 0.43; ^{1}H NMR-1 (250 MHz, DMSO-d_6) δ 8.18 (2H, d, ***m*-PNP-H** × 2), 7.75 (1H, d, **NHAc**), 7.14 (2H, d, ***o*-PNP-H** × 2), 5.47 (1H, d, **4-OH**), 5.24 (1H, d, **H-1**), 4.65 (1H, t, **6-OH**), 4.47 (1H, q, lactyl **CH**-CH_3), 3.30-3.75 (≈9H with

overlap from H_2O at 3.32 ppm, m, **H-2, H-3, H-4, H-5, H-6** × 2, lactyl CH-CO_2CH_3 at 3.63 ppm), 1.78 (3H, s, **NHAc**), 1.29 (3H, d, lactyl CH-**CH_3**); ^{1}H NMR-2 (200 MHz, acetone-d_6 (1 g) and $CDCl_3$ (2-4 drops) until dissolved) δ 8.11-8.19 (2H, m, $J_{m,o}$ = 9.3 Hz, ***m*-PNP-H** × 2), 7.66 (0.4H, d, $J_{NH,2}$ = 8.5 Hz, NHAc), 7.16-7.27 (2H, m, $J_{o,m}$ = 9.3 Hz, ***o*-PNP-H** × 2), 5.34 (1H, d, $J_{1,2}$ = 8.2 Hz, **H-1**), 4.62 (1H, q, $J_{CH,Me}$ = 6.9 Hz, lactyl **CH**-CH_3), 3.55-4.00 (≈9H with overlap from H_2O at 3.59 ppm), m, **H-2, H-3, H-4, H-5, H-6** × 2, lactyl CH-CO_2CH_3 at 3.68 ppm), 1.91 (3H, s, **NHAc**), 1.33 (3H, d, $J_{Me,CH}$ = 6.8 Hz, lactyl CH-**CH_3**);
^{13}C NMR (50.3 MHz, acetone-d_6 (1 g) and $CDCl_3$ (2-4 drops)) δ 175.1 (lactyl **CO_2CH_3**), 171.9 ($NHCOCH_3$), 163.2 (**PNP-O-C1**), 143.1 (**PNP-C4-NO_2**), 126.3 (**PNP-C3** × 2), 117.5 (**PNP-C2** × 2), 99.4 (**C-1**), 81.9 (lactyl **C**H-CH_3), 77.6, 76.3 (**C-3, C-5**), 71.5 (**C-4**), 61.8 (**C-6**), 55.2 (**C-2**), 52.3 (lactyl CH-CO_2**C**H_3), 23.2 ($NHCOCH_3$), 19.2 (lactyl CH-**CH_3**); ESMS: i) exact mass calcd for $C_{18}H_{24}N_2O_{10}$ (M) 428.1431, found 428.24; ii) exact mass calcd for $C_{12}H_{19}NO_7$ (M - [$OC_6H_4NO_2$] - [H]) 289.1162, found 289.17. NOTE: fragmentation was dependent on the cone voltage applied during acquisition; the relative intensity of (M) increased with decreasing cone voltage.

Anal. Calcd for $C_{18}H_{24}N_2O_{10}$ (428.396): C, 50.47; H, 5.65; N, 6.54. Found: C, 50.59; H, 5.80; N, 6.52.

_p_-Nitrophenyl-2-acetamido-3-O-(_D_-1-carboxyethyl)-β-_D_-glucopyranoside Sodium Salt (10) (β-PNP-MurNAc-$CO_2^-Na^+$)

To Bzi-β-PNP-MurNAc **8** (70.0 mg, 0.139 mmol) was added THF (3.0 mL) with stirring, then 2 N HCl (3.0 mL) was slowly added dropwise and then additional THF (1.5 mL) was added with clearing of the solution after 1-2 hr. The reaction was stirred at room temperature for 24 hr after which MQW (5 mL) was added and the pH of the solution was brought to 6.5 (pH meter) with 1 N NaOH (≈ 6-7 mL). The solvent was evaporated to remove THF and then lyophilized. The slightly yellow solid was suspended in MeOH (10 mL) and the solution was filtered through a pipette with a cotton plug. The filtrate was evaporated and the MeOH trituration was repeated to remove as much NaCl as possible.

The solid was dissolved in MQW (≈ 3.5 mL) that had been boiled for 10 min and then cooled (the product has increased solubility in boiled MQW over non-boiled MQW). The sample was centrifuged and applied in two individual runs (≈ 1.8 mL each) to a μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge) using the linear gradient given below with detection at 210 and 300 nm:

Time	Flow Rate (mL/min)	% MQW	% CH_3CN
0	2	100	0
2	5	100	0
35	5	100	0
40	5	5	95
55	5	5	95
59	5	100	0

The eluent over 20 to 32 min was collected and pooled from both runs and lyophilized yielding 56 mg (0.106 mmol (based on a MW of 530.82, see below), 76%) of a white solid. ^{1}H NMR (200 MHz, D_2O (≈ 0.8 mL) and 1 drop acetone as reference (2.04 ppm)) δ 8.00-8.08 (2H, m, $J_{m,o}$ = 9.3 Hz, ***m*-PNP-H** × 2), 6.94-7.02 (2H, m, $J_{o,m}$ = 9.3 Hz, ***o*-PNP-H** × 2), 5.12 (1H, d, $J_{1,2}$ = 8.3 Hz, **H-1**), 4.14 (1H, q, $J_{CH,Me}$ = 6.9 Hz, lactyl **CH**-CH_3), 3.75-3.89†, 3.45-3.70‡ (2H†, m; 4H‡, m; **H-2**, **H-3**, **H-4**, **H-5**, **H-6** × 2), 1.84 (3H, s, **NHAc**), 1.18 (3H, d, $J_{Me,CH}$ = 6.9 Hz, lactyl CH-**CH_3**); ^{13}C NMR (50.3 MHz, D_2O (≈0.8 mL) and 1 drop acetone as reference (29.8 ppm)) δ 180.7 (lactyl **CO_2**$^-$$Na^+$), 174.4 (**NHCOCH$_3$**), 161.4 (**PNP-O-C1**), 142.1 (**PNP-C4-NO_2**), 125.6 (**PNP-C3** × 2), 116.0 (**PNP-C2** × 2), 98.6 (**C-1**), 79.6 (lactyl **CH**-CH_3), 77.5, 75.7 (**C-3**, **C-5**), 68.7 (**C-4**), 60.0 (**C-6**), 53.7 (**C-2**), 21.9 (**NHCOCH$_3$**), 18.4 (lactyl CH-**CH_3**); ESMS: i) exact mass calcd for $C_{17}H_{21}N_2O_{10}Na$ 436.1094, found 436.28, ii) exact mass calcd for $C_{17}H_{22}N_2O_{10}$ (the free acid which is generated from the acidic ESMS delivery solvent) 414.1274, found 414.29, iii) exact mass calcd for $C_{17}H_{22}N_2O_{10}$·$H^{35}Cl$ 450.1041, found 450.28, iv) exact mass calcd for $C_{17}H_{22}N_2O_{10}$·$H^{37}Cl$ 452.1011, found 452.25.

Anal. Calcd for $C_{17}H_{21}N_2O_{10}Na$ (436.350): C, 46.79; H, 4.85; N, 6.42; Na, 5.27. Found: C, 38.11; H, 4.56; N, 5.00; Na, 6.25.
Anal. Calcd for $C_{17}H_{21}N_2O_{10}Na$·NaCl·$2H_2O$ (530.82): C, 38.47; H, 4.75; N, 5.28; Na, 8.66.

The free acid of **10** can be obtained as described above by including 0.1% TFA in the mobile phases during RP-C18 chromatography and the following gradient:

Time	Flow Rate (mL/min)	% MQW†	% CH_3CN†
0	6	100	0
20	6	100	0
40	5	40	60

† containing 0.1% TFA

The acid elutes at ca. 36 min using this gradient over the μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge).

3.2.6.4. Synthesis of Substituted β-PNP-MurNAc Derivatives.

β-*p*-Nitrophenyl-4,6-O-benzylidene-N-acetyl-muramyl-*L*-alanine methyl ester (11) (Bzi-β-PNP-MurNAc-*L*-Ala-CO_2Me)

To a stirring solution of Bzi-β-PNP-MurNAc **8** (100 mg, 0.20 mmol) in distilled DMF (3.4 mL) was added BOP (88 mg, 0.20 mmol), HOBt·H_2O (31 mg, 0.20 mmol) and 4-methyl morpholine (0.033 mL, 0.3 mmol). The solution was stirred for 15 min afterwhich *L*-Ala-OCH_3·HCl (84 mg, 0.60 mmol) and DIPEA (0.104 mL, 0.60 mmol) were added and the solution was stirred for 1 hr. The solvent was removed under vaccum and the residue was repeatedly washed with CH_2Cl_2 and the solvent evaporated. The glutinous residue was suspended in CH_2Cl_2:MeOH (9:1, 1.5-2.0 mL, NOTE: a flocculent precipitate may form which is not the product) and chromatographed over silica gel (2.8 × 10 cm, ≈ 15 g dry mass of gel) using CH_2Cl_2:MeOH (9:1) and gravity flow collecting ≈ 1 mL fractions. Fractions containing the product (21-29, monitored by TLC_2) were pooled, the solvent was evaporated and the product was dried under vacuum yielding 94 mg (0.16 mmol, 80%) of a dense white solid.

R_f (TLC_2, CH_2Cl_2:MeOH, 9:1) 0.72; ^{1}H NMR (200 MHz, DMSO-d_6) δ 8.20 (2H, d, $J_{m,o}$ = 9.2 Hz, ***m*-PNP** × 2), 8.11 (1H, d, $J_{NH,2}$ = 8.9 Hz, **NH**Ac), 7.50 (1H, d, $J_{NH,\alpha}$ = 7.4 Hz, Ala **NH**), 7.37-7.48 (5H, m, **Ph-H** × 5), 7.24 (2H, d, $J_{o,m}$ = 9.2 Hz, ***o*-PNP** × 2), 5.72 (1H, s, Ph-C-**H**), 5.41 (1H, d, $J_{1,2}$ = 8.3 Hz, **H-1**), 4.20-4.35†, 4.06-4.12‡, 3.80* (2H†, m, 1H from Ala **Hα**; 2H‡, m, 1H from lactyl C**H**-CH_3, $J_{CH,Me}$ = 6.7 Hz; 4H*, br s; **H-2**, **H-3**, **H-4**, **H-5**, **H-6** ×2), 3.61 (3H, s, CO_2CH_3), 1.79 (3H, s, NH**Ac**), 1.31 (3H, d, $J_{Me-\alpha}$ = 7.3 Hz, Ala CH_3), 1.24 (3H, d, $J_{Me,CH}$ = 6.7 Hz, lactyl CH-CH_3); ^{13}C NMR (50.3 MHz, DMSO-d_6), δ 172.7, 171.8, 170.1 (lactyl CH-**C**O, NH**C**OCH_3, Ala-CO_2CH_3), 161.7 (**PNP-O-C1**), 142.2 (**PNP-C4-**NO_2), 137.4 (**Ph-C1**), 128.9 (**Ph-C4**), 128.1 (2C), 125.9 (4C) (**Ph-C2** × 2, **Ph-C3** × 2, **PNP-C3** × 2), 116.7 (**PNP-C2** × 2), 100.2, 98.1 (Ph-**C**, **C-1**), 79.6, 78.9, 77.2 (lactyl **C**H-CH_3, **C-4**, **C-3**), 67.5, 65.8 (**C-5**, **C-6**), 54.5 (**C-2**), 52.0 (Ala-**C**α), 47.3 (Ala-CO_2**C**H_3), 22.9 (NHCO**C**H_3), 18.9 (lactyl CH-**C**H_3), 16.9 (Ala-**C**H_3).

Anal. Calcd for $C_{28}H_{33}N_3O_{11}$ (587.584): C, 57.24; H, 5.66; N, 7.15. Found: C, 57.04; H, 5.81; N, 6.95.

β-*p*-Nitrophenyl-N-acetyl-muramyl-*L*-alanine methyl ester (12) (β-PNP-MurNAc-*L*-Ala-CO_2Me)

Anhydrous MeOH (10 mL) was cooled to 0 °C under Ar and freshly distilled acetyl chloride (0.5 mL) was added by syringe. Bzi-β-PNP-MurNAc-*L*-Ala-CO_2Me **11** (87 mg, 0.15 mmol) was added at once and the reaction allowed to warm to room temperature and then stirred for 1 hr (clearing of the solution ensues after 15 min). Reagent grade MeOH (10 mL) was added, the solvent was evaporated and the process was repeated (3 ×). The brown oil was dissolved in a minimal amount of CH_2Cl_2:MeOH (9:1) and chromatographed on silica gel (≈ 6 g dry mass) using the same solvent system and gravity flow collecting 1 mL fractions. Fractions containing the product (8-18, monitored by TCL_2) were pooled and the solvent removed *in vacuo* giving 50 mg of an off-white solid.

A portion of the partially purified product (45 mg, 90% of the total; the remainder was utilized to develop HPLC conditions) was suspended in reagent grade MeOH (1.5 mL) and centrifuged. The sample was applied in two individual runs (≈ 0.8 mL each) to a μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge) using the linear gradient given below with detection at 210 and 300 nm:

Time	Flow Rate (mL/min)	% MQW	% CH_3CN
0	2	80	20
2	6	80	20
40	6	65	35
44	6	5	95
50	6	5	95
54	6	80	20

The eluent over 27 to 31 min was collected and pooled from both runs. The fraction was evaporated to remove CH_3CN and lyophilized yielding 42 mg (0.084 mmol÷90% of total = 0.093 mmol, 62%) of a white solid.

R_f (TLC_2, CH_2Cl_2:MeOH, 9:1) 0.35; ^{1}H NMR (acetone-d_6 (≈ 0.8 g) and $CDCl_3$ (2-3 drops) until dissolved) δ 8.12-8.20 (2H, m, $J_{m,o}$ = 9.3 Hz, ***m*-PNP** × 2), 7.17-7.25 (2H, m, $J_{o,m}$ = 9.2 Hz, ***o*-PNP** × 2), 5.30 (1H, d, $J_{1,2}$ = 8.4 Hz, **H-1**), 4.19-4.41 (2H, 2 overlapping q, $J_{\alpha,Me}$ = 7.3 Hz, $J_{CH,Me}$ = 6.8 Hz, Ala **Hα** (4.35 ppm) and lactyl **CH**-CH_3 (4.28 ppm)), 4.15 (1H, dd, *J*= 9.8 Hz , 8.5 Hz, **H-3** or **H-4**), 3.89 (1H, $J_{6a,6b}$ = 11.3 Hz, **H-6a**), 3.51-3.74 (12H - ≈5H from H_2O at 3.61 ppm ≈7H, **H-2**, **H-3** or **H-4**, **H-5**, **H-6b**, CO_2CH_3 (at 3.67 ppm)), 1.87 (3H, s, **NHAc**), 1.40 (3H, d, $J_{Me,\alpha}$ = 7.3 Hz, Ala CH_3), 1.35 (3H, d, $J_{Me,CH}$ = 6.7 Hz, lactyl CH-CH_3);

^{13}C NMR (50.3 MHz, acetone-d_6 (≈ 0.8 g) and $CDCl_3$ (3 drops)), δ 174.4, 174.0, 172.1 (lactyl CH-**C**O, NH**C**OCH_3, Ala-**C**O_2CH_3), 163.1 (**PNP-O-C1**), 143.2 (**PNP-C4-NO$_2$**), 126.3 (**PNP-C3** × 2), 117.5 (**PNP-C2** × 2), 99.3 (**C-1**), 83.1 (lactyl **C**H-CH_3), 78.4, 77.6 (**C-3**, **C-5**), 70.1 (**C-4**), 61.8 (**C-6**), 55.3 (**C-2**), 52.5 (Ala-**C**α), 48.6 (Ala-CO_2**C**H_3), 23.1 (NHCO**C**H_3), 19.5 (lactyl CH-**C**H_3), 17.2 (Ala-**C**H_3); ESMS: i) exact mass calcd for $C_{21}H_{29}N_3O_{11}$ (M) 499.1802, found 499.39; ii) exact mass calcd for $C_{15}H_{24}N_2O_4$ (M - [$OC_6H_4NO_2$] - [H]) 360.1533, found 360.35. NOTE: fragmentation was dependent on the cone voltage applied during acquisition; the relative intensity of (M) increased with decreasing cone voltage.

Anal. Calcd for $C_{21}H_{29}N_3O_{11}$ (499.475): C, 50.50; H, 5.85; N, 8.41. Found: C, 50.32; H, 5.93; N, 8.26.

***L*-Ala-*D*-*iso*-Glu(αOCH_3)-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala-OCH_3 (Peptide 4)**

Anhydrous MeOH (10 mL) was cooled to 0 °C under Ar and freshly distilled acetyl chloride (0.5 mL) was added by syringe. Peptide **1** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala, 25 mg, 0.035 mmol) was added at once and the reaction allowed to warm to room temperature and then stirred for 2 hr (NOTE: **1** is partially insoluble initially but complete dissolution ensues within 5 min). Reagent grade MeOH (5 mL) was added, the solvent was evaporated and the process was repeated (3 ×). The residue was suspended with sonication in MQW (5 mL) and then lyophilized. The solid was dissolved in reagent grade MeOH (0.5 mL), centrifuged, and was applied to a μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge) using the linear gradient given below with detection at 210 and 265 nm:

Time	Flow Rate (mL/min)	% MQW†	% CH_3CN†
0	6	75	25
5	6	75	25
35	6	45	55

† containing 0.1% TFA

The eluent over 26 to 32 min was collected and the solvent removed *in vacuo* (NOTE: the product precipitates in MQW as the CH_3CN is predominantly evaporated). The product was washed with reagent grade MeOH and then dried under vaccum yielding 25 mg (0.034 mmol, 97%) of a glassy solid.

^{1}H NMR (250 MHz, methanol-d_4) δ 8.13 (0.6H, amide protons, partially exchanged), 7.78†, 7.63‡, 7.27-7.40* (2H†,d; 2H‡,d; 4H*, m; FMOC-Ph-**H** × 8), 4.33-4.52†, 4.15-4.26‡, 4.00*

(5H†, m; 2H‡, m; 1H*, q; α-**H** × 5, FMOC-Ph-C**H**-CH$_3$), 3.70, 3.67 (6H, 2 s, CO_2CH_3 × 2), 3.10 (2H, t, K-**H**ε × 2), 2.86 (0.2H, br s, α-**NH**$_2$ partially exchanged), 2.19-2.35 (3H, m, E-**H**γ × 2, E-**H**β$_1$), 1.75-1.92 (2H, m, E-**H**β$_2$, K-**H**β$_1$), 1.60-1.70 (1H, m, K-**H**β$_2$), 1.46-1.59 (5H, m, K-**H**δ × 2, A_1-**CH**$_3$ (d at 1.53 ppm)), 1.30-1.43 (8H, m, K-**H**γ × 2, A_2-**CH**$_3$ and A_3-**CH**$_3$ (d's at 1.41 and 1.36 ppm)).

β-*p*-Nitrophenyl-4,6-O-benzylidene-N-acetyl-muramyl-*L*-Ala-*D*-iso-Glu(αOCH$_3$)-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala-OCH$_3$ (13)

The coupling of peptide **4** with Bzi-β-PNP-MurNAc **8** was performed with modifications to the methodology described previously by Adam (1992). To a dry mixture of **8** (23.5 mg, 47 μmol), DMAP (5.7 mg, 47 μmol) and HBTU (17.7 mg, 47 μmol) was added CH_2Cl_2:CH_3CN (1:4, 2 mL) and then DIPEA (22 μL, 126 μmol) was added afterwhich the solution cleared. Peptide **4** (23.0 mg, 31 μmol) was suspended with sonication into the same solvent (1 mL) and DIPEA (6 μL, 34 μmol) was added and this suspension was added to the former solution. An additional 1 mL of solvent was used to rinse the flask in which peptide **4** was suspended and then transferred to the reaction (which remains cloudy after addition of the peptide) and the reaction was stirred at 25 °C. The progression of the reaction was monitored by HPLC using the μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge). An aliquot of the reaction (25 μL) was mixed with DMSO (25 μL) giving a clear solution which was centrifuged and applied to the column using the linear gradient below with detection at 215 and 300 nm:

Time	Flow Rate (mL/min)	% MQW†	% CH_3CN†
0	6	70	30
10	6	70	30
50	6	30	70
52	6	5	95
58	6	5	95
60	6	70	30
70	6	70	30

† containing 0.1% TFA

At this scale of chromatography, the chromatograms at each wavelength were very clean and the reactants were well resolved in the gradient with retention times of 29.4 min (peptide **4**), 34.0 min (**8** or CO_2-R activated **8**) and 46.5 min (coupled product **13**). The reactants were identified by the characteristics of the absorption spectra of the individual peaks.

When HPLC analysis indicated complete disappearance of peptide **4** (21 hr), the solvent was evaporated and the residue was triturated with MQW. This was performed by suspending the solid in MQW (5 mL) with sonication and then collecting the solid by centrifugation using two microcentrifuge tubes (1.5 mL) and disposing the supernatant. The pellet in each tube was then repeatedly triturated with MQW (1 mL × 5) and then the pellets were collected and dried under vaccum. If the solid was not dried, then addition of DMSO to solvate the crude material for subsequent preparative HPLC purification would result in a very viscous solution.

Each pellet was suspended in DMSO (0.6 mL), centrifuged and was applied in three individual runs (i.e. 3 × 0.2 mL) to the μBondapak™ C18 Prep-Pak™ cartridge (with the Guard-Pak™ cartridge). The column was eluted using the linear gradient below with detection at 215 and 300 nm:

Time	Flow Rate (mL/min)	% MQW†	% CH_3CN†
0	6	60	40
10	6	60	40
45	6	25	75
47	6	5	95
53	6	5	95
56	6	70	30

† containing 0.1% TFA

The product peak eluting over 35 to 40 min was collected and were pooled from the 6 individual runs (NOTE: there is a small peak which precedes the product which was avoided and the product peak tended to exhibit tailing and only the eluent described was collected). The solvent was removed *in vacuo*, the residue was washed with MeOH and the product dried under vacuum giving 14 mg (11.4 μmol, 37%) of a dense solid. The ^1H NMR of **13** was obtained (see Fig. 3.36 in section 3.3.4) in DMSO-d_6 afterwhich **13** was recovered from the NMR tube (using MeOH as a transfer solvent) and dried under vacuum to remove residual DMSO-d_6.

β-*p*-Nitrophenyl-*N*-acetyl-muramyl-*L*-Ala-*D*-iso-Glu(αOCH$_3$)-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala-OCH$_3$ (14)

To a stirring suspension of **13** (14 mg, 11.4 umol) in anhydrous MeOH (5.0 mL) cooled to 0 °C under Ar was added AcCl (0.25 mL) by syringe afterwhich the reaction was allowed to warm to room temperature (NOTE: the starting material does not dissolve initially but the suspension clears dramatically over 2 hr). The progression of the reaction

was monitored by HPLC using the μBondapak™ C18 Radial-Pak™ cartridge. An aliqout of the reaction (20 μL) was mixed with DMSO (50 μL), centrifuged and 20 μL was applied to the column using the linear gradient below with detection at 215 and 300 nm:

Time	Flow Rate (mL/min)	% MQW†	% CH_3CN†
0	1	70	30
5	1	70	30
30	1	45	55
32	2	5	95
37	2	5	95
39	2	70	30

† containing 0.1% TFA

After 4.5 hr of stirring, HPLC analysis of the reaction indicated essentially complete disappearance of **13** (R_T = 23.5 min) with concomitant production of the debenzylidated product **14** (R_T = 11.8 min). Reagent grade MeOH (2 mL) was added, the solvent was evaporated and the process was repeated (3 x). The slightly beige solid was dissolved in DMSO (0.6 mL), centrifuged and was applied in three individual runs (i.e. 3 x 0.2 mL) to the μBondapak™ C18 Prep-Pak™ cartridge (with the Guard-Pak™ cartridge). The column was eluted using the linear gradient below with detection at 215 and 300 nm:

Time	Flow Rate (mL/min)	% MQW†	% CH_3CN†
0	6	65	35
10	6	65	35
40	6	35	65
42	6	5	95
47	6	5	95
49	6	65	35

† containing 0.1% TFA

The product peak eluting over 28 to 31 min was collected and were pooled from the 3 individual runs (<1 mg of the starting material **13** is recovered and elutes over 40.5 to 42 min). The solvent was removed *in vacuo*, the residue was washed with MeOH and the product dried under vacuum giving 10 mg (8.8 μmol, 77%) of a white solid. The ^1H NMR of **14** was obtained (see Fig. 3.36 in section 3.3.4) in DMSO-d_6 afterwhich **14** was recovered from the NMR tube (using MeOH as a transfer solvent) and dried under vacuum to remove residual DMSO-d_6.

β-*p*-Nitrophenyl-N-acetyl-muramyl-*L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala-*D*-Ala (15)

The simultaneous Fmoc cleavage and dimethyl ester hydrolysis of **14** was performed in the presence of Cs_2CO_3 using methodology developed by Kaestle et al. (1991). To a stirring suspension of compound **14** obtained from above (10 mg, 8.8 μmol) in a solvent mixture of MQW (0.75 mL) and reagent grade MeOH (1.75 mL) was added Cs_2CO_3 (50 mg, 2% w/v). Compound **14** is insoluble initially but the reaction mixture clears with complete solubilization occurring within 3 hr. The progression of the reaction was monitored by HPLC using the μBondapak™ C18 Radial-Pak™ cartridge. An aliqout of the reaction (5 μL) was mixed with MQW containing 0.1% TFA (20 μL), centrifuged and 20 μL was applied to the column using the linear gradient below with detection at 215 and 300 nm:

Time	Flow Rate (mL/min)	% MQW†	% CH_3CN†
0	0.75	90	10
5	0.75	90	10
45	0.75	50	50
50	0.75	5	95
50.1	2	5	95
54	2	5	95
57	2	90	10

† containing 0.1% TFA

The retention times of the reaction products were identified by the characteristics of their absorbtion spectra and were: compound **14** (R_T = 36.3 min); the cleaved Fmoc group (R_T = 44 min); the product **15**, and its mono- and dimethyl ester eluting over 6-9 min. The HPLC analysis indicated that the Fmoc removal from **14** was complete after 3-4 hr however hydrolysis of the methyl esters required an extended reaction time. After 27 hr of stirring at room temperature, the reaction was deemed complete. The solvent was evaporated to remove MeOH and then lyophilized giving a slightly yellow solid.

The solid was suspended in 4 mL of MQW, centrifuged and was applied in four individual runs (i.e. 4 × 1.0 mL) to the μBondapak™ C18 Prep-Pak™ cartridge (with the Guard-Pak™ cartridge). The column was eluted using the linear gradient below with detection at 215 and 300 nm:

Time	Flow Rate (mL/min)	% MQW†	% CH_3CN†
0	6	90	10
10	6	90	10
30	6	80	20
32	6	5	95
37	6	5	95
40	6	90	10

† containing 0.1% TFA

The product peak eluted between 27-29 to 30-32 min (i.e. some small variability was observed) for the four individual runs and the fractions from each run were pooled. The solvent was evaporated to remove CH_3CN and then lyophilized giving 6.2 mg of a flocculent solid.

The purity of **15** was assessed by CZE (conditions are given in the legend to Fig. 3.38) and it was determined that *p*-nitrophenol was present as a contaminant. Previous chromatography using the μBondapak™ C18 Radial-Pak™ cartridge (flow rate: 0.75 mL/min; gradient: 0-20 min, 0-20% CH_3CN (0.1% TFA)) revealed that *p*-nitrophenol eluted prior to **15** with retention times of 13.5 min and 16.6 min respectively. It was possible that i) *p*-nitrophenol was generated during the reaction and that the presence of Cs_2CO_3 in the crude mixture may have resulted in an increased retention of *p*-nitrophenol and hence, its co-elution with **15** during the preparative purification or ii) **15** was purified free of *p*-nitrophenol and that the contaminating *p*-nitrophenol may have been produced from decomposition of **15**.

As such, the 6.2 mg of partially purified material was dissolved in MQW (1.0 mL) and chromatographed over the μBondapak™ C18 Prep-Pak™ cartridge (with the Guard-Pak™ cartridge) employing the same conditions described above. The eluent between 26 to 30 min was collected and was lyophilized without prior evaporation of CH_3CN (the solvent required repeated freezing). The chromatogram also indicated a small peak (R_T = 23.6 min, λ_{max} = 400 nm) that was well resolved from the eluent containing the product. Following the first lyophilization, the material was transferred to a tared vial using MQW (≈ 5 mL) and lyophilized to dryness yielding 5.4 mg (6.1 μmol, 69%) of a white solid. The product was characterized by NMR, MALDI MS and CZE (see Figs. 3.36, 3.37 and 3.38 in section 3.3.4). Compound **15** was stored under Ar in darkness at -20 °C and showed no decomposition over a period of 5 months.

3.2.7. Enzymic Assays with Commercial and Synthetic Substrates

3.2.7.1. Chitooligosaccharides and Nitrophenyl Glycosides

i) Chitooligosaccharides $(GlcNAc)_n$

Stock solutions of $(GlcNAc)_{n=2-6}$ (10 mM in MQW), LaL (1 mg/mL in MQW), HEWL (1 mg/mL in MQW) and 1.0 M KPB, pH 7.0 were prepared under aseptic conditions using filter sterilized buffers and MQW. LaL incubations with $(GlcNAc)_n$ were performed in 50 mM KPB (pH 7.0) at sugar concentrations of 5 mM and an enzyme concentration of 50 μg/mL and were prepared by mixing appropriate amounts of the stock solutions ($(GlcNAc)_n$, 50 μL; LaL, 5 μL; KPB, 5 μL; MQW, 40 μL). As a control, identical incubations with $(GlcNAc)_5$ and $(GlcNAc)_6$ at 1 mM were performed using 50 μg/mL of HEWL ($(GlcNAc)_n$, 10 μL; HEWL, 5 μL; KPB, 5 μL; MQW, 80 μL). In each case appropriate enzyme blanks were also prepared (for LaL blank: $(GlcNAc)_n$, 50 μL; KPB, 5 μL; MQW, 45 μL; for HEWL blank: $(GlcNAc)_n$, 10 μL; KPB, 5 μL; MQW, 85 μL). Samples were incubated at 37 °C for 24 hr. Aliquots (8 μL) of the incubations were filtered (as described in 3.2.2 *ii)*) and 5 μL was chromatographed over the Showdex® OHpak Q-801 gel permeation column. The column was heated at 55 °C and was eluted with MQW at a flow rate of 0.5 mL/min with detection at 195 nm over 40 min. Following the 24 hr period, aliquots from the LaL incubations (1 μL) were removed and the bacteriolytic activity was measured utilizing the turbidimetric assay (2.2.5 *i)*) which indicated that enzyme activity was sustained over the period of incubation.

ii) Nitrophenyl Glycosides

The following nitrophenyl glycosides at 5 mM were incubated with LaL as described above (total volume of 0.5 mL, performed in duplicate): *p*-nitrophenyl-β-D-*N,N'*-diacetylchitobiose, *p*-nitrophenyl-β-D-*N,N',N''*-triacetylchitotriose, *p*-nitrophenyl-β-D-glucose, *p*-nitrophenyl-α-D-glucose, *p*-nitrophenyl-β-D-galactose, and *o*-nitrophenyl-β-D-galactose. After 24 hr, an equal volume of 0.1 N NaOH or 1 M Na_2CO_3 was added, and the absorbance of the solution was immediately measured to detect either released *p*-nitrophenol (λ_{max} 400 nm) or *o*-nitrophenol (λ_{max} 420 nm) against appropriate blank solutions lacking LaL.

*iii) Co-incubation of Peptide **3**, GlcNAc and $(GlcNAc)_5$-PNP with LaL*

The limited solubility of *p*-nitrophenyl-penta-N-acetyl-β-*D*-chitopentaose ($(GlcNAc)_5$-PNP) does not permit its preparation in aqueous solutions at concentrations greater than 1 mM (1.16 mg/mL). The following were prepared under aseptic conditions

using filter sterilized buffers and MQW. $(GlcNAc)_5$-PNP (1.1 mg) was dissolved in a solution consisting of peptide **3** (200 μL from a 100 mM solution in MQW), 1.0 M KPB, pH 7.0 (50 μL) and MQW (750 μL) yielding ≈ 1mM $(GlcNAc)_5$-PNP, 20 mM peptide **3** in 50 mM KPB, pH 7.0 (1.0 mL total volume, solution A; limited quantities of peptide **3** prohibited preparation of a larger volume of solution A). Solution A was divided into two aliquots of 450 μL and 550 μL. To the 450 μL aliquot was added LaL (0.5 mg of a 91.6 % w/w powder) yielding solution B (solution A containing ≈ 1 mg/mL LaL). To the 550 μL aliquot of solution A was added solid GlcNAc (2.4 mg, to 20 mM) yielding solution C. A portion of solution C (100 μL) was removed as an enzyme blank and to the remaining 450 μL of solution C was added LaL to ≈ 1 mg/mL (as described above) yielding solution D. In addition $(GlcNAc)_5$-PNP (1.1 mg), LaL (1.1 mg of the powder) was dissolved in 50 mM KPB, pH 7.0 (1.0 mL) yielding ≈ 1mM $(GlcNAc)_5$-PNP and 1 mg/mL LaL in 50 mM KPB, pH 7.0 (solution E). To summarize, the solutions of interest are:

Solution B: 1mM $(GlcNAc)_5$-PNP, 20 mM peptide **3** and 1 mg/mL LaL (450 μL)

Solution D: 1mM $(GlcNAc)_5$-PNP, 20 mM peptide **3**, 20 mM GlcNAc and 1 mg/mL LaL (450 μL)

Solution E: 1mM $(GlcNAc)_5$-PNP and 1 mg/mL LaL (1.0 mL)

Solution C: 1mM $(GlcNAc)_5$-PNP, 20 mM peptide **3** and 20 mM GlcNAc (100 μL).

Each of the solutions of interest were incubated at 37 °C. Over the course of 6 days, aliquots (50 μL) of the solutions were removed, filtered (as described in 3.2.2 *ii)*) and chromatographed over the Showdex® OHpak Q-801 gel permeation column (only 30 μL aliquots from solution C were chromatographed). The column was heated at 55 °C and was eluted with MQW at a flow rate of 0.5 mL/min with detection at 300 nm over 50 min. The areas under the curves corresponding to $(GlcNAc)_5$-PNP and GlcNAc-PNP (retention times of 17.1 and 37.4 min respectively) were obtained using the integration option available with the Waters 994 detector. In some cases, the nature of the peak corresponding to GlcNAc-PNP was not conducive to detection by the integrator. When this occurred, an expanded version of the chromatogram including the peak was printed and photocopies were generated. The peak was cut from three copies and the average mass of the paper was recorded (M1). Similarly, a region of known area was cut from three copies of the same chromatogram expansion and the average mass of the paper obtained (M2). The area of the GlcNAc-PNP peak was obtained from M1/M2 × area of M2.

Incubations were analyzed by determining the ratio of the area resulting from GlcNAc-PNP to $(GlcNAc)_5$-PNP. The commercial preparation of $(GlcNAc)_5$-PNP was found to contain trace amounts of GlcNAc-PNP. To obtain the initial ratio, a solution of $(GlcNAc)_5$-

PNP (1 mM in 50 mM KPB, pH 7.0) was freshly prepared and an aliquot (100 μL) was chromatographed as described above. The ratio of the two areas measured was 0.0008 ± 0.0005 (averaged from two individual runs).

Again, the incubations were periodically monitored utilizing the turbidimetric assay (2.2.5 *i)*) to measure bacteriolytic activity. A slight reduction in activity was observed between the fourth and sixth day.

3.2.7.2. β-PNP-MurNAc Analogues

i) Analogues ***9***, ***10*** *and* ***12***

9: β-PNP-MurNAc-CO_2CH_3

10: β-PNP-MurNAc-CO_2Na

12: β-PNP-MurNAc-*L*-Ala-OCH_3

Stock solutions of analogues **9**, **10** and **12** (10 mM in MQW; NOTE: analogues **9** and **12** required mild heating and extensive vortexing to dissolve), LaL (8 mg/mL in MQW), HEWL (8 mg/mL in MQW), 1.0 M KPB, pH 7.0 and 0.5 M sodium acetate, pH 5.0 were prepared under aseptic conditions using filter sterilized buffers and MQW. LaL incubations were performed in 50 mM KPB (pH 7.0) while incubations with HEWL were performed in 25 mM sodium acetate (pH 5.0). Compounds were assayed (performed in triplicate) at a concentration of 5 mM and an enzyme concentration of 1.0 mg/mL and were prepared by mixing appropriate amounts of the stock solutions (analogue, 400 μL; LaL or HEWL, 100 μL; KPB or NaOAc, 40 μL; MQW, 260 μL). Appropriate enzyme blanks (performed in duplicate) were also prepared (analogue, 400 μL; KPB or NaOAc, 40 μL; MQW, 360 μL). Samples were incubated at 37 °C for 47-48 hr. Aliquots of the samples (200 or 500 μL) were mixed with 0.1 N NaOH and the absorbance at 400 nm was immediately measured. Prior to adding the base, aliquots of the incubations (1 μL) were removed and LaL activity was measured utilizing the turbidimetric assay (2.2.5 *i)*) and HEWL activity was measured using a suspension of *M. lysodeikticus* cells (diluted in 50 mM KPB, pH 7.0 to an $OD_{450} \approx 0.7$) which indicated that the enzyme activities were sustained over the period of incubation.

ii) Analogue ***15*** *(β-PNP-MurNAc-L-Ala-D-iso-Glu-L-Lys-D-Ala-D-Ala)*

A 10 mM stock solution of analogue **15** was prepared by dissolving **15** (0.46 mg) into MQW (52 μL). A Cahn C-31 Microbalance in the laboratory of Dr. B. Taylor (Dept. of Biology, University of Waterloo) was employed to accurately measure the quantity of **15**.

In addition, stock solutions of LaL (5 mg/mL in MQW), HEWL (10 mg /mL in 50 mM NaOAc, pH 5.0), 0.2 M KPB, pH 7.0 and 50 mM sodium acetate, pH 5.0 were prepared. Incubation of **15** with LaL was achieved by mixing appropriate amounts of the stock solutions (**15**, 15 μL; KPB, 7.5 μL; LaL, 3 μL; MQW, 4.5 μL) to prepare 5 mM **15**, 0.5 mg/mL LaL in 50 mM KPB, pH 7.0 (30 μL total volume) and an enzyme blank (**15**, 15 μL; KPB, 7.5 μL; MQW, 7.5 μL). Incubation of **15** with HEWL was achieved by mixing appropriate amounts of the stock solutions (**15**, 10 μL; NaOAc, 8 μL; HEWL, 2 μL; MQW) to prepare 5 mM **15**, 1.0 mg/mL HEWL in 25 mM NaOAc, pH 5.0 (20 μL total volume) and an enzyme blank (**15**, 5 μL; NaOAc, 5 μL). All samples were incubated at 37 °C for either 24 hr (LaL) or 48 hr (HEWL).

The reactions were analyzed for released *p*-nitrophenol by capillary electrophoresis. The entire enzyme and blank samples from the LaL incubations were filtered (as described in section 3.2.2 *ii)*) while the HEWL samples were mixed with an equal volume of MQW and then filtered. Samples and blanks were subjected to CZE using the 24 cm × 25 μM I.D. coated cartridge at 25 °C and the BioRad Basic Protein Analysis Buffer (0.3 M sodium borate, pH 8.5). Samples were introduced through a 20 psi × sec injection and were electrophoresed under a positive to negative polarity at 12 kV for 10 min. The detector was set to measure the absorbance at 400 nm from 0 to 5 min and 300 nm from 5 to 10 min. Under these conditions, migration times of 2.58 min and 7.40 min were observed for *p*-nitrophenol and **15** respectively.

The sensitivity to detection of *p*-nitrophenol by CZE was estimated as follows. Solutions of PNP (10 μM) in 50 mM KPB (pH 7.0) or 25 mM NaOAc (pH 5.0) were prepared and subjected to CZE under the conditions described above. With a 20 psi × sec injection, the peak height at 400 nm was 0.0011 ± 0.0003 absorbance units. At this wavelength, the baseline noise was less than 0.00005 absorbance units. It is then possible to calculate the minimum concentration of PNP in solution that could be detected. If one considers that a peak at twice the intensity of the baseline noise could be readily visualized then; $0.0011/(2 \times 0.00005) = 11$ and a concentration of one-eleventh of that injected, 10 μM/11 = 0.91 μM or ≈ 1uM can be detected. As such, if 1 μM PNP can be detected and considering that incubations were conducted with **15** at a concentration of 5 mM, then release of PNP from 0.02% (1 μM/5 mM) of **15** would be sufficient.

3.2.8. Chemical Modification

EDC (1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide) and EAC (1-(4-azonia-4,4-dimethylpentyl)-3-ethyl-carbodiimide iodide) were dried over P_2O_5 prior to use. The preparation of MES buffer is described in Appendix A. The following stock solutions were prepared:

(A) 100 mM MES, pH 5.8

(B) 100 and 200 mM EDC or 100 mM EAC in solution A

(C) 0.8 or 8 mg/mL LaL in solution A

In a typical experiment, LaL was incubated at 0.16 or 1.6 mg/mL (20 μL of the appropriate solution C) in the presence of 10-100 mM of the selected modifying agent (10-70 μL of the appropriate solution B) in 100 mM MES, pH 5.8 at room temperature (solution A was added to bring the final volume to 100 μL). Identical controls lacking the modifying agent were also prepared. After the addition of the enzyme solution as the final reagent, timing of the incubations was initiated.

At various time intervals, aliquots from the samples were removed and diluted with stabilizing buffer (Appendix A) that had been dispensed into microcentrifuge tubes ahead of time and maintained on ice. For samples containing 1.6 mg/mL LaL, a 2 μL aliquot was diluted with 998 μL stabilizing buffer while 5 μL aliquots from samples containing 0.16 mg/mL were diluted with 245 μL stabilizing buffer (in each case, a 3.2 ng/μL solution of LaL resulted). Immediately following their dilution, an aliquot from the diluted sample (4 μL) was assayed using the turbidimetric assay as described earlier (3.2.3.2). All assays were conducted using the substrate cell #13 preparation. After thawing, the substrate cells were utilized for no more than 30-40 min, afterwhich they were discarded and a freshly thawed aliquot of substrate cells was utilized. Background lysis of the substrate cells (blank) was determined using only stabilizing buffer (4 μL).

The average activity of the control samples (no modifying agent) was obtained. Residual activity remaining in the samples containing modifying agent were determined as follows:

$$\frac{\text{(activity of specific time point)} - \text{(blank)}}{\text{(average activity of control)} - \text{(blank)}} \times 100\%$$

Protection of LaL from inactivation by $(GlcNAc)_n$ was achieved by inclusion of the required amount of solid $(GlcNAc)_n$ to the desired concentration with the samples.

3.3. RESULTS AND DISCUSSION

The underlying and central theme of the work described in this chapter concerns our investigations towards elucidating the structural requirements of compounds that would firstly act as ligands for LaL and secondly, as substrates. These subjects have received limited attention in previous investigations with LaL and interestingly, with other phage lysozymes. We chose to explore and to identify any of the fundamental interactions between LaL and the two constituent components of the *E. coli* peptidoglycan, namely the saccharide and the peptide, employing both kinetic and biophysical methods. The structural mimicry of β(1→4) oligomers of N-acetylglucosamine to the glycan component prompted a detailed analysis of their interactions with LaL while studies to investigate the peptide component required the synthesis of peptidoglycan peptidyl mimics.

3.3.1. Substrate Cells and the Turbidimetric Assay

3.3.1.1. Chloroform Treatment and Evaluation of Substrate Cells

Activity measurements on LaL were determined turbidimetrically by monitoring its bacteriolytic activity on chloroform-sensitized *E. coli* MG1655 cells (substrate cells). As discussed earlier, sensitization of the bacterial cells is required to weaken the integrity of the outer membrane such that an exogenous source of LaL has access to the peptidoglycan. Initially, substrate cells were prepared following the somewhat limited description outlined with the first reported purification of LaL (Black & Hogness, 1969a). This involved the suspension of exponentially grown cells in chloroform saturated substrate cell buffer (5 mM EDTA, 60 mM KPB, pH 7.0) at 25 °C, followed by immediate chilling to 0 °C and centrifugation of the cells. However, many attempts with this methodology culminated with inconsistent results and cells that were only weakly sensitive to lysis by LaL and/or not favorable for use in kinetic measurements. It became evident that an alternative protocol was required.

Several important modifications were implemented (detailed in 3.2.3.1) which have resulted in greater success. Careful control and monitoring of the chloroform treatment was imperative. Briefly, *E. coli* cells were grown and harvested at an $OD_{600} \approx 0.8$. The cells were then thoroughly suspended in chloroform-saturated substrate cell buffer and were maintained at room temperature throughout the chloroform treatment. It is believed that some of the inconsistencies described above resulted from inadequate saturation of the buffer, in which cooling the solution would lessen the extent of saturation. Others have also adopted a similar strategy (Koteswara Rao & Burma, 1971; Jensen & Kleppe, 1972; Caldentey & Bamford, 1992). The response of the cells was then monitored with the

addition of an amount of LaL which is known by experience, to cause essentially instantaneous (< 15 s) clearing of optimally sensitized cells. One such progression is shown in Fig. 3.6. As is illustrated, extended duration of treatment typically rendered the cells more prone to lysis. Treatment was terminated when the absorbance of the cell suspension would decrease from ≈ 0.7 and level to ≈ 0.2 in less than 2 min (curve 3).

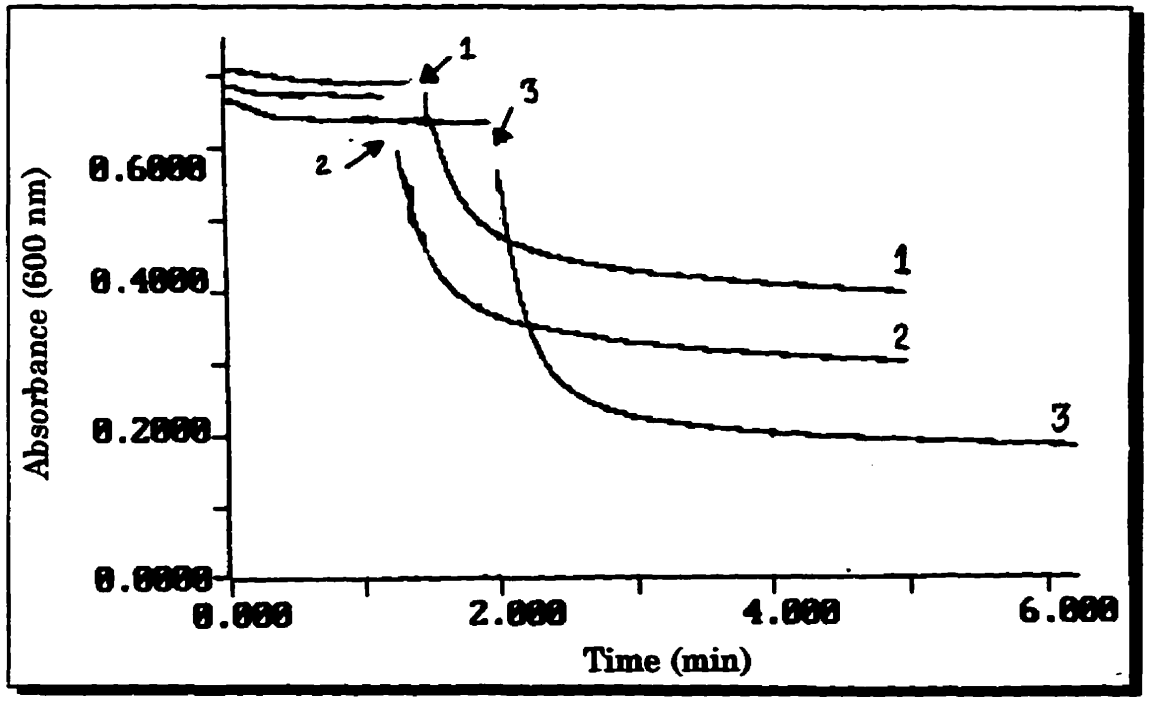


Figure 3.6. Chloroform treatment of *E. coli* MG1655. The curves indicate the propensity to clearing of the cells to LaL following treatment with chloroform for (1) 5 min, (2) 10 min (3) and 18 min.

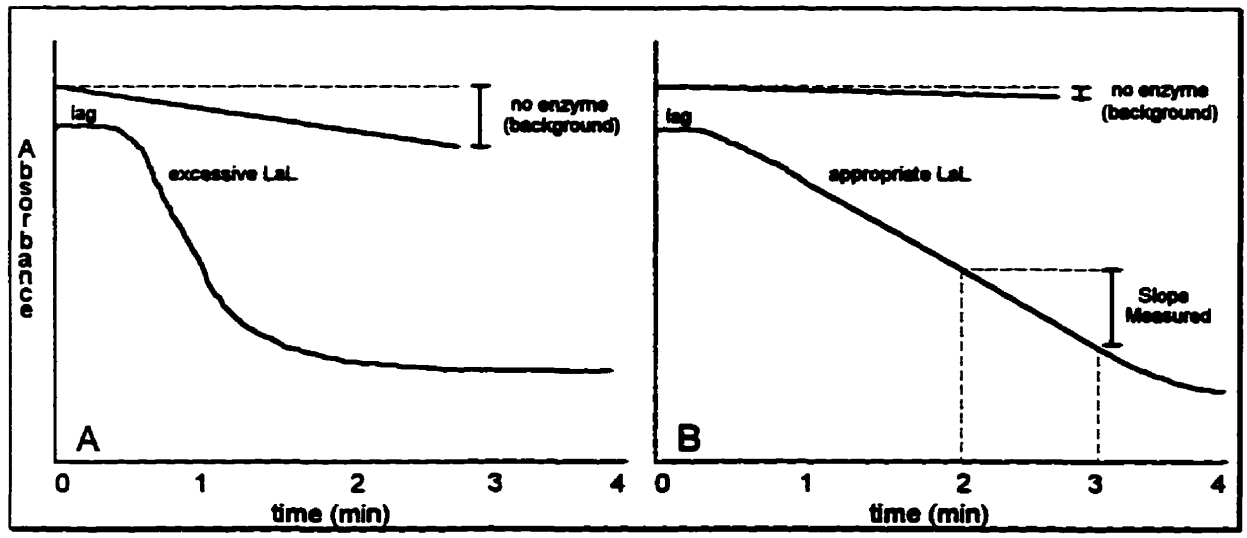
With some attempts, the cells did not respond adequately (as demonstrated in Fig. 3.6) to this initial treatment. In these cases, an additional amount of chloroform would be added to the cells suspended in the original chloroform-saturated buffer and the progression of treatment again monitored until the desired response was observed. It was important to cool the cells and remove as much of the residual chloroform after the treatment. This residual chloroform would be concentrated with the cell pellet during the centrifugation step required to collect the cells and would adversely affect them. The cells would then be washed and frozen as concentrated aliquots.

Each preparation of substrate cells would then be evaluated as to their potential to serve for purposes of activity measurements. To this end, substrate cells were required to

possess certain desirable characteristics. The first of these was the facile dilution of the concentrated substrate cells into an assay buffer. The resulting suspension should have the appearance of a customary *E. coli* culture and should lack any particulate matter (clumps of cells) which, if present, will obstruct the transmission of light during absorbance measurements. Meticulous and careful suspension of cell pellets before and after chloroform treatment is quintessential for reducing these complications. At times, it was found that cells which responded well to chloroform treatment, and were frozen and then thawed for evaluation, would not suspend easily into buffer or would undergo spontaneous lysis. This may have been a consequence of further wearing of the cell wall from the freezing and thawing process which had already been weakened from the exposure to chloroform.

If the first criterion is met, then it was important that the substrate cells demonstrated a usable region of linear response to LaL in addition to a minimal degree of spontaneous background lysis. These can be better appreciated with the following description of the turbidimetric assay. The observed activity is the lysis of the substrate cells resulting from the action of LaL on the peptidoglycan. However, before activity can be observed, a sufficient degree of degradation to the peptidoglycan surrounding each cell must be attained in order for the cell to lyse. As such, there exists a lag period in which LaL acts upon the peptidoglycan yet the cells remain "intact" and do not lyse until the peptidoglycan structure has been sufficiently weakened. Furthermore, the process of cell lysis will generate cellular debris (e.g. vesicles) which will contribute to increasing the absorbance of the solution. Therefore, the absorbance will decrease linearly over a period of time before the accumulation of cellular debris progressively slows this rate and the curve no longer remains linear.

The curves in Scheme 3.1 demonstrate unfavorable (A) and favorable (B) responses of substrate cells. If excessive LaL is added (curve A), then the balance between lysis and accumulation of cellular debris occurs too rapidly in order to obtain a descriptive and accurate measure of activity. In this situation, the linear portion is too small and the time interval in which linearity exists will be variable, thereby impeding the ability to obtain reproducible and consistent measures of activity. This is extremely important if a comparison in activity between, for example, LaL in the absence and presence of an inhibitor was of interest. Instead, an amount of LaL would be added (curve B) such that the curve would possess a linear portion extending for approximately 3 minutes before linearity was lost. As will be discussed below, each substrate cell preparation was unique with respect to this "appropriate" amount. Activity was always expressed as the slope in the curve between 2 and 3 minutes, not only to standardize measurements but also to



Scheme 3.1.

permit sufficient time for the passing of the lag period during which the absorbance did not decrease.

The background lysis in Scheme 3.1 (A) is too high while in (B) the background is acceptable (a decrease between 0.005 and 0.01 au/min was the upper limit of acceptability). Minimal background lysis was essential for kinetic measurements for the following reasons. There existed only a defined range of activity that could be expressed from the assay. The upper limit of measurable activity would be the maximal rate of decrease which maintained linearity between the chosen interval of 2 and 3 minutes (this rate would reflect both the spontaneous and enzyme induced lysis). Typically, a 0.1 au/min decrease was this maximal rate. The lower limit would be the background lysis of the cells. If, for example, the background lysis was 0.02 au/min, which is unfavorable, then activity could only be measured to a level of 0.2 of the maximum. For example, if 100% activity is assumed for LaL, then inhibition could only be measured to a level of at most 20% residual activity, as it would be inappropriate to measure an activity that was of the same magnitude as the background lysis.

The turbidimetric assay utilizing the substrate cells is presently the only feasible method available to determine quantitative measures of LaL activity. Although extensive efforts have been directed towards establishing conditions to prepare favourable substrate cells, the methodology was not successful in all attempts. Quite often, substrate cells would be prepared under the conditions developed, yet the cells would not behave as desired. Several attempts were often required before an acceptable lot was obtained. It is suggested from experience, that if confidence is to be guaranteed with activity

measurements, then the substrate cells should be prepared and thoroughly evaluated for the desired characteristics and only then, used for kinetic purposes. Substrate cell preparations of lesser quality can still be employed for qualitative activity measurements, for example during purification of the enzyme.

Despite the shortcomings associated with the preparation of the substrate cells, the turbidimetric assay was nonetheless very effective in measuring LaL activity. It was also necessary to characterize the response of each lot of substrate cells to LaL concentration. To this end, it was important to establish the maximum concentration of LaL that could be used with each lot of substrate cells, or the "appropriate" amount described earlier that resulted with the maximal desired linear response. Solution samples of LaL were prepared in stabilization buffer (0.5 M KCl, 100 mM KPB, pH 7.0, 20% glycerol) which was required to maintain activity at dilute enzyme concentrations (discussed previously in 2.3.5.1). To minimize the delivery of salt from the stabilizing buffer into the assay mixture the concentration of LaL was adjusted in stabilizing buffer such that a 5 μL aliquot would deliver this "appropriate" amount. Then, aliquots of this solution (1-5^{f} μL), were added to a suspension of the substrate cells (total volume of 750 μL) and the absorbance monitored (NOTE: The addition of 1-5^{f} μL of enzyme sample in stabilizing buffer will add 0.7-3.3 mM KCl and 0.1-0.7 mM KPB to the assay mixture. This is expected to result with some inhibition of lysis of the cells due to ionic strength effects. A discussion on the effects of ionic strength on the assay is presented in section 3.3.1.2).

The measurements on two different lots of substrate cells (SC#9 and SC#13) and their response to varying amounts of LaL are illustrated in Fig. 3.7. These are the actual assay curves and have been included so the reader is aware of the type of response typically observed during an assay. It can be seen from the activity curves that both the criteria of linearity and low background lysis have been met. As was described, to avoid the inherent lag period, the rate of decrease in absorbance from the curves was obtained from the slope between 2 and 3 minutes. Figure 3.8 demonstrates the quantitative relationship between this activity and enzyme concentration for four separate preparations of substrate cells. In each case, it can be seen that a very reasonable linear response exists between enzyme concentration and the rate of decrease in absorbance. The results are encouraging and offers credence to the turbidimetric assay as a reliable indicator of LaL activity. The turbidimetric assay was also quite sensitive requiring only nanogram quantities of enzyme. With two of the substrate cell preparations (SC#9 and 13) as little as 1 ng of LaL could be readily detected.

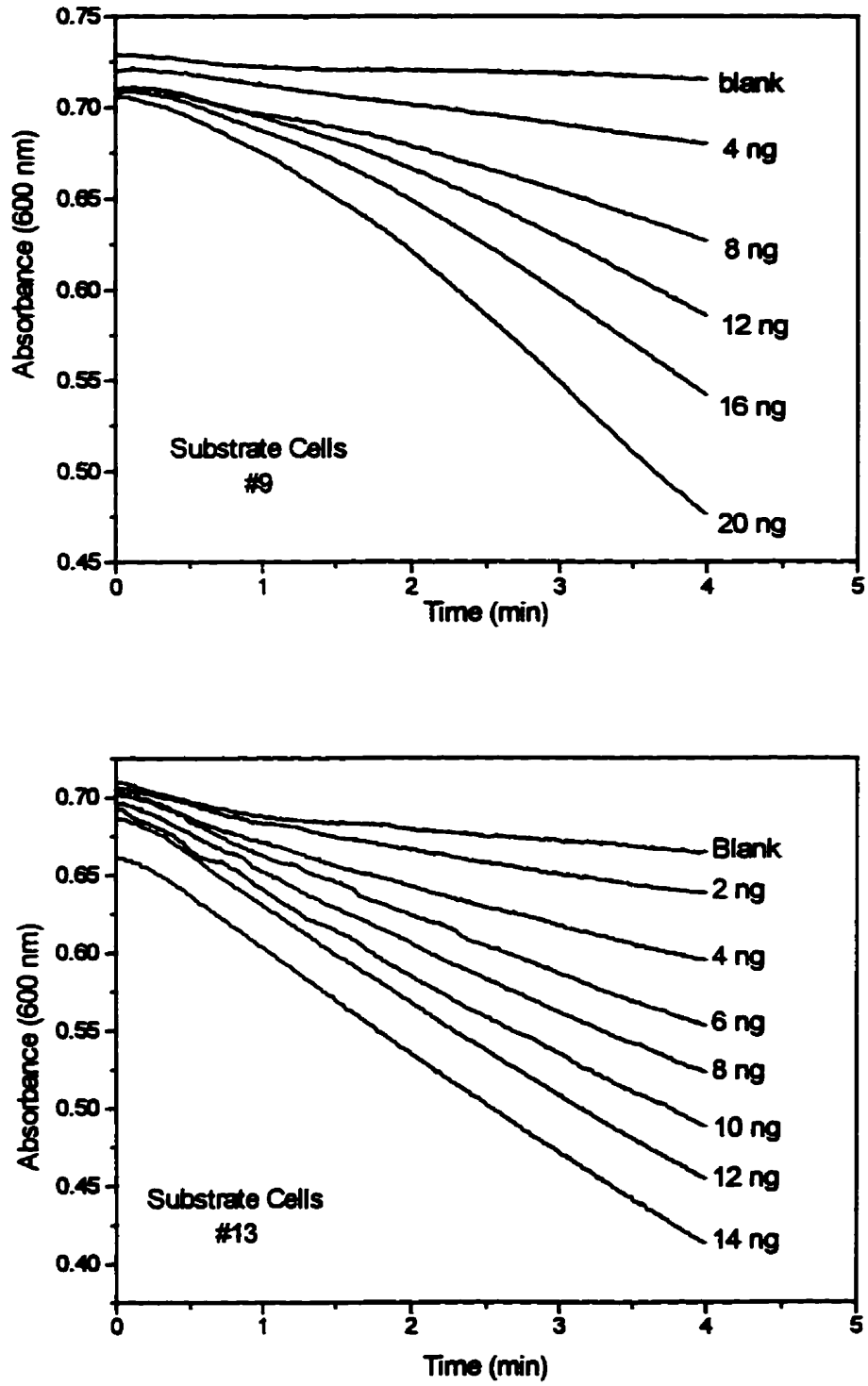


Figure 3.7. Response of substrate cells to LaL concentration (1). Shown are the activity curves for two different substrate cell preparations. The substrate cells were diluted in 50 mM KPB (pH 7.0, 750 µL total volume) and appropriate amounts of LaL were added (as indicated) to initiate cell lysis and the decrease in turbidity was monitored.

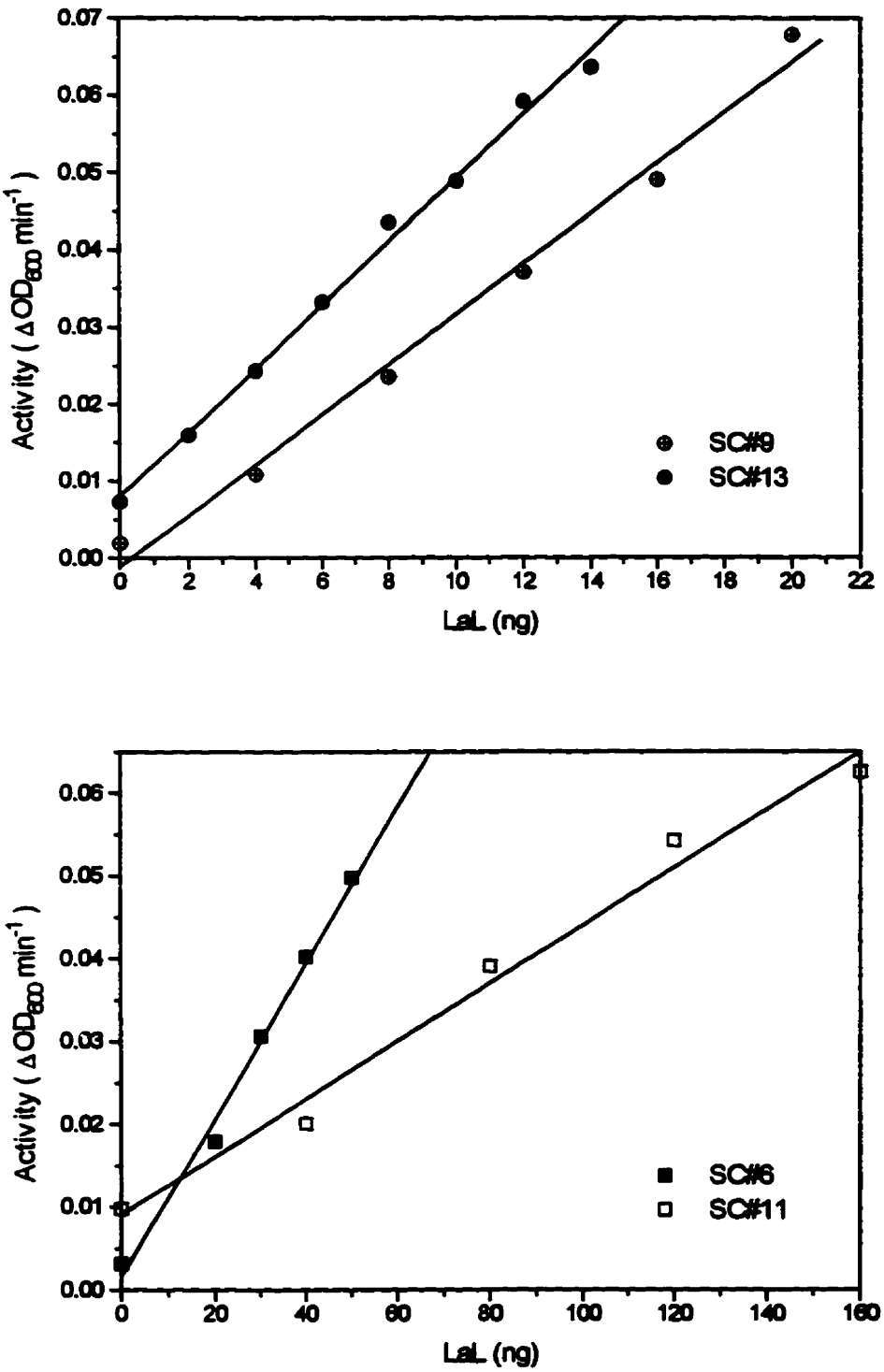


Figure 3.8. Response of substrate cells to LaL concentration (2). Shown are the quantitative relationships between activity and amount of LaL for four different substrate cell preparations. Activity is expressed as the rate of decrease (slope) between 2 and 3 minutes. The top panel represents the activity for the assays shown in Fig. 3.7.

Table 3.6. Comparison of the response of different substrate cell preparations.

Substrate cell Preparation	Response † (× 10^4)	Relative Response	Amount of LaL Required for 1 unit of activity (ng)‡
#13	41.4	1.000	2.42
#9	32.7	0.790	3.06
#6	9.5	0.229	10.53
#11	3.5	0.085	28.57

† Taken from the slopes of the curves from Fig. 3.8. It is the decrease in OD_{600}/min (activity) per 1.0 ng of LaL.
‡ Defining a unit as that amount of LaL resulting in a decrease in OD_{600} of 0.01/min (between 2 and 3 min) using the turbidimetric assay.

There exists obvious differences in the responses of the different substrate cell preparations to LaL concentration (Fig. 3.8). The activities obtained with SC#9 and SC#13 are comparable, yet their susceptibility to lysis by LaL are approximately 5 and 10 times that of SC#6 and SC#11 respectively. A direct comparison of the behavior of these substrate cells is given in Table 3.6. Therefore, the preparation of substrate cells is accompanied with a "batch to batch" variability. This type of variation has also been described for the turbidimetric assay of HEWL using *M. lysodeikticus* cells as the substrate (Shugar, 1952) which have prompted others to report on the inherent deviations of commercial preparations of these cells (Gorin et al., 1971). As was suggested earlier (2.3.1.3) this variability is, therefore, not permissive to the definition of an absolute unit for activity. If arbitrarily designated, then each preparation of substrate cells are found to be unique with respect to the amount of LaL required to produce the activity described as a unit (Table 3.6). For these reasons, any comparative studies on enzyme activity must be performed with the same lot of substrate cells. All kinetic measurements (inhibition, chemical modification, ionic strength effects) with LaL were performed utilizing preparations #6, #9 and #13. Preparation #11 is an example of cells that would be reserved for qualitative measures of activity.

3.3.1.2. Ionic Strength Effects on the Turbidimetric Assay

Lysis of the bacterial substrate cells is in essence an indirect measure of LaL activity. The actual cleavage of the peptidoglycan glycosidic bonds is itself not directly measured. Instead, the consequence of peptidoglycan degradation, cell lysis, serves as the indicator of activity. Because cell lysis is a consequence of the differences between the internal cellular osmotic pressure and the osmotic pressure of the environment, an

alteration in this pressure differential will affect cell lysis and, hence, measures of activity. Bacterial cells in a hypertonic or isotonic medium will not osmotically lyse (although plasmolysis or dehydration of cells will result in a hypertonic environment) while cells in a medium low in solute, will be more disposed to osmotic lysis.

Ionic strength effects were investigated by exposing substrate cells to LaL in buffers of varying concentration. Figure 3.9 (top panels) demonstrates the effect of potassium phosphate (A) and Tris (B) buffer concentration on the turbidimetric assay. In both buffer systems, cell lysis is effectively reduced with increasing buffer concentration. The data are also graphically presented in terms of the standard measure of activity (Fig. 3.9, bottom panel). Since phosphate is a polyprotic buffer, the ionic strength at a given concentration is greater than for the monoprotic Tris buffer and the reduction in lysis or activity is more prevalent for the phosphate system. Similar effects of ionic strength are observed with the turbidimetric assay when potassium chloride or ammonium sulfate are included with the assay buffer (Fig. 3.10). Again, the rate of cell lysis (top panels) and the measured activity (bottom panel) decreases with increasing concentrations of the salts. The effect is stronger with ammonium sulfate since the ionic species of this salt will contribute a greater increase in ionic strength to the solution than an equal concentration of potassium chloride.

The observed inhibitory influence of low concentrations of ammonium sulfate bear some practical importance. It was suggested earlier (refer to section 2.2.5) that when fractions are obtained during the purification of LaL from columns containing high concentrations of salt, smaller aliquots should be employed for activity determinations. This is especially noteworthy when LaL is chromatographed over Phenyl-Superose and is regularly obtained from this column in buffer containing $\approx$ 1 M ammonium sulfate. For example, if a 50 μL aliquot of such a fraction was assayed in the typical manner then the assay mixture will contain approximately 60 mM ammonium sulfate, which will dramatically influence the measurement of activity (Fig. 3.10). This should be kept in mind for future reference if the purification of LaL containing amino acid analogues or mutant forms of the enzyme, which may have diminished activity, is to be performed.

Although the effects of ionic strength have been reasonably explained in terms of the influence that it will have on the lysis of the bacterial cells, the phenomena may be somewhat more elaborate. It has been generally accepted (Katchalsky et al., 1953; Davies et al, 1969; Maurel & Douzou, 1976) that the cell walls of bacteria like *E. coli* and *M. lysodeikticus* are predominantly negatively charged, and surface charge measurements on

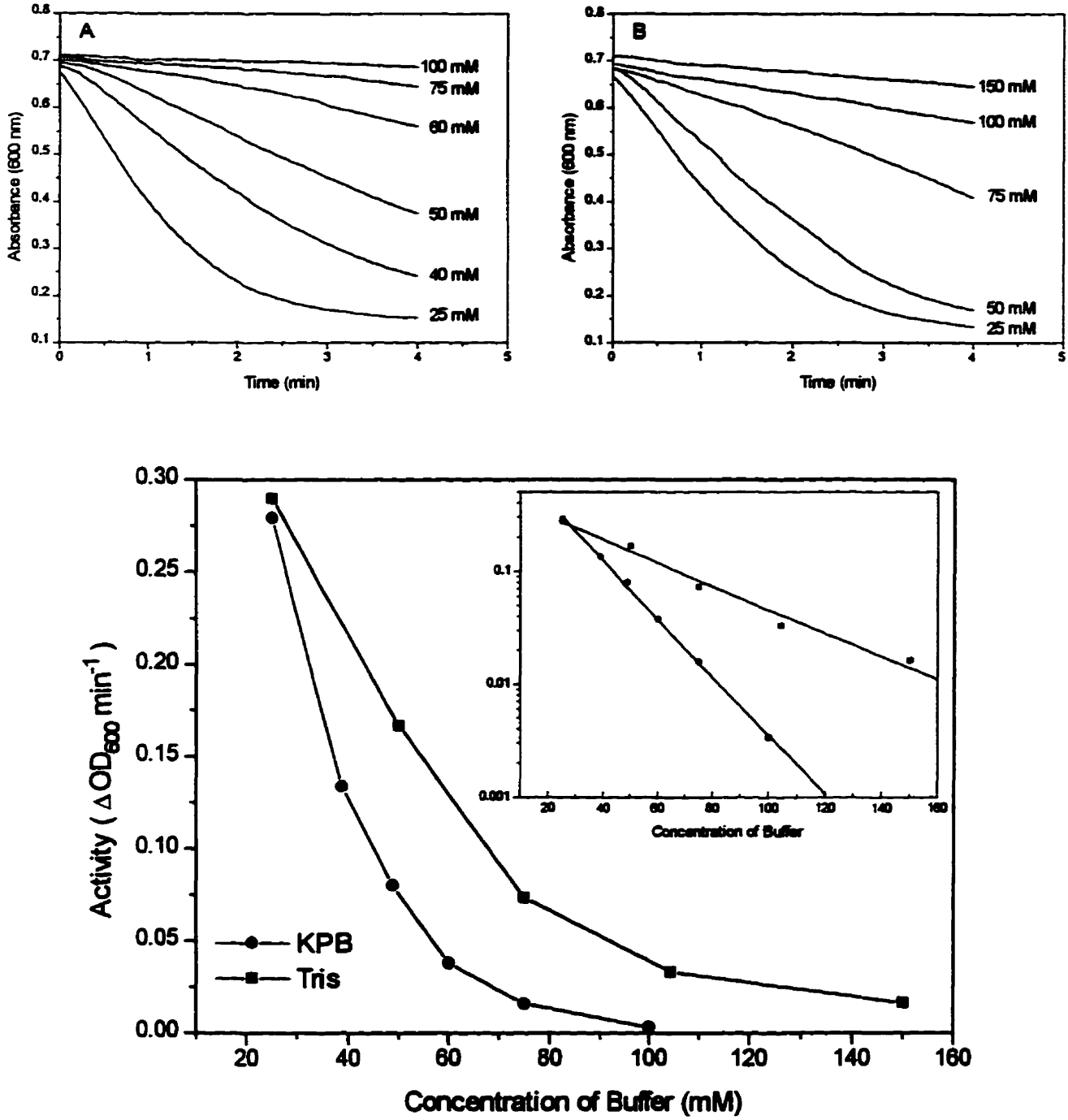
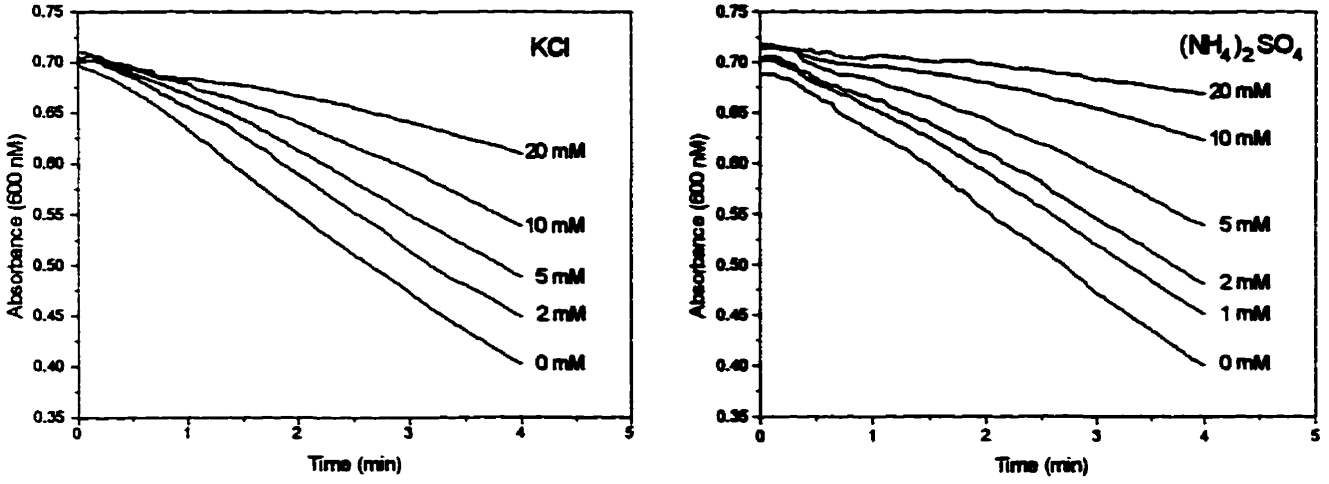


Figure 3.9. Top Panels: Effect of potassium phosphate, pH 7.0 (A) and Tris, pH 7.0 (B) buffer concentration on the turbidimetric assay. Substrate cells #9 were prepared in the two buffers at the concentrations indicated and 10 ng of LaL was added to initiate lysis. Bottom Panel: Effect of buffer concentration presented in terms of measured activity. The activity was taken from the linear portion of the curves shown in the top panels. Inset: semi-logarithmic plot of the data.



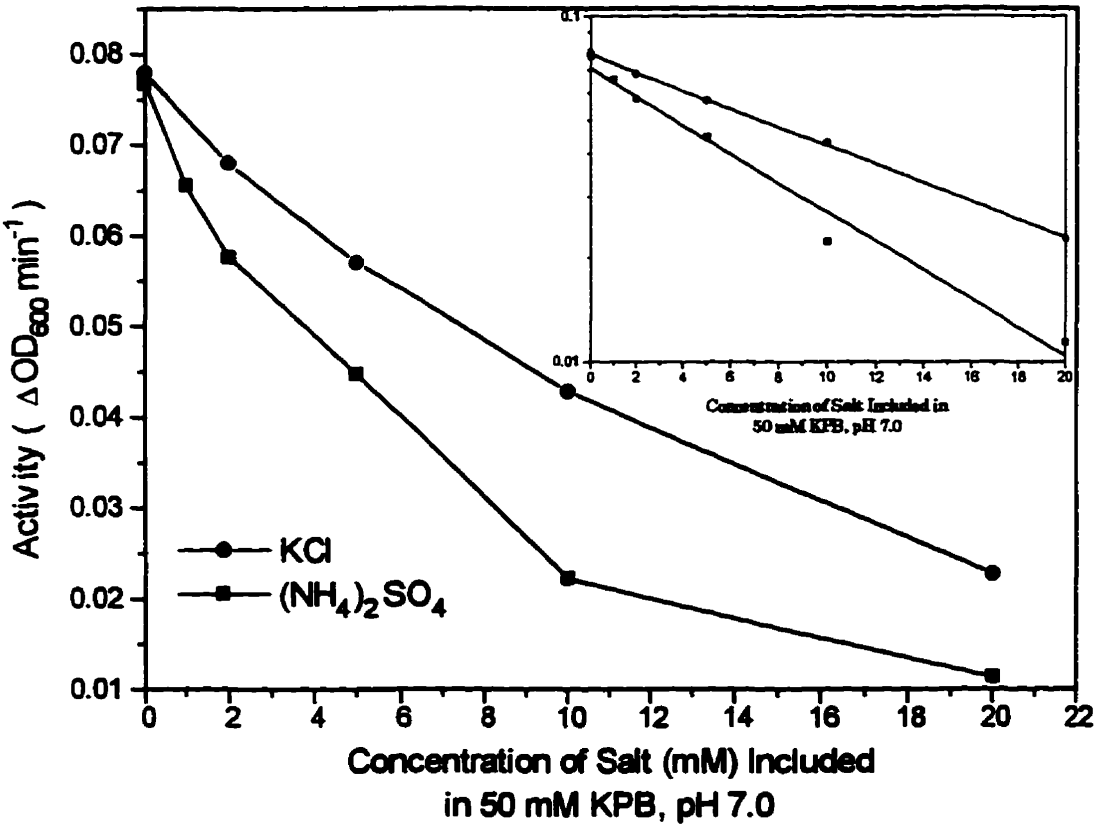
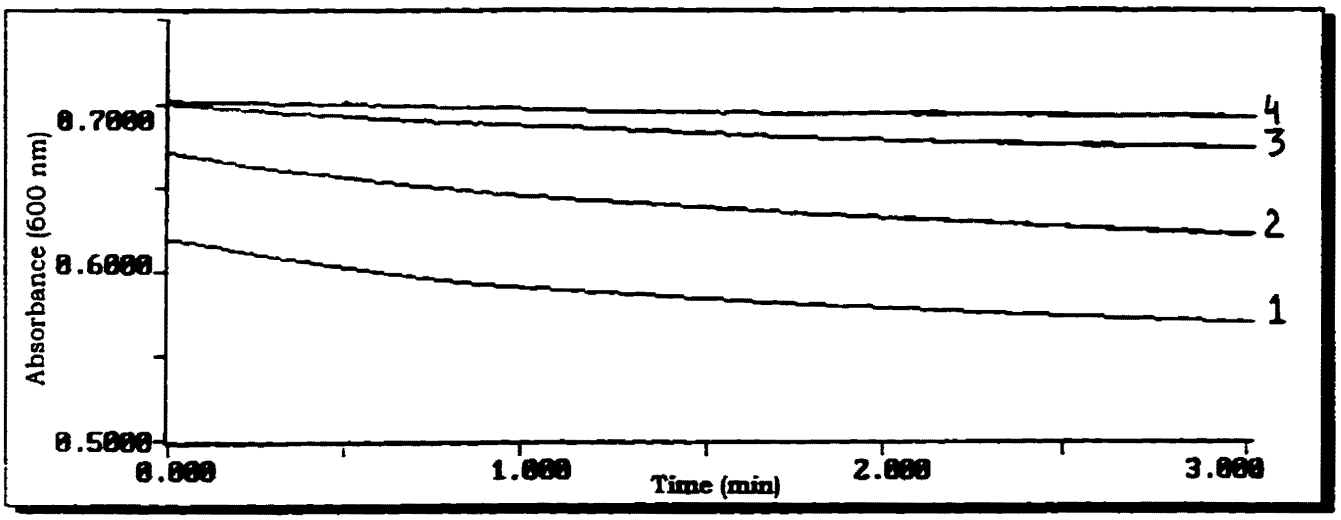


Figure 3.10. Top Panels: Effect of KCl and $(NH_4)_2SO_4$ concentration on the turbidimetric assay. Substrate cells #9 were prepared in 50 mM KPB, pH 7.0 with the additional salt as indicated and 10 ng of LaL was added to initiate lysis.
Bottom Panel: Effect of salt concentration in terms of measured activity. The activity was taken from the linear portion of the curves shown in the top panels. Inset: semi-logarithmic plot of the data.

M. lysodeikticus have experimentally substantiated this net negative charge density (Price & Pethig, 1986). In a detailed kinetic study of the lytic action of HEWL towards *M. lysodeikticus*, Maurel and Douzou (1976) have elaborated upon earlier efforts (Davies et al., 1969; Yong Chang & Carr, 1971) and concluded that the pH and ionic strength dependencies of the lytic activity could be explained by taking into account the electrostatic potential associated with the negatively charged cell wall. The results are rather intricate but are essentially two-fold; i) the electrostatic potential of the cell wall influences the physio-chemical properties of its microenvironment and ii) the basic nature of HEWL (and most lysozymes p$I \approx$ 10-11) will result with its electrostatic attraction to the cell wall. The first point implies that the pH near the cell wall, regardless of the pH of the environment, must always be of the order near or below the pKa of the Glu35 (≈6.1-6.7), the general acid involved in catalysis (the required protonation of Glu35 cannot be expected at pH >8 or 9). Secondly, at higher pH values and particularly at low ionic strength, attraction between the cell wall and lysozyme is stronger, K_m(app) decreases with decreasing ionic strength and maximal *lytic* activity occurs at pH values as high as 9 or 10. However, as ionic strength is increased, the electrostatic potential is decreased, K_m(app) increases, and the pH optimum is shifted to lower pH. This can also be appreciated with the fact that maximal activity (not lytic activity) with small neutral substrates such as $(GlcNAc)_n$ is observed at pH 5.2 and is independent of ionic strength (Banerjee et al., 1973).

It is also of interest that HEWL is inactive towards *M. lysodeikticus* in distilled water and requires activation with low salt concentrations (Yong Chang & Carr, 1971). At extremely low ionic strengths, it has also been suggested that a repulsive force may exist between the cell wall and HEWL (Price & Pethig, 1986). These authors propose that a local region of negative charge on the enzyme molecule, which was associated with the ionized Asp52, is responsible for this repulsion and might act to assist in the release of the enzyme from the cell wall after lysis. There also appears to be a requirement of a minimal amount of salt to maintain an active conformation of HEWL (Douzou & Petsko, 1984). The effect of salt concentration on the stability of HEWL has been extended by NMR studies to show how salt influences the p*K*a of acidic residues (Abe et al., 1995).

In our system the effect of increasing ionic strength on the turbidimetric assay is believed to result, at least in part, from the inhibition of bacterial lysis. Nonetheless, the effects may be more specific in that electrostatic attractions between LaL and the *E. coli* cells may be affected. We have also shown that the spontaneous background lysis is dependent on the concentration of buffer into which the cells were diluted. As demonstrated in Fig. 3.11, the amount of clearing of the substrate cells is dependent on



Buffer	Curve	ΔOD_{600} between 2 and 3 min ($\times 10^4$)	Relative Decrease
MQW	1	95 ± 3	1.4
25 mM KPB, pH 7.0	2	104 ± 5	1.6
50 mM KPB, pH 7.0	3	67 ± 15	1.0
100 mM KPB, pH 7.0	4	34 ± 14	0.5

Figure 3.11. Spontaneous lysis of *E. coli* substrate cells. Substrate cells #13 (40 μL) were diluted into the indicated buffers (710 μL) and the decrease in absorbance was monitored (performed in triplicate). Representative curves are illustrated above and the average ± s.d. of the measured rates are tabulated.

the strength of the buffer into which they are diluted (as would be performed in a typical assay). The standard buffer employed in the turbidimetric assay was 50 mM KPB, pH 7.0. Compared to this buffer, the relative lysis of the cells is reduced in 100 mM KPB yet is increased when measured in MQW or 25 mM KPB (Fig. 3.11, see Table). Notice as well how the initial absorbance of the cells in MQW or 25 mM KPB (curves 1 and 2) is lower than those observed for 50 and 100 mM KPB (curves 3 and 4). This demonstrates the greater degree of initial spontaneous lysis which occurred during the time between dilution of the cells and initiation of the absorbance measurements (≈ 5 s). These results suggests that the chloroform-treated cells are indeed more sensitive to osmotic lysis at lower ionic strengths.

It is of some interest that the influence of buffer composition on bacterial lysis has not been taken into consideration in describing the lytic activity of HEWL. Furthermore,

the salt induced inhibition of the lytic activity of protein P15 (a muramidase from phage PRD1) on chloroform treated *E. coli* cells received no attention as to the effect of the salts themselves on cell lysis (Caldentey et al., 1994). Regardless, when a substrate as complex as a bacterial cell is used to measure an enzyme's activity, it is to be expected that descriptions of the factors affecting the enzyme activity are themselves complex. It is for this reason that there exists a strong interest in the development of well defined substrates for use in detailed kinetic measurements with lysozymes.

3.3.2. Interactions of Chitooligosaccharides with LaL

3.3.2.1. Kinetic Observations

Our preliminary efforts were to establish whether oligomers of N-acetyl-D-glucosamine would serve as substrates for LaL. Previously, $(GlcNAc)_3$ and $(GlcNAc)_5$ were concluded not to be substrates for LaL, based on the detection of the assumed increase in reducing groups that would result from hydrolysis (Black & Hogness, 1969a). However, reducing group analysis would not detect products containing 1,6-anhydro linkages, and these studies could therefore have been in error. We have extended this analysis with $(GlcNAc)_n$ (n = 2-6) by using gel permeation HPLC to monitor product formation. In each case, 5 mM of the saccharide was incubated with LaL for 24 hr. No conversion of $(GlcNAc)_n$ into correspondingly smaller saccharides ($(GlcNAc)_x$, x<n) was observed. The analysis of the incubations of LaL with $(GlcNAc)_5$ and $(GlcNAc)_6$ are illustrated in Fig. 3.12 A. Similar chromatograms were obtained for LaL incubated with $(GlcNAc)_{2\text{-}4}$ in which only the original saccharide was observed. Although the chromatograms are presented at full scale, the ordinate axis was expanded to the level of baseline noise for observation. In parallel experiments with HEWL performed under the same conditions, both $(GlcNAc)_5$ and $(GlcNAc)_6$ at concentrations of 1 mM were clearly and essentially completely hydrolysed into expected products (Fig. 3.12 B).

The activity of LaL on some simple nitrophenyl glycosides (*p*-nitrophenyl-β-D-*N,N'*-diacetylchitobiose, *p*-nitrophenyl-β-D-*N,N',N''*-triacetylchitotriose, *p*-nitrophenyl-β-D-glucose, *p*-nitrophenyl-α-D-glucose, *p*-nitrophenyl-β-D-galactose, and *o*-nitrophenyl-β-D-galactose) was examined by the detection of released nitrophenol. In each case, nitrophenol was not detected and therefore, none of these compounds appear to serve as substrates for LaL.

Although LaL did not accept oligomers of GlcNAc as substrates, we have found that the bacteriolytic activity of LaL was inhibited by these chitooligosaccharides. Inhibition was determined with the turbidimetric assay under a specified set of experimental conditions to reduce as best as possible, any factors which might result with improper

Figure 3.12. Action of LaL and HEWL on $(GlcNAc)_5$ and $(GlcNAc)_6$.

Chitooligosaccharides were incubated with LaL or HEWL (50 μg/mL) in 50 mM KPB, pH 7.0 at 37 °C for 24 hr. Aliquots (5 μL) from the individual incubations were subjected to chromatography over the Showdex® OHpak Q-801 (8 mm × 50 cm) gel permeation column heated at 55 °C using MQW and a flow rate of 0.5 mL/min with detection at 195 nm.

The retention times of the individual saccharides under the chromatographic conditions outlined above are:

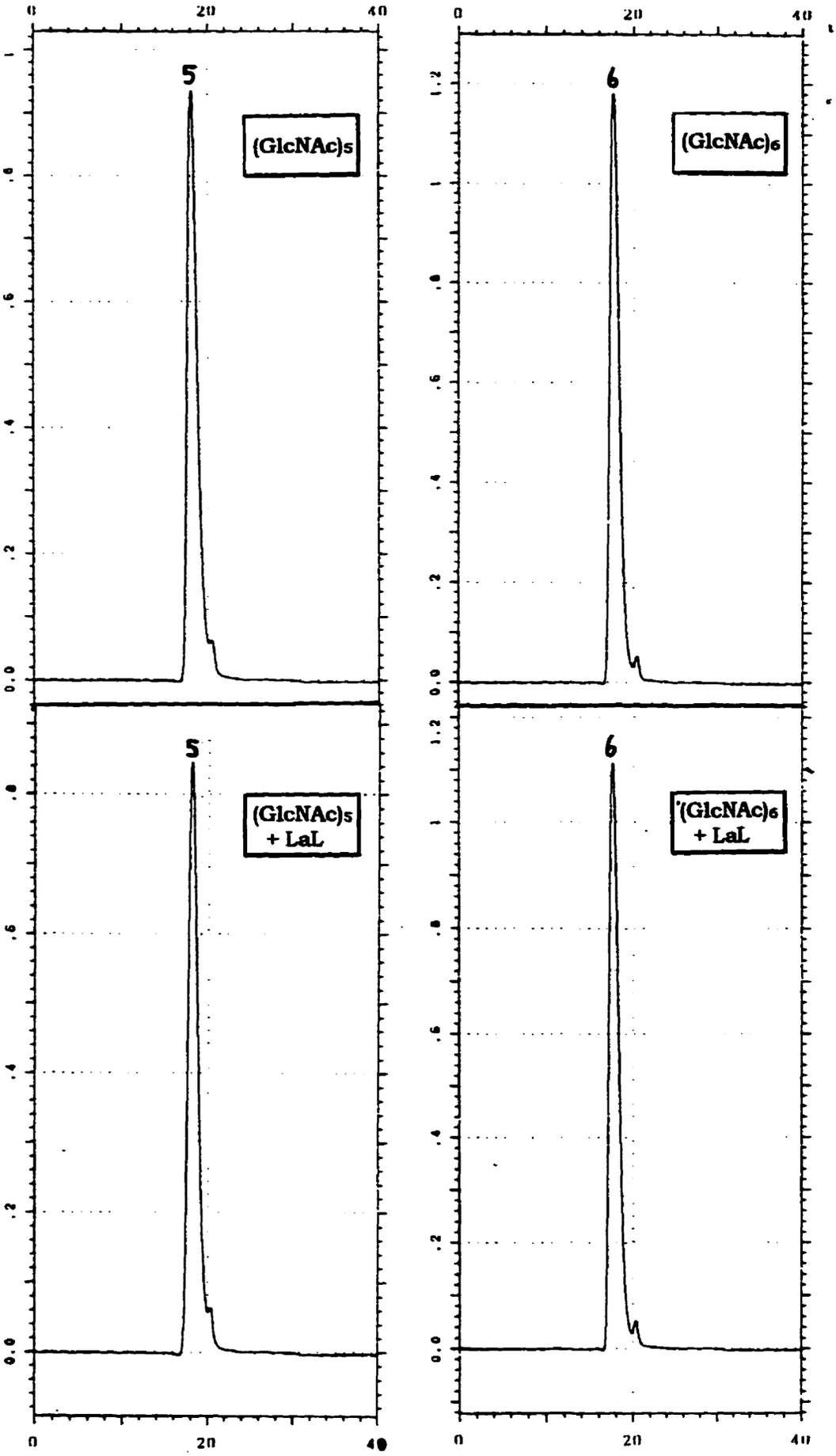
$(GlcNAc)_n$	Time (± 0.05 min)
n = 1	26.7
n = 2	22.5
n = 3	20.1
n = 4	18.8
n = 5	18.1
n = 6	17.7

(A) 24 hr chromatograms of 5 mM $(GlcNAc)_5$ and $(GlcNAc)_6$ in the absence (top panels) and presence (bottom panels) of LaL.

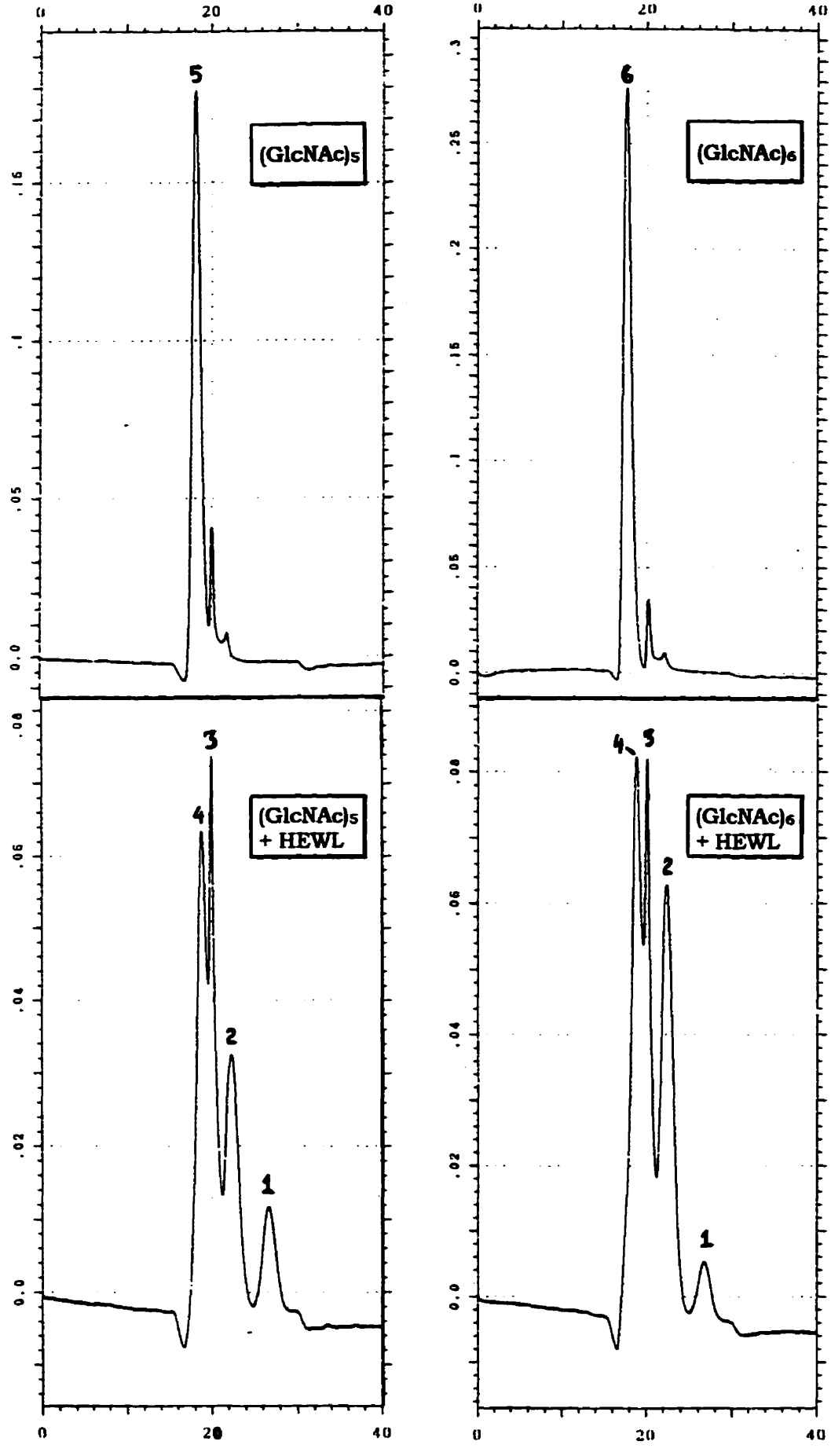
(B) 24 hr chromatograms of 1 mM $(GlcNAc)_5$ and $(GlcNAc)_6$ in the absence (top panels) and presence (bottom panels) of HEWL.

Peaks are designated as to the oligomeric size (n) of the saccharide. The ordinate and abscissa axes are absorbance (195 nm) and time (min) respectively.

(A)



(B)



descriptions on inhibition. For each saccharide experiment, a freshly thawed sample of substrate cells was used and inhibition was determined by comparison of the observed activities for LaL in the absence and presence of the inhibitor (i.e. the activity representing 100% enzyme activity was found independently for each unique experiment). In addition, the measurements in the absence and presence of inhibitor were performed in an alternating manner in order to offset possible aging affects with respect to the substrate cells after they had been thawed. Finally, solutions of LaL were prepared in stabilizing buffer. To ensure that the lysis measured was not influenced by ionic strength differences, an identical aliquot of enzyme solution or of stabilizing buffer was added to the substrate cells to measure the lysis in the presence and absence of LaL respectively.

A qualitative illustration of how the presence of an inhibitor is observed to affect the turbidimetric assay is shown in Figure 3.13 for the inhibition by $(GlcNAc)_5$. The activity curves demonstrate that the rate of lysis of the substrate cells decreases with increasing inhibitor concentration when compared to the curve generated in the absence of inhibitor. The 2 mM curves were actually obtained from a different experiment (under essentially identical conditions except that a different frozen aliquot of the same substrate cells was used) and is included with the figure for comparative purposes. Essentially similar types of curves were observed for all compounds which demonstrated an inhibitory propensity. Inhibition is then calculated from the selected linear portion of the curves for the no inhibitor and inhibitor present samples following correction for the enzyme blank.

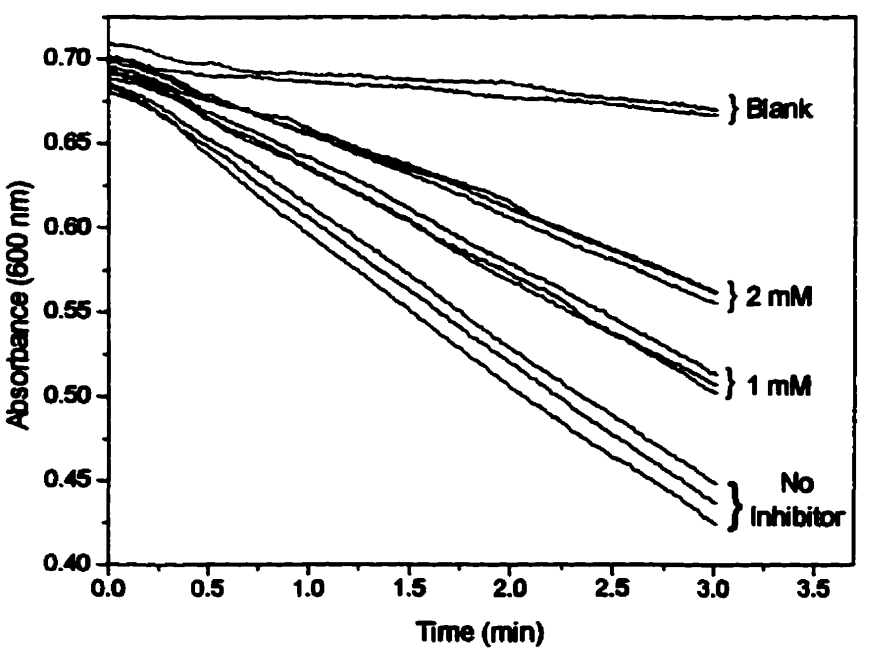


Figure 3.13. Inhibition of the bacteriolytic activity of LaL by $(GlcNAc)_5$.

The results of all the inhibition studies conducted with saccharides are presented in Table 3.7. A graphical presentation of the data concerning the inhibition by only $(GlcNAc)_n$ is given in Fig. 3.14. For each particular saccharide oligomer, inhibition was found to occur in a concentration-dependent manner. As well, enhanced inhibition is observed with increasing oligomeric size of $(GlcNAc)_n$ suggesting that LaL has a higher affinity for the larger saccharide oligomers. Although an exhaustive study including a greater number of concentrations for each saccharide was not performed, the data do present a trend regarding inhibition and oligomeric length. Each of $(GlcNAc)_{4-6}$ demonstrated a comparable level of inhibition to each other (within ≈15%) at concentrations of both 1 and 2 mM. However, substantially higher concentrations of the smaller saccharides were required to obtain these levels of inhibition. For example, inhibition by 5 mM $(GlcNAc)_3$ was nearly but not as efficient as the tetra-, penta- and hexamer at 1 mM and was far less effective than the tetramer at the same concentration of 5 mM. Similarly, 20 mM GlcNAc and 10 mM $(GlcNAc)_2$ demonstrated proportional inhibition to each other yet more weakly than did 10 mM $(GlcNAc)_3$. Therefore, the data

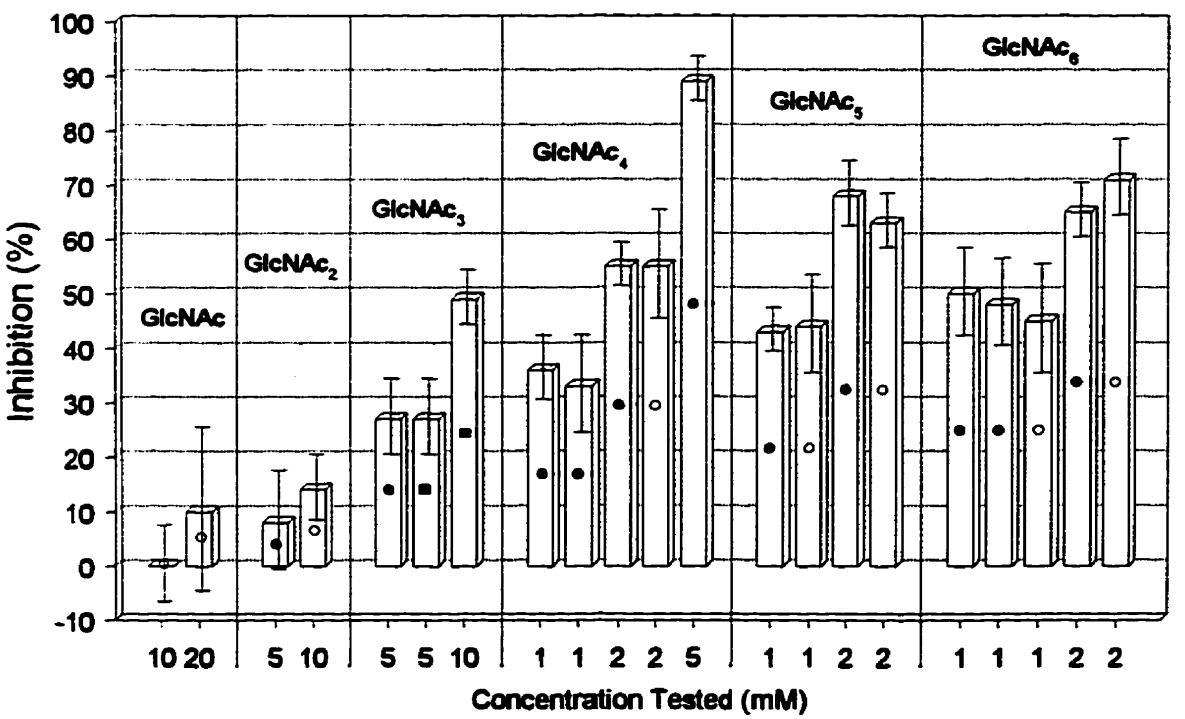


Figure 3.14. Inhibition of the bacteriolytic activity of LaL by $(GlcNAc)_n$. Different substrate cell preparations were used and these are indicated by the symbols within the bars and are: solid circle, SC#6; open circle, SC#9; solid square, SC#13. Refer to Table 3.7 for the specific values and associated errors.

Table 3.7. Inhibition of the bacteriolytic activity of LaL by selected saccharides.

Compound	Concentration Tested (mM)	Replica	SC†	Residual Activity (%)‡	Inhibition (%)*
GlcNAc	10.0	1	9	100 ± 7	0
	20.0	1	9	90 ± 16	10
$(GlcNAc)_2$	5.0	1	6	92 ± 9	8
	10.0	1	9	86 ± 6	14
$(GlcNAc)_3$	5.0	1	6	73 ± 7	27
	5.0	2	13	73 ± 7	27
	10.0	1	13	51 ± 5	49
$(GlcNAc)_4$	1.0	1	6	64 ± 6	36
	1.0	2	6	67 ± 9	33
	2.0	1	6	45 ± 4	55
	2.0	2	9	45 ± 11	55
	5.0	1	6	11 ± 4	89
$(GlcNAc)_5$	1.0	1	6	57 ± 4	43
	1.0	2	9	56 ± 10	44
	2.0	1	6	32 ± 6	68
	2.0	2	9	37 ± 5	63
$(GlcNAc)_6$	1.0	1	6	50 ± 8	50
	1.0	2	6	52 ± 8	48
	1.0	3	9	55 ± 10	45
	2.0	1	6	35 ± 5	65
	2.0	2	9	29 ± 7	71
Maltotetraose	5.0	1	6	101 ± 2	0
	5.0	2	9	101 ± 8	0
	10.0	1	9	106 ± 7	0
Cellotetraose	5.0	1	9	97 ± 11	3
Adjuvant Peptide	4.5	1	6	101 ± 9	0

† The particular substrate cell preparation used.

‡ Values determined by the turbidimetric assay ± absolute standard deviation. A sample calculation is given in Appendix C (section C.2).

* Values possess the same error as the residual activity. The data for $(GlcNAc)_n$ is graphically presented in Fig. 3.14 allowing a visual comparison of the data. Adjuvant peptide is N-acetylmuramyl-*L*-alanyl-*D*-isoglutamine.

obtained would suggest that $(GlcNAc)_{4-6}$ have comparable binding affinities to LaL which are stronger than $(GlcNAc)_{1-3}$. There is also a more evident dependency on inhibition and saccharide length with respect to the monomer through trimer and between the trimer and tetramer. This could indicate that a minimum of 4 saccharide units are required to satisfy interactions with LaL that will result in maximal binding leading to inhibition of peptidoglycan binding and hence, cell lysis with only marginal enhancement occuring when the additional saccharide units of $(GlcNAc)_5$ and $(GlcNAc)_6$ are present.

If one can consider inhibition as a credible reflection of binding affinity, then the results observed with LaL are somewhat different than those observed for HEWL. The association constants for $(GlcNAc)_{3-6}$ for HEWL are each on the order of $\approx 1\times10^5$ M^{-1} and are 2 and 4 orders of magnitude greater than for the dimer and monomer respectively. As such, maximal binding is achieved with the trisaccharide for HEWL and not with a tetrasaccharide as suggested above for LaL. GlcNAc, $(GlcNAc)_2$ and $(GlcNAc)_4$ demonstrate IC_{50} values of 7, 0.6 and 0.15 mM respectively (refer to Table 3.4) which are consistent with the differences in the association constants of these saccharides to HEWL noted above. Although the inhibition by $(GlcNAc)_4$ observed with LaL is within one order of magnitude of the IC_{50} value for the tetramer against HEWL (an IC_{50} value of $\approx$ 2 mM for $(GlcNAc)_4$ with LaL can be extrapolated from the data in Table 3.7), the monomer and dimer inhibit HEWL much more effectively than they do LaL. It is possible that the *E. coli* peptidoglycan will more readily displace the binding of $(GlcNAc)_n$ to LaL than will the peptidoglycan of *M. lysodeikticus* displace the binding of these sugars to HEWL. For this reason, inhibition of cell lysis by chitosaccharides may operate more effectively with HEWL against *M. lysodeikticus* than with LaL against *E. coli* even though their binding affinities to both enzymes may be comparable. It is also possible that the apparent ability of HEWL to simultaneously bind more than one saccharide in which multiple binding is more effective at inhibiting bacterial lysis than a singular binding may not occur in the case of LaL.

Maltotetraose and cellotetraose were chosen as control sugars for the inhibition experiments. Maltotetraose is the $\alpha(1\rightarrow4)$ tetramer of glucose while cellotetraose is the $\beta(1\rightarrow4)$ isomer. Both saccharides lack the 2-acetamido groups and, in the case of maltotetraose, the correct linkage of the peptidoglycan and are therefore suitable controls. Neither tetrasaccharide demonstrated an ability to inhibit the bacteriolytic activity of LaL (Table 3.7). The lack of inhibition by cellotetraose illustrates the importance of the 2-acetamido group for the binding of $(GlcNAc)_n$ to LaL. The adjuvant peptide *N*-acetylmuramyl-*L*-alanyl-*D*-isoglutamine more closely resembles the structure of the

peptide-substituted muramic acid component of peptidoglycan. However, it was not inhibitory towards the bacteriolytic activity of LaL at the concentration tested (Table 3.7), indicating very poor if any affinity for LaL. The cell wall peptide of *E. coli* is composed primarily of *L*-Ala-*D*-*iso*-Glu-*L*-*meso*-DAP-*D*-Ala, and the comparative truncated dipeptide of the adjuvant peptide may partially explain its lack of inhibition. In addition, the α-carboxyl of the *D*-glutamic acid, which is replaced with an amide in the adjuvant peptide, may be important for binding to LaL. Such a conclusion was made concerning the loss of inhibitory action towards T4L of a disaccharide-pentapeptide in which glycine forms an amide bond with the α-carboxyl of the glutamic acid (Jensen et al., 1976). Furthermore, as indicated in the crystal structure of the mutant T4L/peptidoglycan complex (Kuroki et al., 1993; see Fig. 3.4, p. 196), Arg137 and Gln141 interact with this carboxyl group and similar interactions may also be required with LaL.

We are confident in the inhibitory values obtained employing the turbidimetric assay. Not only were inhibition values reproducible when the same substrate cell preparation was used in replica measurements, but the reproducibility is also evident on comparison of values obtained with different substrate cells (Table 3.7). This is encouraging in that any acceptable preparation of substrate cells, when optimized accordingly with respect to their response to LaL, can be used to reliably appraise the inhibitory affect of a compound. Unfortunately however, a moderately high error is associated with the assay. Considering though that the error is propagated through calculations involving the rate of cell lysis due to the enzyme alone, the enzyme in the presence of the inhibitor and the enzyme blank, the error is quite respectable.

3.3.2.2. Fluorescence Studies

The formation of complexes of LaL with N-acetylglucosamine oligosaccharides were investigated by following changes in the intrinsic fluorescence properties of LaL. The results of this study are given in Table 3.8. Lambda lysozyme was found to fluoresce under 280 nm excitation with an emission maximum at 343-345 nm at pH 7.0. Other investigators (Jespers et al., 1992) have reported measuring the emission at 341 nm (with excitation at 280 nm) to follow the denaturation of LaL, but the pH at which this was performed was not given. It is assumed then that the reported 341 nm wavelength represented the emission maximum and therefore, is in close agreement for the emission maximum we have observed.

A trend does become apparent upon examination of the fluorescence changes observed for LaL in the presence of $(GlcNAc)_n$. A progressive decrease in the emission

Table 3.8. Effect of saccharides on the fluorescence of LaL and HEWL.

Compound	Concentration		λ_{max}†	Intensity	Δ Intensity‡	
	Enzyme (μM)	Compound (mM)	(± 1 nm)	(± 0.02) (arbitrary units)	units	%
A. LaL						
$(GlcNAc)_2$	2.0	-	344	0.60		
	2.0	5.0	344	0.61	+ 0.01	2
	2.0	10.0	344	0.63	+ 0.03	5
$(GlcNAc)_3$	2.0	-	345	0.66		
	2.0	5.0	343	0.68	+ 0.02	3
	2.0	10.0	342	0.64	- 0.02	-3
$(GlcNAc)_4$	1.1	-	344	0.57		
	1.1	1.0	342	0.61	+ 0.04	7
	1.1	5.0	338	0.65	+ 0.08	14
$(GlcNAc)_5$	1.1	-	344	0.52		
	1.1	1.0	340	0.54	+ 0.02	4
	1.1	5.0	338	0.57	+ 0.05	10
	2.5	-	343	0.49		
	2.5	10.0	338	0.54	+ 0.05	10
$(GlcNAc)_6$	2.5	-	344	0.50		
	2.5	1.6	339	0.50	0	0
	2.5	8.0	337	0.59	+ 0.09	18
Maltotetraose	2.0	-	344	0.69		
	2.0	5.0	344	0.68	- 0.01	-1
	2.0	10.0	343	0.70	+ 0.01	1
B. HEWL						
$(GlcNAc)_5$	2.5	-	343	0.53		
	2.5	2.0	335	0.62	+ 0.09	17
$(GlcNAc)_6$	2.5	-	343	0.53		
	2.5	2.0	335	0.64	+ 0.11	21

All measurements were performed in 50 mM KPB, pH 7.0.
† The emission maxima with excitation at 280 nm.
‡ The change in the fluorescence intensity at λ_{max} between the liganded and free enzyme.

maximum (λ_{max}) results that is dependent on oligosaccharide length and concentration. Where $(GlcNAc)_2$ caused no change in λ_{max} and $(GlcNAc)_3$ produced only a marginal blueshift, the higher concentrations of the tetramer through hexamer examined decreased λ_{max} from 343-345 nm (in the absence of sugar) to approximately 337-338 nm in their presence. This gives good indication towards a complexation of these saccharides with LaL. The blueshift suggests that the environment around the Trp residue(s) that dominate the fluorescence of LaL decreases in polarity. A shift in λ_{max} to shorter wavelength

associated with ligand binding is often taken as an indication that water is excluded in the complex (Freifelder, 1982). Of the four Trp residues in LaL, it is reasonable to suggest that if one or more line the binding region and are partially exposed to solvent, then the observed blueshift could indicate the displacement of solvent with saccharide. It is also possible that the shift in wavelength could be attributable to conformational changes in the protein which alters the environment around the dominating fluorophores. Since the dimer and trimer do not cause the similar blueshift noted with the larger saccharides, they may only interact with a smaller (or different) portion of the binding region than do the larger saccharides and that the binding regions of the smaller saccharides are removed from influencing these possible Trp residues. It is also possible that the binding regions from dimer to hexamer do indeed overlap but the association of dimer and trimer may be weaker with LaL than with the tetra-, penta- and hexamer which may account for the lack of any substantial shift in emission maxima for the former two sugars. These differences tend to parallel the inhibitory properties in which a notable change in the influence of the saccharide on the measured parameter (inhibition or fluorescence) exists between $(GlcNAc)_3$ and $(GlcNAc)_4$.

Some interesting similarities exist between the fluorescence properties of HEWL and LaL. Traditionally, fluorescence studies with HEWL/$(GlcNAc)_n$ complexes have been performed with excitation at 285 or 290 nm and monitoring emission between 340 to 370 nm (Imoto et al., 1972; Banerjee & Rupley, 1973a). The emission properties of the complexes vary with pH and are strongly quenched at extreme (>10) pH (Lehrer & Fasman, 1967). In order for comparative purposes, fluorescence spectra were obtained for HEWL under the same conditions as those obtained for LaL.

The observed emission maxima of LaL and HEWL are nearly identical (both around 343-344 nm, Table 3.8). This observed emission maximum for free HEWL (with excitation at 280 nm) agrees well with previously reported measurements (Lehrer & Fasman, 1966; Imoto et al., 1971). Of the 6 Trp residues in HEWL, Imoto et al. (1971) have estimated that Trp62 is responsible for 35-38% of the fluorescence in the pH interval of 2-8, while Trp108 is the other major contributor to HEWL fluorescence. Tryptophan 62 is the most exposed of HEWL's 6 Trp residues and makes extensive non-polar contacts with the sugar ring occupying site B of HEWL. Tryptophan 108 on the other hand, is partially buried, emits at a shorter wavelength (Imoto et al., 1971) and makes hydrophobic contacts with the acetamido group of a sugar in site C and to the side chains of Ile58 and Ile98 (Strynadka & James, 1991). Since fluorescence is dominated by Trp residues, the similar emission maxima for LaL and HEWL under identical conditions indicates that the average

environment of fluorescent tryptophan residues could be very similar in both enzymes. This raises the possibility that the roles of Trp residues in the binding regions of both LaL and HEWL could have similar and related functions (see below). There does exist a strong sequence similarity upon comparison of the regions containing Trp62 and Trp63 of HEWL and Trp73 and Trp74 of LaL (Fig. 3.15). The only analogous Trp of LaL remaining that could be related to Trp108 of HEWL is Trp124, but the sequence similarity in these regions of the proteins are not as strong.

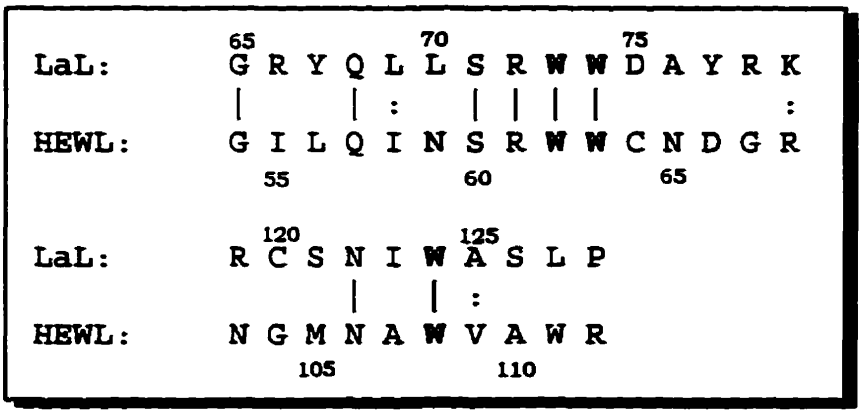


Figure 3.15. Alignment of Trp containing regions of LaL and HEWL.

In protein-carbohydrate interactions, there is strong agreement that van der Waals contact of aromatic side chains, especially of tryptophan, on the saccharide ring is important for carbohydrate recognition and may be a general feature of protein-carbohydrate binding (Quiocho, 1986; Vyas, 1991). The plane of the indole ring typically lies parallel with the plane of the carbohydrate ring at a distance of 3-5 Å, and this type of stacking relationship has been observed in several structural determinations (e.g. Trp17 of arabinose binding protein (Quiocho, 1988; Johnson et al., 1988), and Trp340 of maltodextrin binding protein (Spurlino et al., 1991)). The role of Trp62 in HEWL has been explored by site directed mutagenesis at this position which has been observed to produce marked changes of the activity of the enzyme (Kumagai et al., 1992b, 1993). As well, Trp62 is also considered to maintain or produce the local structure or the critical steric configuration of the protein molecule around this residue that is most suited for binding (Maenaka et al., 1994).

Tryptophan 62 of HEWL is substituted with tyrosine in homologous lysozymes from rat, bovine and human. The emission maximum (at 280 nm excitation) for HuL is 330 nm (Mulvey et al., 1974) as compared to the 344-345 nm emission maxima observed for LaL and HEWL. With the human enzyme, the fluorescence is now dominated by Trp109

(analogous to Trp108 in HEWL) which as mentioned above, is in a more hydrophobic environment and hence the emission is at lower wavelengths. This can again be taken as an indication that there exists a Trp in LaL which may be analogous to Trp62 in HEWL.

In addition, both HEWL (Lehrer & Fasman, 1966) and HuL (Mulvey et al., 1974) show dequenching (i.e. an increase in the quantum yield) in their fluorescence spectra upon complexation with chitooligosaccharides. Related dequenching is also observed for LaL in the presence of $(GlcNAc)_{4\text{-}6}$, although the magnitude of dequenching was found to be less than with HEWL (Table 3.8) and higher concentrations of saccharides appeared to be necessary for dequenching to occur with LaL than with HEWL. No shift in λ_{max} nor any dequenching was observed with LaL in the presence of maltotetraose indicating that complexation was not detected by fluorescence.

In all *c* type lysoyzme sequences known by 1989, an homologous Trp to that of Trp108 in HEWL is conserved (Nitta & Sugai, 1989). In the known crystal structures of the *c* type lysozymes, Trp108 is in van der Waals contact with Glu35 and its hydrophobicity was shown to play an important role in maintaining the abnormally high pKa of Glu35 in HEWL (Inoue et al., 1992a). It has been suggested that the proximity of the carboxyl of Glu35 quenches the fluorescence of Trp108 and that the dequenching upon $(GlcNAc)_n$ complexation is due to conformational changes in which Glu35 moves apart from Trp108 and this relieves the carboxyl quenching (Pecht et al., 1970). The dequenching of Trp108 has also been attributed to deprotonation of Glu35 (Lehrer & Fassman, 1967) which is validated by the known increase in the p*K*a of this group from 6.0-6.1 to 6.5-6.6 (Banerjee et al., 1973, 1974) upon binding of chitosaccharides.

The dequenching observed with some of the saccharides given in Table 3.8 with LaL again raises the possibility that the mechanism of dequenching with LaL could result by related processes described above. If indeed true, then an analogous Trp to Trp108 of HEWL could exist in LaL (see Fig. 3.15 for a possibility) which could play the same role in saccharide binding and maintenance of a high p*K*a value for the assumed catalytic general acid in the active site of LaL. The determination of the crystal structure of LaL will be essential before such suppositions can be confirmed.

Extrinsic fluorescence was also probed using the 4-methylumbelliferyl-β-glycoside of N-acetyl-chitotriose ($(GlcNAc)_3$-4MU). The fluorescence spectra of both $(GlcNAc)_2$-4MU and $(GlcNAc)_3$-4MU excited at 330 nm have a single emission maximum at 375 nm at pH 5.18 (Yang & Hamaguchi, 1980) or at 372 nm when excited at 335 nm at pH 5.75 (Delmotte et al., 1975). Under our experimental conditions, the fluorescence spectra of

$(GlcNAc)_3$-4MU in 50 mM KPB, pH 7.0 was found to give rise to two emission maxima at 376 nm and 395 nm when excited at 330 nm. The emission maximas were found to merge to a broad single maximum at 388 nm when the excitation wavelength was progressively increased to 339 nm. Repeated spectra with different freshly prepared samples of $(GlcNAc)_3$-4MU gave the same results.

Potential binding of $(GlcNAc)_3$-4MU to LaL was monitored by following the spectral changes upon addition of approximately equivalent and excess enzyme. The results given in Table 3.9 demonstrate that only small increases in the intensity of the two emission maxima for the fluorophore occurred upon addition of LaL. As a control, the same experiments were performed under identical conditions with HEWL which revealed that the emission intensities increased to a greater extent. With both LaL and HEWL, there was no shift of the maxima emission wavelengths in the presence of either enzyme. The observed increases in intensities resulting from the presence of HEWL are consistent with those reported previously (the intensity of 2.18 μM $(GlcNAc)_3$-4MU at pH 5.18 increased by ca. 13 and 26% in the presence of 2.39 and 13.7 μM HEWL respectively (Yang & Hamaguchi, 1980); the intensity of 11.3 μM $(GlcNAc)_3$-4MU at pH 5.75 increased by ca. 43% in the presence of 26.4 μM HEWL (Delmotte et al., 1975)). As such, our results appear to confirm that under the experimental conditions used, binding of $(GlcNAc)_3$-4MU to HEWL was occurring.

Table 3.9. Effect of LaL and HEWL on the Fluorescence of $(GlcNAc)_3$-4MU

Enzyme Concentration:	None	2.4 μM		13.7 μM	
		Intensity	% Increase†	Intensity	% Increase†
LaL					
λ_{376}	0.48	0.47	-2	0.51	6
λ_{395}	0.50	0.49	2	0.54	8
HEWL					
λ_{376}	0.52	0.60	15	0.72	38
λ_{395}	0.54	0.64	18	0.79	46

The concentration of $(GlcNAc)_3$-4MU was 2.2 μM in 50 mM KPB, pH 7.0. The excitation wavelength was 330 nm and emission intensities were measured at 376 and 395 nm.
† The percent increase in the intensity relative to free $(GlcNAc)_3$-4MU (i.e. no enzyme).

Van Landshoot et al. (1977) reported that the fluorescence intensity of $(GlcNAc)_3$-4MU decreases with decrease in solvent polarity. The observed increases in intensity suggests that on complexing with HEWL, the environment of the 4MU group is more polar than water. It has been shown that $(GlcNAc)_3$-4MU binds predominantly in a non-

productive manner to sites A-D of HEWL and also productively to sites B-E which will result in hydrolysis of the substrate to give $(GlcNAc)_3$ and 4MU (Yang & Yamaguchi, 1980). Since non-productive binding predominates, the terminal 4MU group is expected to rest primarily at site D in HEWL. Since only minor changes (as compared to those observed for HEWL) in the intensities were observed in the presence of LaL this suggests that i) $(GlcNAc)_3$-4MU does not bind or binds only weakly to LaL; ii) the saccharide portion binds to LaL such that the 4MU group is still primarily exposed to the solvent; or iii) the entire ligand including the 4MU are in contact with LaL but the protein environment adjacent to the 4MU group does not affect its spectral properties. Even though we have established that $(GlcNAc)_3$ both inhibits the bacteriolytic activity of and increases the thermal stability of LaL and thereby interacts with LaL, it is possible that the 4MU moiety hinders the association of $(GlcNAc)_3$-4MU with LaL.

3.3.2.3. Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) is a powerful method available for detection of temperature induced transitions of macromolecules. The technique has been applied to the study of protein transitions yielding information about protein folding and stability (Privalov, 1989; Dill et al, 1989), domain structure and interaction (Privalov, 1982) and ligand binding (Hinz, 1983; Brandts et al., 1989). To obtain a greater appreciation of the interaction of LaL with chitooligosaccharides, we chose to explore the effect of these saccharides on the thermal stability of LaL.

Initial efforts were directed to establish a concentration of LaL which gave suitable data for analysis. It was suggested by Dr. H. Frey (Dept. of Physics, University of Waterloo) that a concentration of 3.5 mg/mL (196 μM) of LaL should be employed. A typical heat capacity (Cp) endotherm for the thermal denaturation of LaL is shown in Fig. 3.16. The curvature visible in the raw data curves (Fig. 3.16 A) is apparently intrinsic to the calorimeter (Lepock et al., 1993) necessitating baseline correction (Fig. 3.16 B). The transition temperature (T_m) was the parameter chosen to monitor the calorimetric results. The transition temperatures for all curves obtained in this study were found with the Microcal Origin v2.90 software. As seen in Fig. 3.16, the T_m determined by the software was the same for the raw and the processed curves. This was true in all samples measured indicating that curve smoothing and baseline correction did not affect measurement of the transition temperature.

The DSC curves in Fig. 3.16 reveal a single large endotherm at ≈ 49.4 °C (see also Table 3.10) for LaL at 3.5 mg/mL in 50 mM KPB, pH 7.0, indicative of the thermally induced denaturation of LaL. The specific nature giving rise to the other peaks at higher

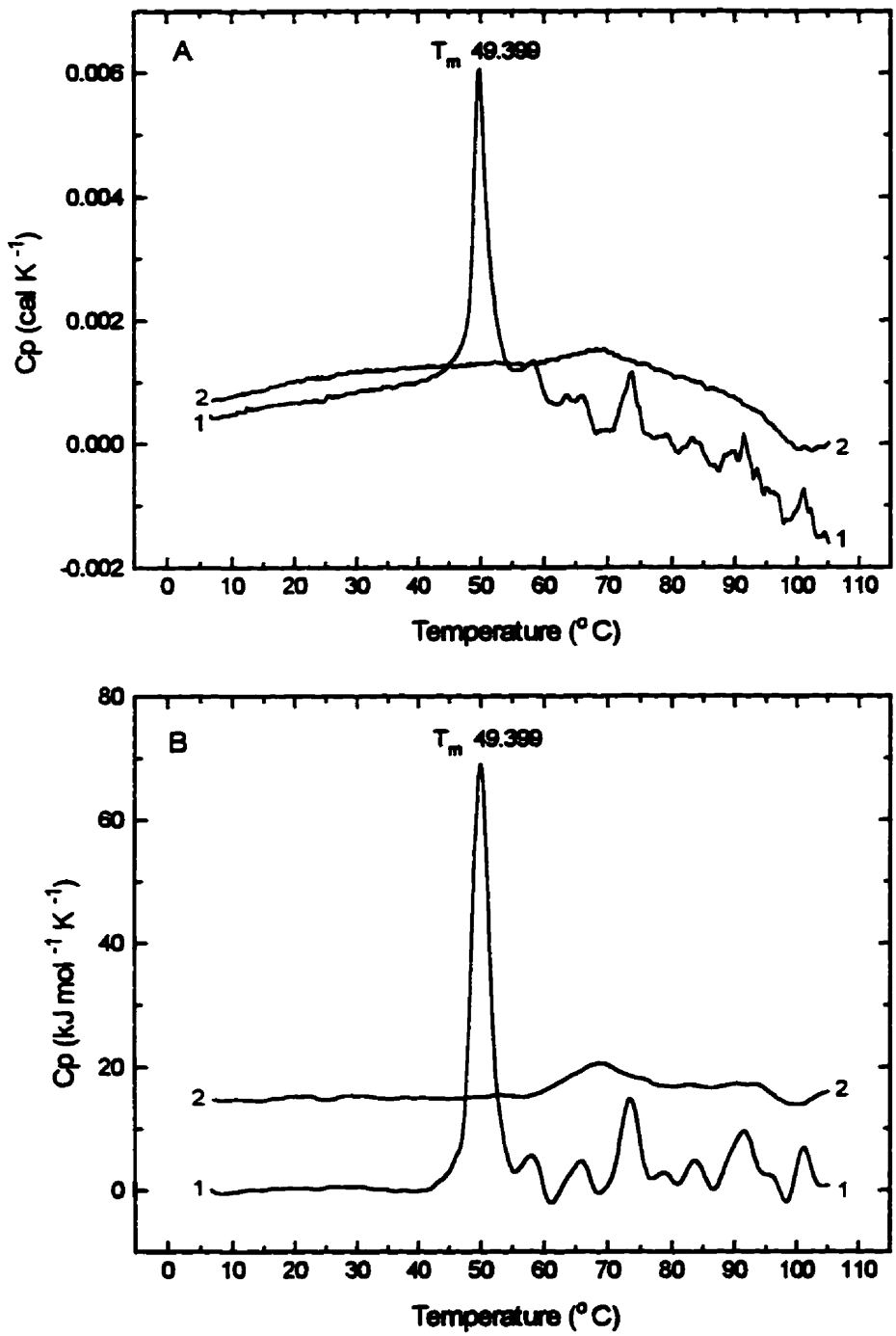


Figure 3.16. Typical DSC curve for the thermal denaturation of LaL. Protein concentration was 3.5 mg/mL (196 μM) in 50 mM KPB, pH 7.0. (A) Raw data. (B) Data generated following curve smoothing, baseline correction and normalization to protein concentration. The sample was scanned at 1 K/min giving the initial scan (curves 1) then cooled before rescanning under the same conditions to give the rescan (curves 2). The rescan in (B) was intentionally raised for clarity from its original ordinate position at 0 Cp.

temperatures is not clear but could possibly represent transitions of aggregates formed following the initial denaturation of LaL. As well, no explanation can be given as to the observation that the profiles of these "higher temperature peaks" were not consistent from sample to sample. In some cases, the curve was virtually flat at the higher temperatures while in others, this region of the curve resembled that shown in Fig. 3.16. In all cases however, the appearance of the large endotherm was consistent.

The effect of $(GlcNAc)_5$, $(GlcNAc)_3$ and maltotetraose (again chosen as a control) on the transition temperature of LaL is shown in Table 3.10 and Fig. 3.17. From triplicate measurements at a protein concentration of 3.5 mg/ml, LaL exhibits a transition at 49.37 ± 0.07 °C. At concentrations of 5 mM, the presence of $(GlcNAc)_5$ increased the T_m by 1.37 ± 0.09 °C, the presence of $(GlcNAc)_3$ increased the T_m by 0.55 ± 0.10 °C whereas the presence of maltotetraose resulted in a 0.33 ± 0.13 °C increase in T_m (Table 3.10). As such, the presence of each of the saccharides investigated increased the thermal stability of LaL. In parallel with their inhibitory properties, $(GlcNAc)_5$ appears to have a greater stabilizing effect on LaL than that of $(GlcNAc)_3$, once again suggesting that stronger interactions occur with an increase in oligomeric size. The increase in T_m observed with maltotetraose, although significant when compared with the effect of $(GlcNAc)_3$, is likely the result of non-specific interactions with the enzyme in a manner that polyhydroxylated compounds, such as glycerol and other sugars, stabilize proteins (Wiseman, 1978; Timasheff & Arakawa, 1990). The lack of observed inhibition by maltotetraose on the bacteriolytic activity of LaL (see Table 3.7) is consistent with such non-specific interactions.

Table 3.10. Transition temperatures obtained for LaL in the presence and absence of selected saccharides.

Sample	T_m (° C)[†]	Δ T_m (°C)
LaL	49.37 ± 0.07	-
LaL + 5 mM Maltotetraose	49.70 ± 0.11	0.33 ± 0.13
LaL + 5 mM $(GlcNAc)_3$	49.92 ± 0.07	0.55 ± 0.10
LaL + 5 mM $(GlcNAc)_5$	50.74 ± 0.05	1.37 ± 0.09

[†] Values represent the average ± standard deviation for 3 independent scans.

Under the same experimental conditions, the stability of HEWL was increased substantially more by $(GlcNAc)_3$ than was LaL. As shown in Fig. 3.18, the T_m for HEWL was increased by 3.9 °C from 73.0 °C for the enzyme alone to 76.9 °C for the enzyme in

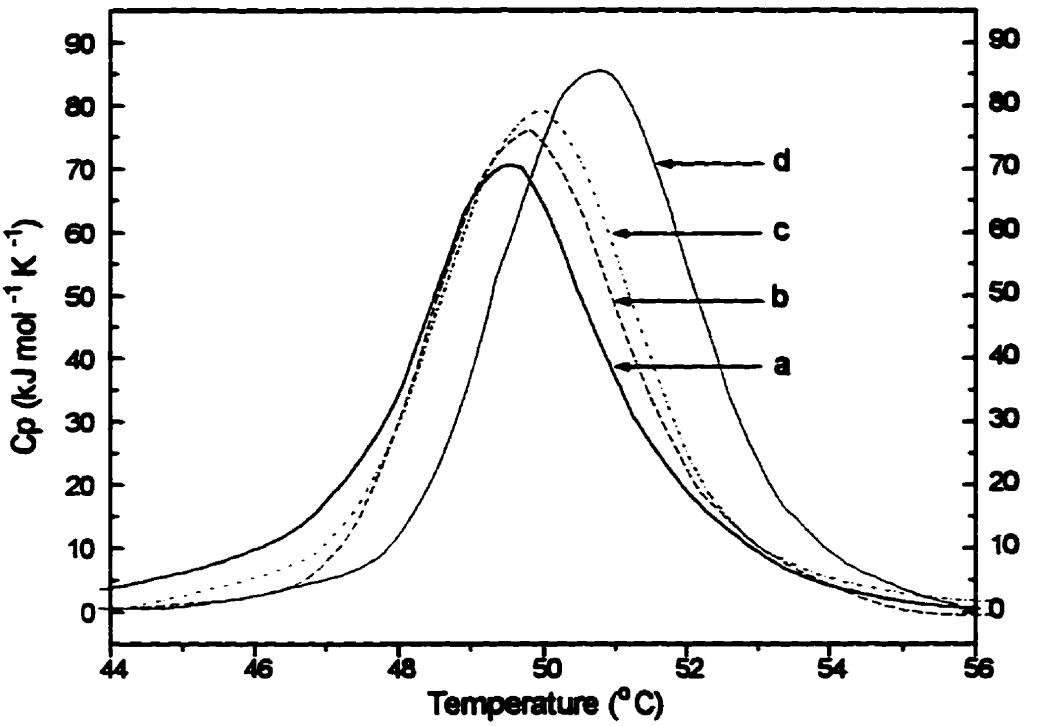


Figure 3.17. Representative DSC scans observed for LaL in the presence of selected saccharides. Protein concentration was 196 μM in 50 mM KPB, pH 7.0. (a) LaL alone; (b) LaL and 5 mM maltotetraose; (c) LaL and 5 mM $(GlcNAc)_3$; (d) LaL and 5 mM $(GlcNAc)_5$. See Table 3.10. for specific T_m values.

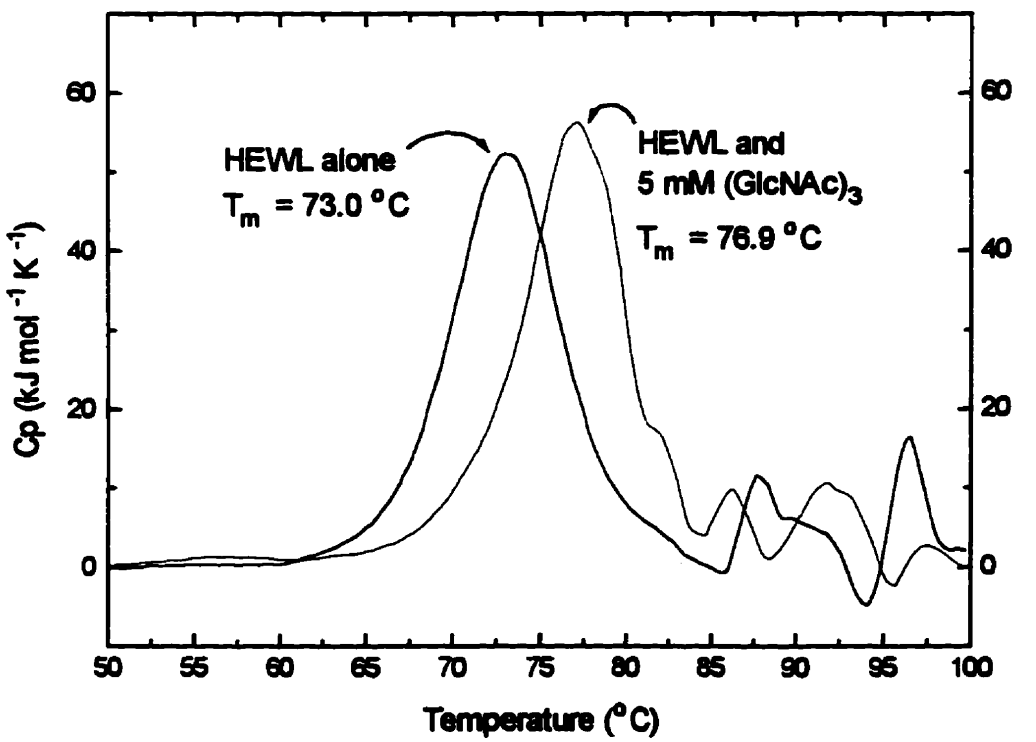


Figure 3.18. DSC scans for HEWL in the presence and absence of $(GlcNAc)_3$. Protein concentration was 4.0 mg/mL (278 μM) in 50 mM KPB, pH 7.0.

the presence of 5 mM $(GlcNAc)_3$. The observed T_m for HEWL falls into a range of previously reported measurements using different techniques (see Table XIX in Imoto et al., 1972; Radford et al., 1991). As such, it is apparent that HEWL is a more thermally stable protein than is LaL. As well, the much larger increase in T_m induced by $(GlcNAc)_3$ with HEWL (3.9 °C) than with LaL (0.55 °C) suggests that this saccharide is more able to stabilize the structure of HEWL and therefore, must have stronger or more elaborate interactions with HEWL than with LaL.

Our differential scanning calorimetry investigations with LaL also provided information on the reversible denaturation of the enzyme. When scanned to 105 °C, the rescan was irreversible in all cases (eg. see Fig. 3.16). However, when the first heating is brought to only 51 °C (approximately 80% completion), partial reversibility was observed (Fig. 3.19). The heat absorbed for the rescan in Fig. 3.19 was between 84-93 % of the heat absorbed for initial scans of other samples (in which the first scan was brought to completion) and the transition occurred at 50.7 °C. The transition temperature was also dependent on the protein concentration and it was determined that the T_m increased with decreasing protein concentration. Single determinations with protein samples run at 7.0 and 1.95 mg/mL (in 50 mM KPB, pH 7.0) gave transition temperatures of 48.8 °C and 51.2 °C respectively.

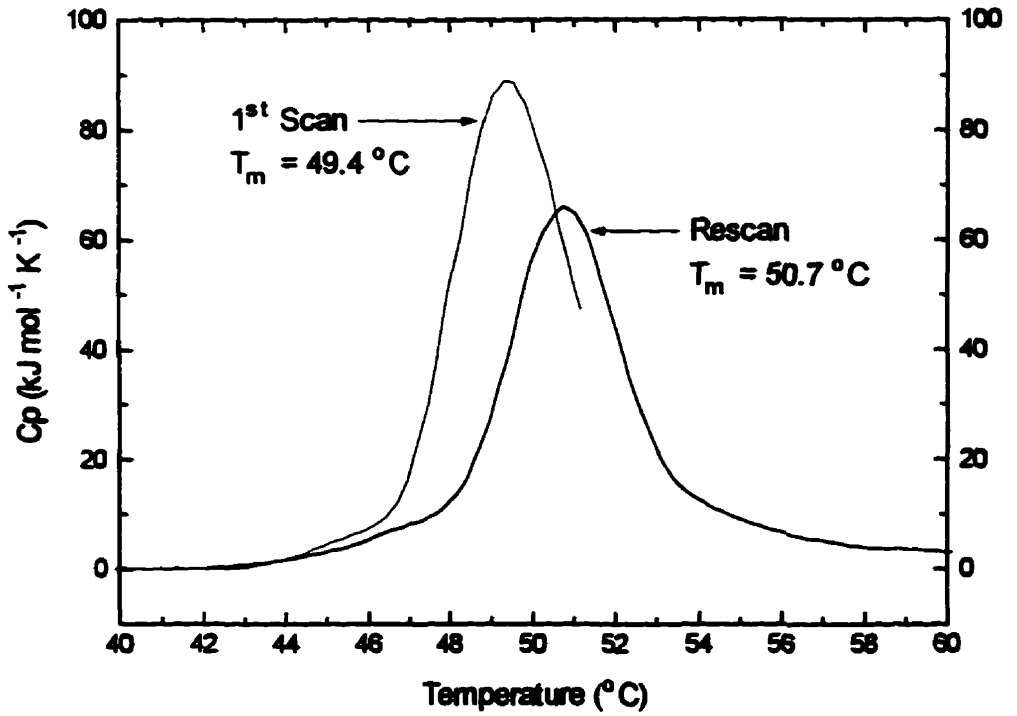


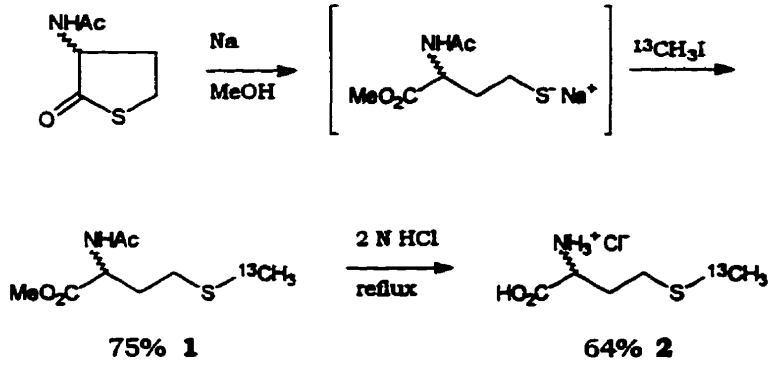
Figure 3.19. DSC scans demonstrating the partial reversibility of the thermal denaturation of LaL. The first scan was brought to 51 °C, the sample was cooled *in situ*, and then the rescan was obtained.

The energetics of protein unfolding and protein-ligand equilibria can in principle be derived from direct calorimetric measurements, yielding the thermodynamic parameters ΔG°, ΔH° and ΔS° involved in these processes. However, the procedures involved in the determination of these values often involves fitting of data to appropriate models (e.g. van't Hoff enthalpy change) and proper deconvolution and manipulation of the data, each requiring a certain amount of intuitive expertise. After personal consultation with Drs. H. Frey and J. Lepock (Dept. of Physics, University of Waterloo) and realization of the type of analysis required, it was decided at the time that it would not be feasible to determine the aforementioned thermodynamic parameters and we chose to describe our results with transition temperatures. Furthermore, a more complete calorimetric study would involve multiple concentrations of selected ligands, requiring amounts of protein that would have demanded constant devotion of efforts towards purification of LaL. The DSC data presented above had already consumed a substantial quantity of 125 mg of LaL which was purified at levels of 6-7 mg/L from *E. coli* TG-1/pHDM10.

3.3.2.4. NMR Studies on [methyl-^{13}C]Methionine Labelled LaL

Protein NMR is frequently adopted as an effective method to provide information on protein-ligand interactions and has been applied to explore complexes of GlcNAc oligomers with lysozymes from hen egg white (Szilágyi & Forgó, 1991; Fukamizo et al., 1992; Lumb et al., 1994), guinea egg white (Fukazimo et al., 1991) and human (Ohkubo et al., 1991). The protein NMR assignment of LaL is, as of yet, not available preventing comparable studies. We have, however, chosen to introduce methionine into LaL in which the methyl group is labelled with ^{13}C as a marker for NMR detection. This approach was chosen for several reasons; introduction of a ^{13}C probe enhances and permits the detection of the signals of the carbon and protons attached to the enriched carbon in a [^{1}H-^{13}C]HMQC experiment with relatively small quantities of protein; of its 158 amino acids, LaL contains only three methionines separated at positions 1, 14, and 107 thereby simplifying the spectra and not limiting the probe to only a specific region of the protein; the proton methyl signals of Met105 of HEWL are shifted upon binding $(GlcNAc)_n$; other proteins and enzymes having methionines in their active/binding sites and that are involved in saccharide chemistry have been found such as the arabinose-binding protein (Vermersch et al., 1991), the *E. coli* exomuramidase Slt70 (Thunnissen et al., 1995a) and T4L (Kuroki et al., 1993); and this work was a prelude to our studies on the incorporation trifluoromethionine into LaL presented in Chapter 4. The functional importance of methionine in proteins is also elaborated upon in Chapter 4.

The ^{13}C-enriched methionine analogue is commercially available however we were able to prepare the amino acid effectively (Scheme 3.2) based on methodology previously developed in our laboratory (Houston, 1992). In the presence of sodium methoxide, N-acetyl-*D,L*-homocysteine thiolactone is converted to the thiolate *in situ* which is followed by S-alkylation with ^{13}C-methyl iodide giving the fully protected racemic amino acid **1**. Acid hydrolysis of the N-acetyl and methyl ester groups of **1** afforded *DL*-[*methyl*-^{13}C]methionine hydrochloride **2** in an overall yield of 48%.



75% **1** 64% **2**

Scheme 3.2.

In order to incorporate the ^{13}C-enriched methionine into LaL, the expression plasmid pHDM10 was transformed into *E. coli* DH93, a methionine auxotroph that is deficient for the two methionine synthetases (*metE* and *metH*) in *E. coli*. It was necessary to determine the required amount of methionine supplementation that would support normal growth of this strain. The effect of *L*-methionine supplementation on the growth of *E. coli* DH93/pHDM10 is shown in Figure 3.20. From this study, it was determined that a supplementation at concentrations of 0.2 mM or higher provided an adequate source of methionine that sustained maximal growth especially during the exponential phase when the cells would be subjected to induction of LaL synthesis. As is expected, there is an observed dependence on the growth of the auxotroph for concentrations below 0.2 mM. For this reason, during preparation of labelled LaL, the auxotroph was grown in media supplemented with 0.4 mM *DL*-[*methyl*-^{13}C]methionine hydrochloride. It is of interest to note that the generation time of the auxotroph (≈ 55 min) was less than for the parent non-auxotrophic strain MG1655 (≈100 min). This illustrates a metabolic devotion for the

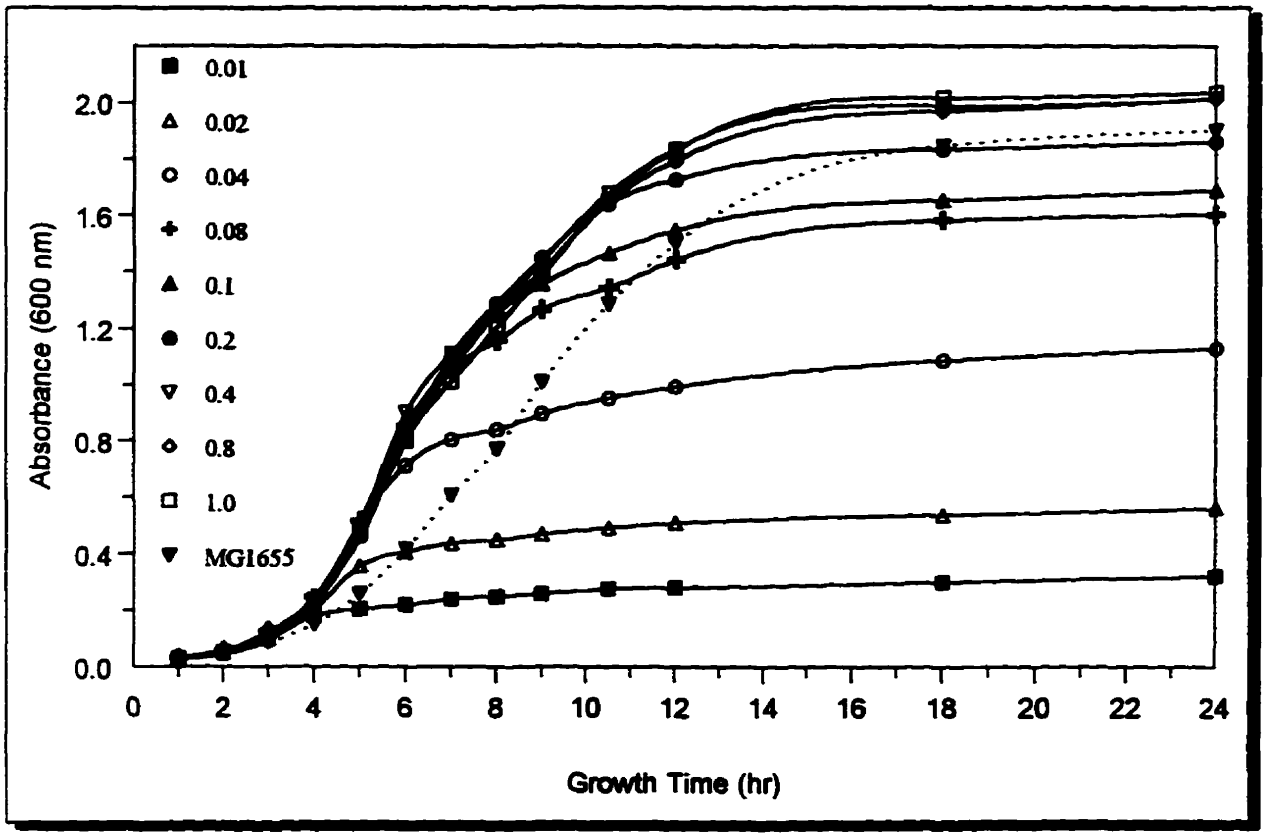


Figure 3.20. Dependence on methionine supplementation for the growth of *E. coli* DH93 harboring pHDM10. Cultures were grown in M9 medium at 37 °C supplemented with glucose, ampicillin and various concentrations of *L*-methionine as indicated by the symbols (in mM). Also shown is the growth curve of *E. coli* MG1655 grown in M9 medium supplemented with glucose.

de novo synthesis of methionine required for cell propagation which is attenuated if an exogenous source of methionine is supplied.

Growth and induction of *E. coli* DH93/pHDM10 in M9 minimal medium containing 0.4 mM *DL*-[*methyl*-^{13}C]methionine hydrochloride was performed under the standard conditions. The enriched enzyme, ^{13}C-Met-LaL, could be purified following the established protocols at a level of 0.5 mg/L and appeared homogeneous by SDS-PAGE analysis. Since only [*methyl*-^{13}C]methionine was present in the medium, it is expected that the enrichment of the methionines in ^{13}C-Met-LaL isolated approaches 100%. Enrichment of the methyl groups is clearly demonstrated in the [^{1}H-^{13}C]HMQC spectra of ^{13}C-Met-LaL (Fig. 3.21), in which three well resolved signals arising from each of the three methionine residues in LaL are observed. The HMQC experiment will provide a signal (or correlation) for each unique C-H pair (or C-H_n group if the *n* protons are chemical shift equivalent) at

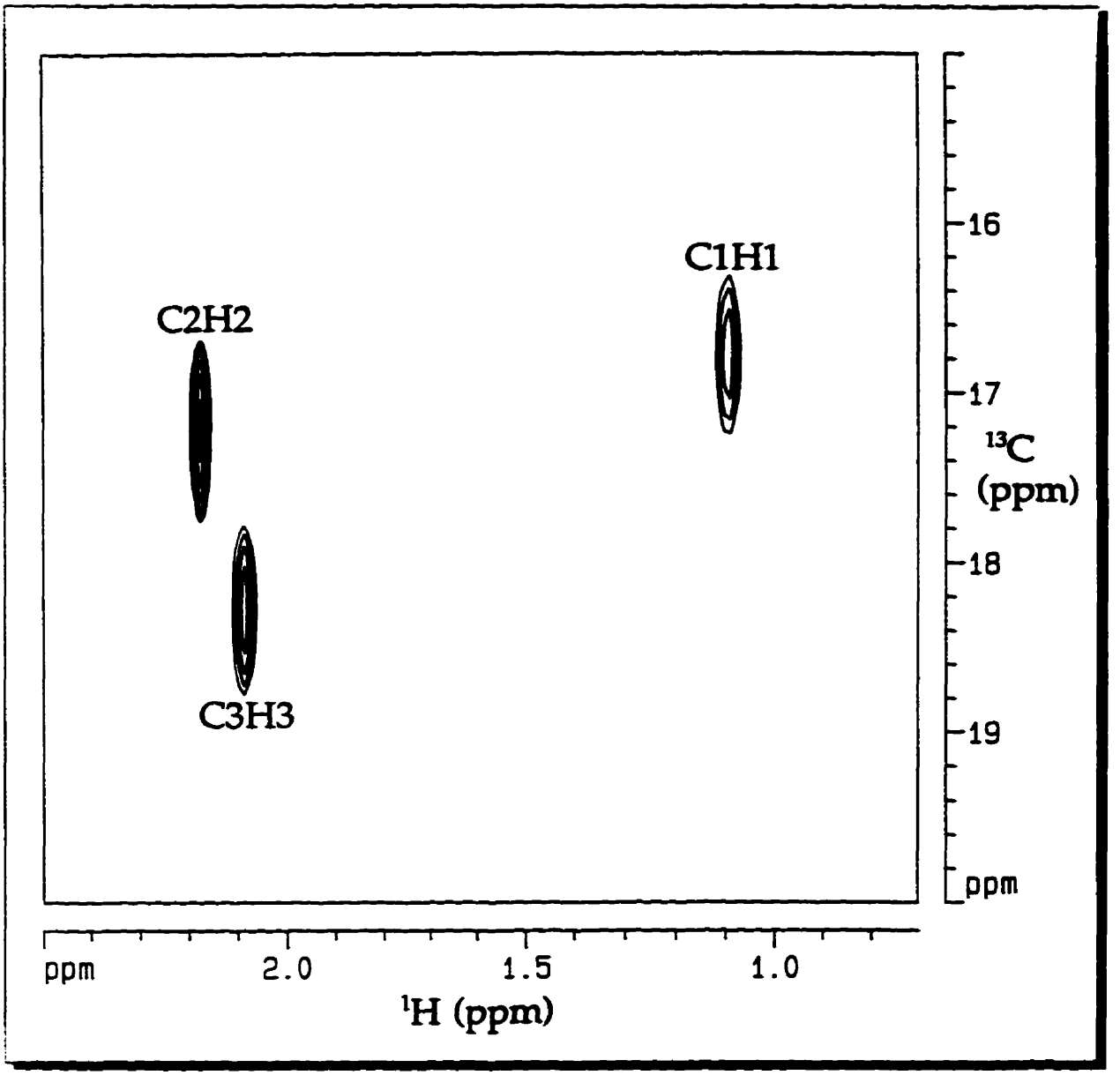


Figure 3.21. [^{1}H-^{13}C]HMQC spectrum of [*methyl*-^{13}C]methionine labelled LaL illustrating the signals arising from the ^{13}C-enriched methyl groups of the three methionine residues of LaL. The signals are arbitrarily designated as indicated and is referenced to internal TSP (0.00 ppm).

the respective chemical shifts of the proton and carbon. As the methionyl methyl groups of ^{13}C-Met-LaL are enriched with ^{13}C, their sensitivity to detection will be approximately 100 times that of the other C-H pairs in the protein molecule. Attempts to assign the signals shown in Figure 3.21 to the specific methionine positions by application of other two dimensional NMR techniques were unsuccessful and we have arbitrarily designated them as C1H1, C2H2 and C3H3. The proton and carbon resonances are in good

agreement for chemical shifts reported for the methionine methyl group in other proteins (Matta et al., 1981; Sailer et al., 1993; Bruix et al., 1993; Macquaire et al., 1993; Barden & Kemp, 1993).

The results of the NMR experiments with ^{13}C-Met-LaL in the absence and presence of $(GlcNAc)_5$ and maltotetraose are summarized in Table 3.11. Within the uncertainty of assignments, little or no variation in the chemical shifts of ^{13}C-Met-LaL were observed for the two protein and buffer concentrations examined. The presence of $(GlcNAc)_5$ produced a change in chemical shift in C1 and C3 in what would appear to be a concentration dependent manner. $(GlcNAc)_5$ at 2 and 5 mM lowered the chemical shift of C1 by approximately 0.1 ppm and at 10 mM, lowered it by 0.15 ppm. Conversely, the chemical shift of C3 was raised by 0.05 ppm with 2 mM $(GlcNAc)_5$ and by approximately 0.15 ppm in the presence of 5 and 10 mM. No significant change in the proton chemical shifts could be noted. Although the change in chemical shifts resulting from the presence of $(GlcNAc)_5$ are small, they do however suggest a change in the local environment around these methionine residues, either incurred through the direct binding of $(GlcNAc)_5$ in a vicinity nearby to the methionines or from conformational changes in the protein structure induced upon binding. No change in the chemical shifts were observed for the methionine methyl groups in the presence of maltotetraose suggesting that these specific interactions do not occur with this sugar.

Table 3.11. Effect of $(GlcNAc)_5$ and maltotetraose on the chemical shifts (in ppm) of the ^{1}H- and ^{13}C-nuclei of the methionine methyl groups in ^{13}C-Met-LaL. The significant changes are indicated in bold type.

Sample	C1	H1	C2	H2	C3	H3
LaL alone [a]	16.80	1.07	17.20	2.17	18.26	2.07
LaL + 2 mM $(GlcNAc)_5$ [a]	**16.70**	1.06	17.19	2.17	**18.31**	2.07
LaL + 5 mM $(GlcNAc)_5$ [a]	**16.72**	1.05	17.22	2.17	**18.41**	2.07
LaL + 10 mM $(GlcNAc)_5$ [a]	**16.65**	1.05	17.22	2.17	**18.40**	2.07
LaL alone [b]	16.81	1.07	17.21	2.16	18.27	2.06
LaL + 5 mM Maltotetraose [b]	16.79	1.07	17.20	2.16	18.25	2.07
LaL + 10 mM Maltotetraose [b]	16.80	1.07	17.21	2.16	18.26	2.06

All signals relative to internal TSP (0.00 ppm).
[a] 0.39 mM LaL, 128 mM KPB (pH 7.0).
[b] 0.22 mM LaL, 87 mM KPB (pH 7.0).

At the time when ^{13}C-Met-LaL was prepared from *E. coli* DH93/pHDM10, plasmid pLR102 had yet to be constructed. Plasmid pLR102 was introduced into *E. coli* B834(λDE3), a commercially available (Novagen) derivative strain of BL21(λDE3) that is auxotrophic for methionine. *E. coli* B843/pLR102 was intended primarily for the incorporation of the trifluoromethionine into LaL (Chapter 4) but it also allowed for the purification of ^{13}C-Met-LaL at the same high levels as obtained for purification of *wt* LaL from *E. coli* BL21(λDE3)/pLR102. Growth and induction of *E. coli* B834/pLR102 in M9 minimal media supplemented with 0.2 mM *L*-[*methyl*-^{13}C]methionine and purification following the established protocol provided ^{13}C-Met-LaL with a yield of 31 mg/L (a total of 46 mg of ^{13}C-Met-LaL was purified from this system). It is expected again that the level of incorporation of ^{13}C-Met approaches 100%. The calculated mass for LaL in which the three methionines have been substituted with [*methyl*-^{13}C]methionine is 17828.1940 g·mol^{-1}, as opposed to 17825.2174 g·mol^{-1} for *wt* LaL. As seen in Fig. 3.22, the mass obtained for ^{13}C-Met-LaL purified from *E. coli* B834/pLR102 is in complete agreement with the expected mass. As was discussed earlier (2.3.2.2), peak "B" in Fig. 3.22 is believed to arise from a species in which the initiator methionyl residue has been removed by methionine amino peptidase. Loss of the N-terminal ^{13}C-Met would reduce the mass of ^{13}C-Met-LaL by 132.1908 mass units to 17696.0032 Da, which is again in complete agreement with the observed mass for this species.

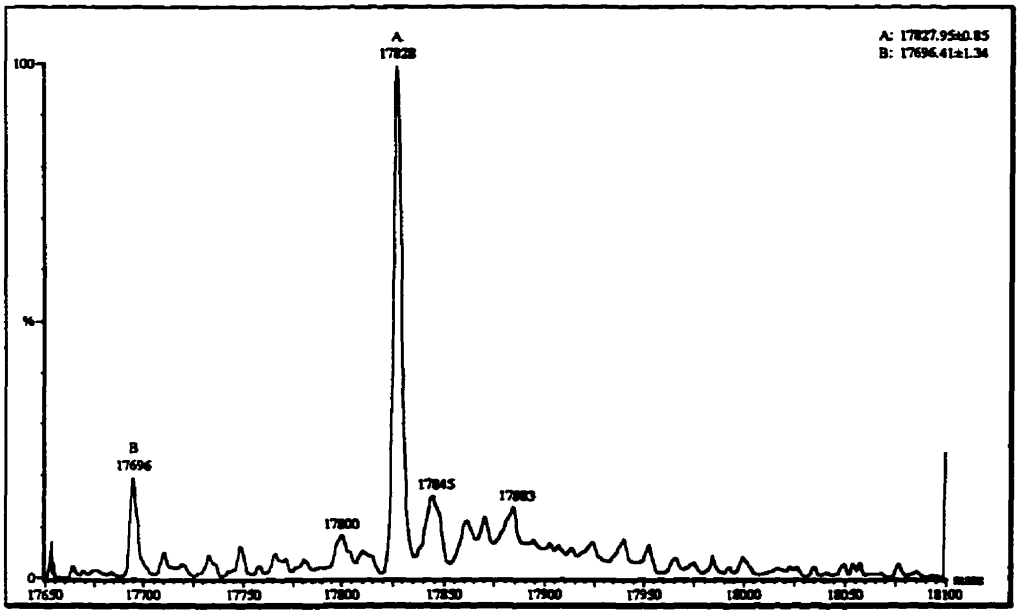


Figure 3.22. ESMS spectra of ^{13}C-Met-LaL purified from *E. coli* B834/pLR102. The sample was desalted by reverse phase chromatography and the spectra was resolved to a resolution of 1 Da.

3.3.2.5. Chitooligosaccharide Protection of Carboxyl Modification

The involvement of carboxyl groups in the catalytic mechanism of LaL was investigated with the carboxyl-group-specific water-soluble carbodiimides EAC (1-(4-azonia-4,4-dimethylpentyl)-3-ethyl-carbodiimide iodide) and EDC (1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide). LaL was incubated with the reagents at pH 5.8 in the absence of an added nucleophile and residual activity over time was determined by the turbidimetric assay. Both EDC and EAC resulted in time- and dose-dependent inactivation of LaL. Semi-logarithmic plots of the time course of inactivation of LaL by various concentrations of EDC and EAC were linear, indicating that the inactivation obeys pseudo-first-order kinetics (Fig. 3.23).

The reader will notice that inactivation of LaL by EDC was performed with a 1.6 mg/mL enzyme solution while a 0.16 mg/mL enzyme solution was employed for the EAC experiments. Initially during method development to establish suitable conditions for inactivation studies involving EDC, higher concentrations of LaL were employed with the intent of nullifying any time-dependent loss in activity observed for more dilute preparations of the enzyme (see 2.3.5.1). However both EDC and EAC were found to promote crosslinking of LaL, possibly from reaction between a nucleophilic group from one enzyme molecule and a carbodiimide activated (i.e. the *O*-acylisourea) carboxyl group of another. In an experiment, LaL (0.016, 0.16 and 1.6 mg/mL) was incubated with EDC or EAC (50 mM) for 1 hr and then samples were subjected to SDS-PAGE (results not shown). It was evident that at the concentration of 1.6 mg/mL, enzyme dimers, trimers and tetramers were formed (dimer > trimer > tetramer) in the presence of the reagents in which the total amount of oligomerization is estimated to be roughly 20-30%. EAC caused slightly more oligomerization than did EDC from the analysis. The formation of oligomeric species was not detected with EDC or was essentially eliminated (less than 5%) with EAC when the reagents were incubated with LaL at the lower concentrations of 0.016 and 0.16 mg/mL. For this reason, subsequent experiments were performed with LaL at a concentration of 0.16 mg/mL. No loss in activity for the control samples of LaL at this concentration was noted over the time period of incubation.

Lambda lysozyme was more rapidly inactivated by EAC than by EDC. When the times required to obtain 50% inactivation are plotted as a function of reagent concentration, the enhanced inactivation of LaL by EAC can be easily appreciated (Fig. 3.24). The superior reactivity of EAC can be attributed to its existence as only the non-cyclic carbodiimide whereas EDC exists as a mixture of two cyclic forms (Tenforde et al., 1972). As illustrated in Scheme 3.3, EDC can undergo tautomerization between the linear

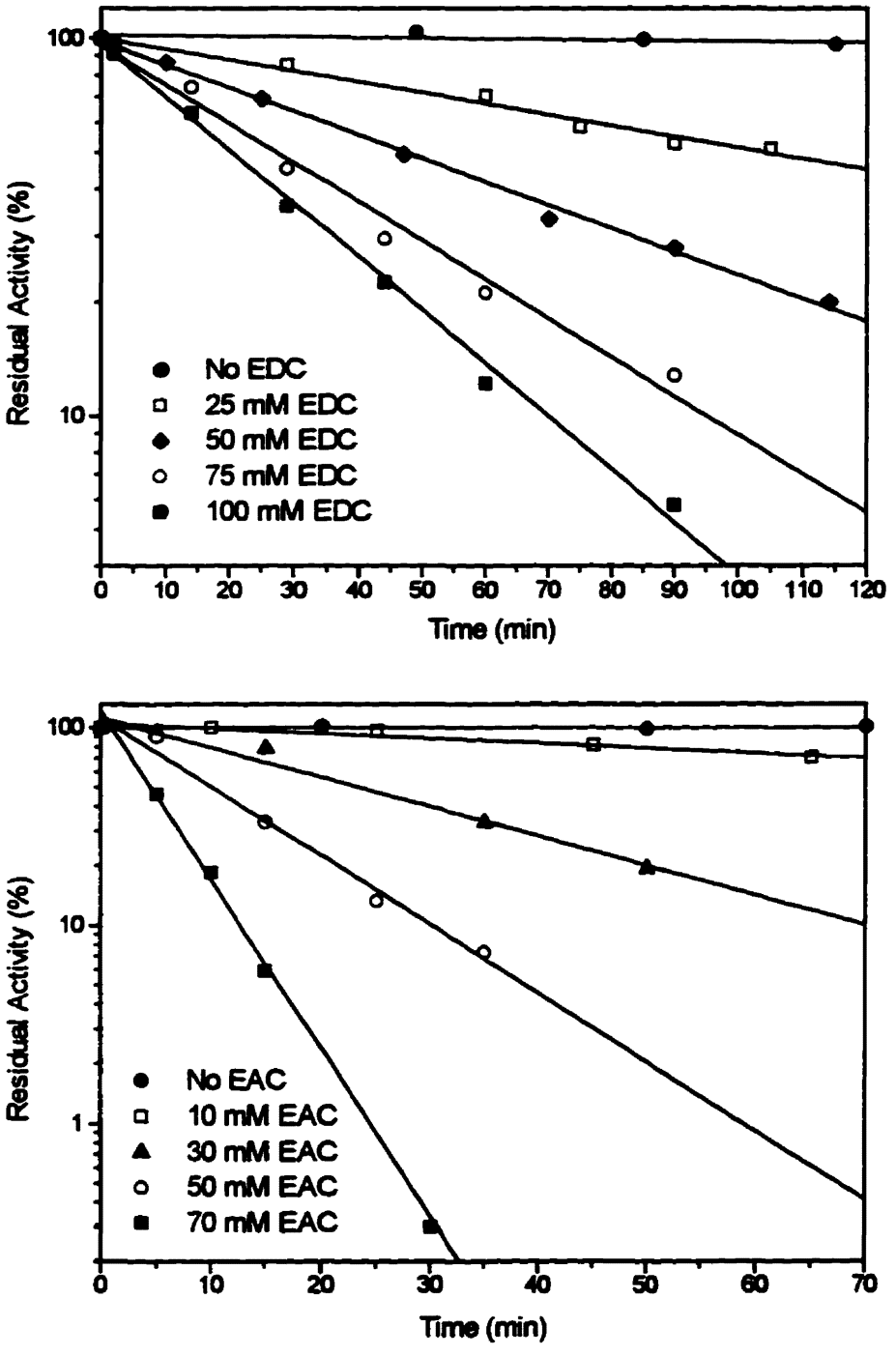


Figure 3.23. Inactivation of LaL by EDC and EAC. LaL was treated with various concentrations of the carbodiimides as indicated and at an enzyme concentration of 1.6 mg/mL with EDC and 0.16 mg/mL with EAC in 100 mM MES, pH 5.8. Samples were removed from the reaction mixtures at the indicated times, diluted with stabilizing buffer and immediately assayed for residual activity.

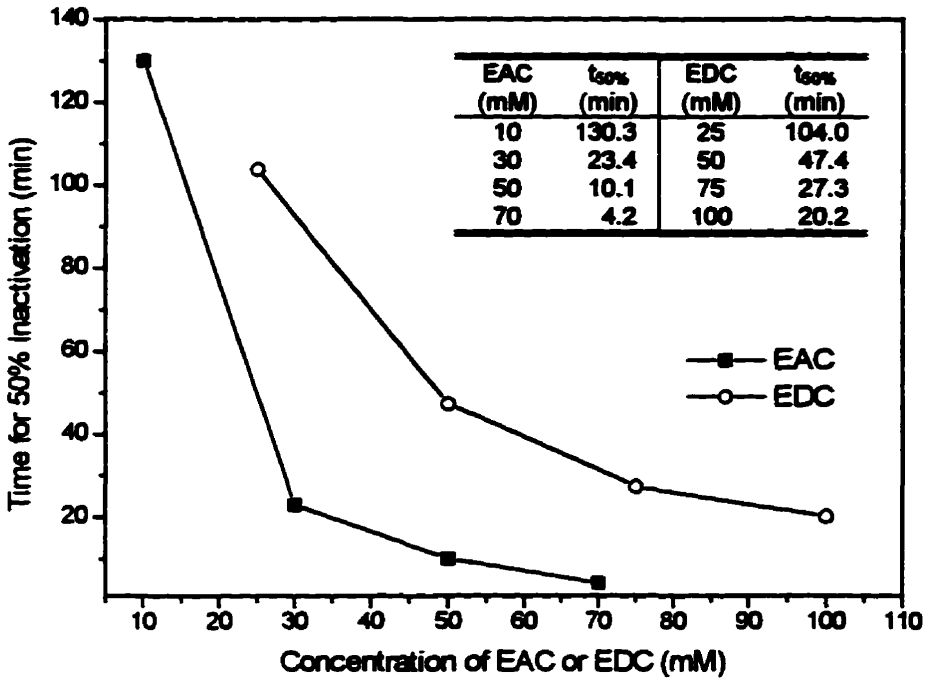
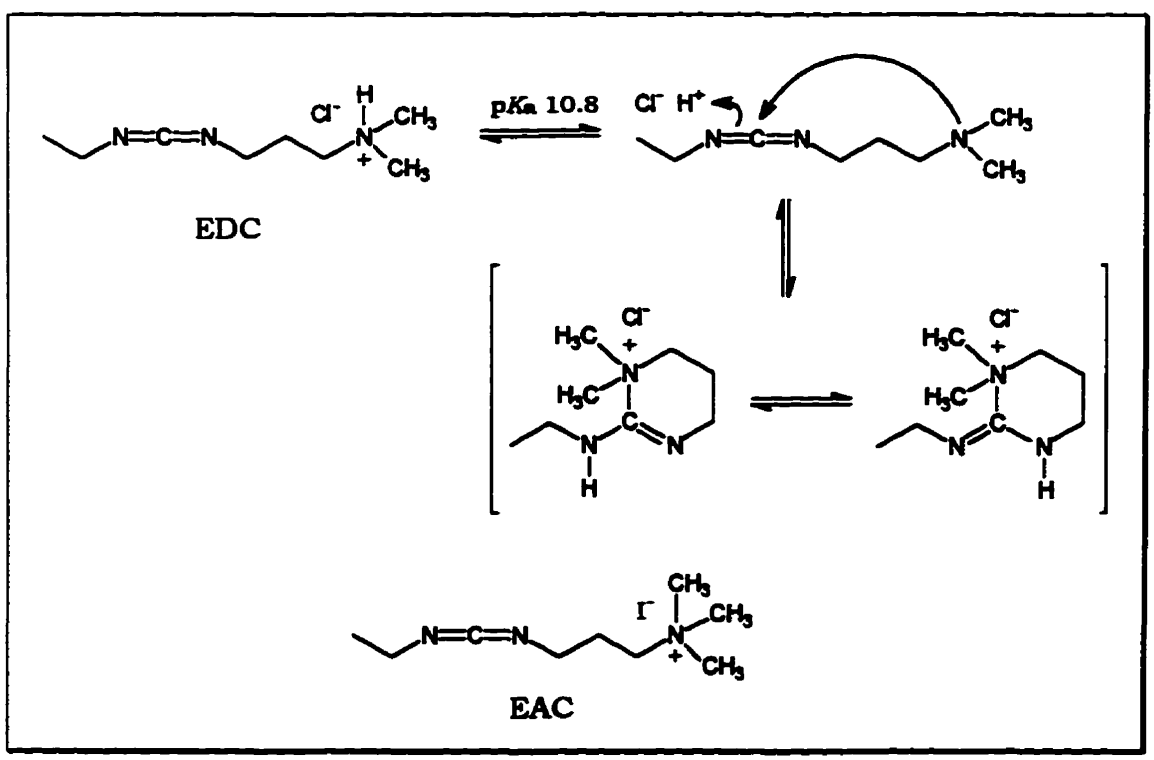


Figure 3.24. Comparison of the inactivation of LaL by EAC and EDC. The times for 50% inactivation of enzyme activity were obtained from the curves in Fig. 3.23 and were plotted as a function of reagent concentration.



Scheme 3.3.

and cyclic form, and at neutral pH, exists predominantly (94%) in the cyclic form (Tenforde et al., 1972). Conversely, EAC, the methiodide analogue derivative of EDC, has no ionizable proton and therefore exists exclusively in the linear carbodiimide form (Scheme 3.3). Since the linear form of the carbodiimide is that which reacts with a carboxyl group, EAC will be more reactive than EDC towards carboxyl group modification. It has also been suggested that rearrangement of the *O*-acylisourea to the stable *N*-acyl derivatives will be favoured over hydrolysis of the *O*-acylisourea if attack by water is hindered because of steric considerations (Timkovich, 1977). As such, the additional methyl group present on EAC may lead to a higher rate of rearrangement over hydrolysis of the *O*-acylisourea formed with EAC than the *O*-acylisourea formed with EDC.

The apparent first-order rate of inactivation of LaL by EAC and EDC permits determination of the order of inactivation with respect to reagent concentration by the method of Levy et al. (1963). The apparent first order rate constant (k') of inactivation will depend on the concentration of the reagent and can be expressed as

$$k' = k\,[\text{EDC or EAC}]^n$$

where n is the average order of the reaction with respect to concentration of reagent. Taking the logarithm of both sides yields

$$\log k' = \log k + n \log [\text{EDC or EAC}].$$

Therefore a plot of log k' against log [EDC or EAC] will have a slope equal to n. Using the reciprocal of the time required to reach 50% inactivation ($1/t_{50\%}$) to represent k', the analysis of the data by this method is presented in Figure 3.25. In each case, the points fit extremely well to a straight line and in which inactivation by EAC yielded a slope of 1.7 while a slope of 1.2 was obtained for inactivation by EDC. The results indicate that at an average of at least one to two molecules of carbodiimide bind to one molecule of LaL to affect inactivation and suggest involvement of carboxyl residues in the mechanism of LaL. The results with EDC and EAC do not rule out involvement of essential tyrosine (Carraway & Koshland, 1968) or cysteine (Carraway & Triplett, 1970) residues as the side chains of these amino acids are also susceptible to modification with carbodiimides to form stable adducts but not as readily as with carboxyl groups.

In order to determine if the carboxyl groups modified by EDC or EAC are located in the active site of LaL, modifications were also performed in the presence of the chitooligosaccharides. The effect of the chitooligosaccharides protection of LaL from

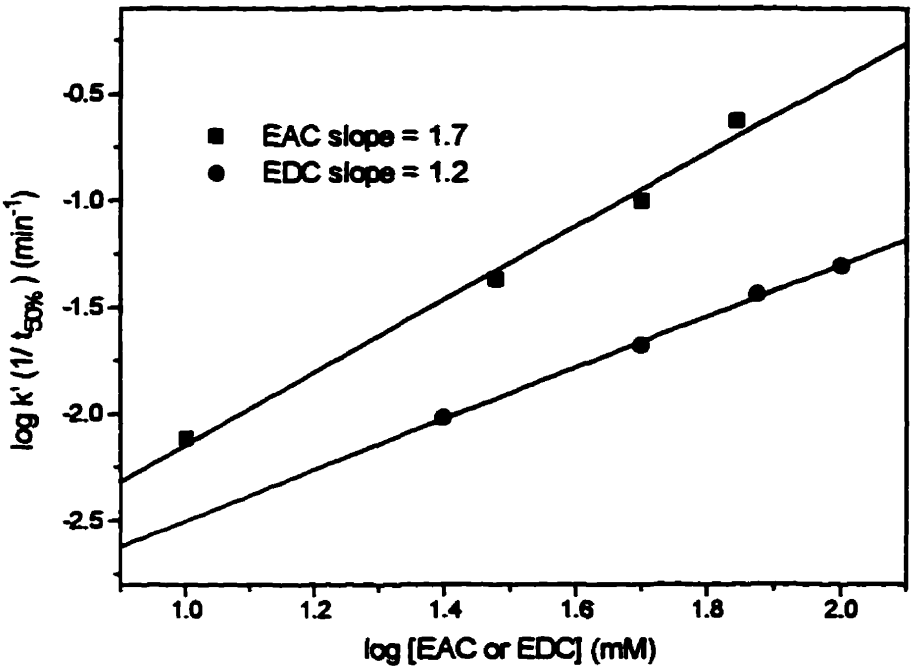


Figure 3.25. Order of the inactivation of LaL by EAC and EDC.

inactivation by EDC is illustrated in Figure 3.26. When compared to the inactivation performed in the presence of EDC only, 20 mM concentrations of $(GlcNAc)_2$ and $(GlcNAc)_3$ were equally and only minimally effective in preventing the inactivation by EDC. However, a substantial improvement in protection is observed when the inactivation was conducted in the presence of $(GlcNAc)_4$ and $(GlcNAc)_5$. Whereas the residual activity of LaL was only 16% after 120 min of reaction in the presence of 50 mM EDC, 87 and 91% enzyme activity remained after the same time of incubation with EDC in the presence of 20 mM $(GlcNAc)_4$ and $(GlcNAc)_5$ respectively (Fig. 3.26). Therefore, the tetra- and pentasaccharide prevent the access of EDC to a particular region(s) of the enzyme in which are located one or more specific acidic residues that are important for the enzyme's bacteriolytic activity. $(GlcNAc)_5$ was less able to prevent inactivation of EAC. In the presence of 50 mM EAC, enzyme activity was reduced to 16% after approximately 19 min, whereas after this same time, the residual activity with $(GlcNAc)_5$ present at 20 mM was itself only 60% (Fig. 3.27).

It was suggested earlier that a disparity may exist between the general interactions of the different chitooligosaccharides. In both the inhibition and fluorescence studies, marked differences in the measured parameters were observed between the group of $(GlcNAc)_n$ where $n \leq 3$ and the group of $(GlcNAc)_n$ where $n \geq 4$. These differences are

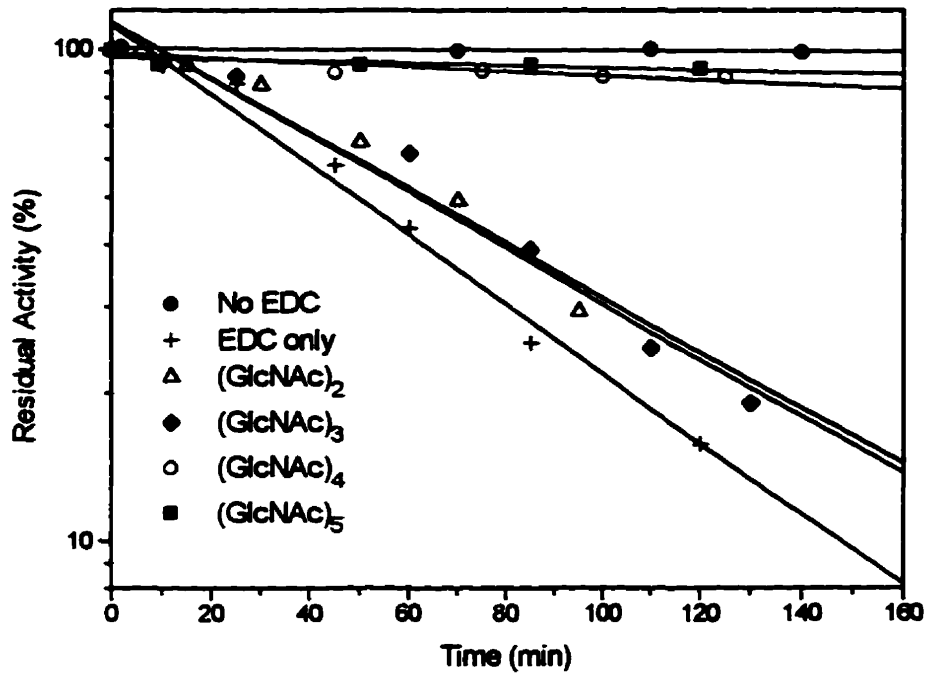


Figure 3.26. Protection of LaL from EDC inactivation by $(GlcNAc)_n$. LaL (1.6 mg/mL) was treated with EDC (50 mM) in the absence and presence of 20 mM $(GlcNAc)_n$ as indicated in 100 mM MES, pH 5.8. At the indicated times, samples were removed from the reaction mixtures diluted with stabilizing buffer and immediately assayed for residual activity.

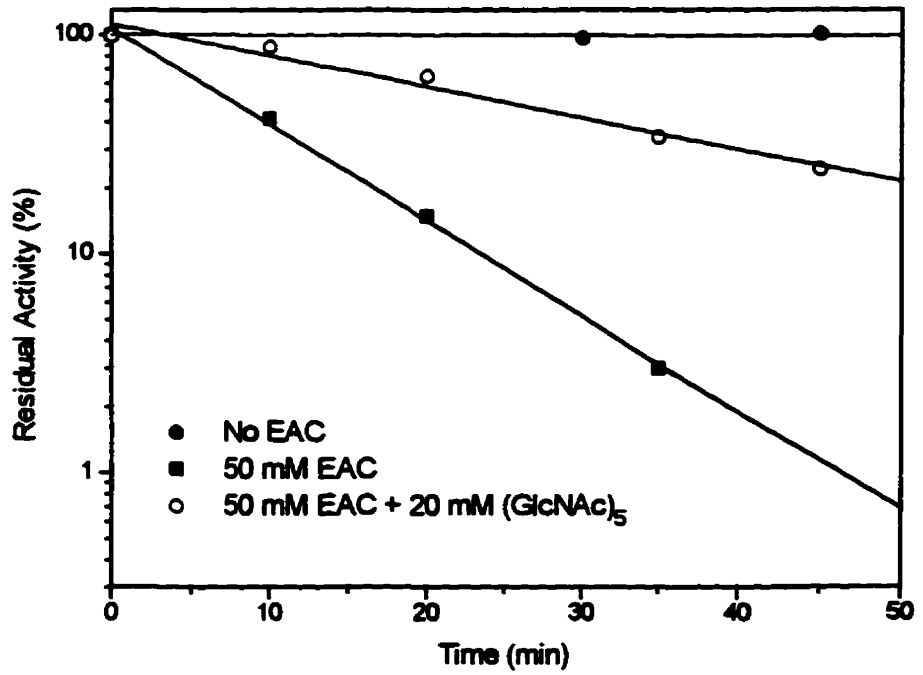


Figure 3.27. Protection of LaL from EAC inactivation by $(GlcNAc)_5$. LaL (0.16 mg/mL) was incubated with EAC as indicated under the same conditions as described in the legend above.

again strikingly apparent in the observed protection of EDC inactivation conferred by these two groups (Fig. 3.26). It appears that the preferred binding regions of the two groups are distinct. The larger saccharides must in essence, engage a larger region of the enzyme upon binding, while the smaller saccharides are assumed to occupy part of but accordingly smaller portions of this region. As indicated by the selective protection, part of this binding region, which is not occupied by $(GlcNAc)_3$ but is with $(GlcNAc)_4$ appears to contain a functionally important glutamyl or aspartyl residue.

It is tempting to speculate as to the nature of the specific residue in LaL whose modification does lead to inactivation. However, there is sufficient experimental support to consider Glu19 as a potential candidate. The involvement of Glu19 and of Asp34 in the mechanism of LaL have been implicated by mutations at these positions (Jespers et al., 1992). If it is assumed that Glu19 acts as the general acid then, in analogy to HEWL and other glycosidases, this group should have an elevated pK_a to fulfill this role. At the pH at which reactions with EDC and EAC were performed (pH 5.8), only a limited number of the 7 glutamic acid and 13 aspartic acid residues of LaL are expected to be protonated. According to a report on the mechanism of the reaction of carbodiimides with proteins, the reaction with a carbodiimide is critically dependent on the protonated form of an acidic residue. A protonated acid would simultaneously protonate the carbodiimide nitrogen atom activating it for nucleophilic attack by the resultant carboxylate anion (Chan et al., 1988). From this reasoning, modification of acidic residues in LaL will be restricted to those having an elevated pK_a and that remain protonated at pH 5.8, possibly the catalytic general acid.

The reduced reactivity for EDC compared to EAC described earlier can be interpreted to indicate that the observed order of inactivation by EDC of 1.2 reflects modification primarily of this assumed catalytic general acid. The higher reactivity of EAC could result in the modification of acidic residues in LaL that are less or even not susceptible to reaction with EDC. Consequently, the larger value of 1.7 determined as the order of inactivation with respect to EAC might reflect modification of the same catalytic acid achieved with EDC in addition to another, slightly less reactive carboxyl group. The reduced ability of $(GlcNAc)_5$ to protect LaL from modification by EAC than by EDC supports this argument. The carbohydrate binding region of LaL may indeed be equally protected from modification by EDC or EAC in the presence of $(GlcNAc)_5$. However, if a second acidic residue which is preferentially reactive to EAC is important for peptidoglycan binding, perhaps for interactions with the peptide portion, then $(GlcNAc)_5$ binding would be ineffective towards protection of this residue. Therefore, the enhanced inactivation observed with EAC could also be explained if functional residues in both the

carbohydrate and peptide binding domains of LaL are modified. The results presented in the following section suggest a possible existence and importance of a peptide binding domain in LaL.

The chemical modification studies, although preliminary, do suggest the involvement of functionally important acidic residues, possibly in the active site of the enzyme. Future efforts should include a study on the dependence of inactivation on pH to identify the pH at which maximal inactivation occurs permitting a more explicit description of the pK_a of the critical reactive groups. This study, however, would also require optimization of the turbidimetric assay with each of the different buffer systems necessary for this work. Investigations including protection with peptides may be useful to identify if EAC does indeed modify a residue(s) outside of the carbohydrate binding region which is important for peptidoglycan binding. A parallel study involving simultaneous protection by peptide and $(GlcNAc)_n$ may also yield additional information.

The classical approach of performing inactivation in the presence and absence of an active site directed inhibitor, followed by proteolysis of the modified proteins and sequencing or ESMS analysis of the peptide fragments is imperative to identify the specific acidic residues involved in the mechanism of LaL. Similar studies involving use of a radioactive carbodiimide would lend itself to the identification of the total number of carboxyls modified in the absence and presence of selected chitooligosaccharides. This study could indicate whether particular acidic residues show preferential protection by $(GlcNAc)_n$ depending on the oligomeric size. If found, this would suggest that different regions of the enzyme are involved in the association with the different oligomers and would substantiate the claims made regarding the enhanced protection observed for the larger saccharides. Preliminary efforts toward these studies have been initiated with EAC but have been impeded by difficulties in purifying the modified proteins from the carbodiimide reagent.

3.3.3. Interactions of Peptides with LaL

The complex and intricate structure of peptidoglycan identified for various microorganisms would suggest that phage lysozymes will demonstrate specific interactions that are unique to their bacterial host's peptidoglycan. LaL may also recognize other components of the peptidoglycan than simply the polysaccharide backbone, and these interactions may be essential for recognition or to impart activity. In order to probe these possible interactions with LaL, peptidyl mimics of the *E. coli* cell wall peptide were prepared. The *E. coli* cell wall peptide is composed primarily of *L*-alanyl-*D*-isoglutamyl-*L*-*meso*-diaminopimelyl-*D*-alanine. The peptides we chose to prepare were of this sequence except that *L*-Lys was used in place of *meso*-DAP since a suitably protected DAP derivative was not commercially available.

3.3.3.1. Peptide Synthesis and Characterization

Peptides were prepared using solid-phase peptide synthesis methodology (Merrifield, 1966) and the BOC functionality as the α-amino protecting group. A summary of the peptide syntheses performed and the recovery of the desired purified peptides is given in Table 3.12. The first step in the synthesis was the introduction of the C-terminal *D*-Ala onto the respective resins. Different resins were used with the intention of possibly identifying one that gave the best and highest yields of the desired product peptides. Introduction of BOC-*D*-Ala using diisopropylcarbodiimide and HOBt onto the hydroxymethyl resin resulted in a higher substitution (based on the theoretical potential substitution of the resins) than that obtained for the introduction of the cesium salt of BOC-*D*-Ala onto the Merrifield (chloromethyl) and PAM (bromomethyl) resins by the method of Gisin (1973). The coupling yields by the latter method were unexpectedly low considering the near quantitative yields reported previously using this method (Gisin, 1973; Sparrow, 1976). It is possible that the attachment yields are somewhat lower than expected if the theoretical substitutions (provided by the manufacturers) given for the commercial resins were overestimated.

Peptide elongation was achieved under standard conditions involving BOC deprotection by TFA, neutralization of the peptidyl-resin with DIPEA and coupling with an excess of the succeeding BOC-amino acid following activation of its carboxylic acid. Both BOP/HOBt and DCC/HOBt were utilized as the carboxyl activating agents in the respective syntheses. Coupling was found to be more efficient using DCC/HOBt. For the synthesis of the tetrapeptide **3** using DCC/HOBt, the coupling reactions went to completion (as monitored by disappearance of free amino groups) in less than 1 hr.

Table 3.12. Summary of Peptide Syntheses.

Peptide and MW (Da)	Resin[a]	Initial Scale[b]	Attachment of BOC-*D*-Ala to Resin		Coupling Method	Fmoc Analysis Yield[c,d]	Purified Peptide Yield[c,e]
			Method	Yield[c]			
1 *L*-Ala-*D*-*iso*-Glu-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala (710.78)	Hydroxymethyl	4.00	DIPC/ HOBt	2.85 (71%)	BOP/ HOBt	2.23 (56%)	0.91 (23, 41%)
1 *L*-Ala-*D*-*iso*-Glu-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala (710.78)	Merrifield	4.52	Cs Salt	1.66 (37%)	BOP/ HOBt	1.73 (38%)	0.53 (12, 31%)
2 *L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala-*D*-Ala (488.54)	-	-	-	-	-	-	88%
3 *L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala (417.46)	PAM	0.68	Cs Salt	0.32 (47%)	DCC/ HOBt	0.37 (54%)	0.13 (19, 35%)

[a] The resin used for the synthesis.

[b] The scale of the synthesis (in mmol) based on the amount and theoretical substitution of the resin used.

[c] Yields are given in mmol. The percentage yield in parantheses is based on the initial scale.

[d] Quantitated during the synthesis at the stage of BOC-*L*-Lys(Fmoc)-*D*-$(Ala)_n$-Resin.

[e] First value in parentheses is based on the initial scale; the second value is based on the yield obtained from the Fmoc analysis.

For peptide **2**, the yield represents that obtained following Fmoc removal and purification from peptide **1**.

Refer to section 3.2.5.1 for other details concerning the methodology employed. DIPC is diisopropylcarbodiimide.

Synthesis of peptide **1** using BOP/HOBt required greater time (2-3 hr) to reach completion of the coupling reactions. For the two syntheses of peptide **1** using BOP/HOBt, the coupling of BOC-*L*-Lys(Fmoc)-OH did not go to completion even under extended periods of coupling (> 15 hr). Inefficient coupling of the Lys derivative could result in a reduction in the final yield of peptide **1** due to loss of the alanine residues as 2,5-diketo-3,6-dimethylpiperazine (Fig. 3.28 A, R_1=R_2=CH_3). This loss through diketopiperazine cleavage (Stewart & Young, 1984) will reduce the total substitution on the resin and may have been significant owing to the extended reaction time required for the coupling of the Lys derivative.

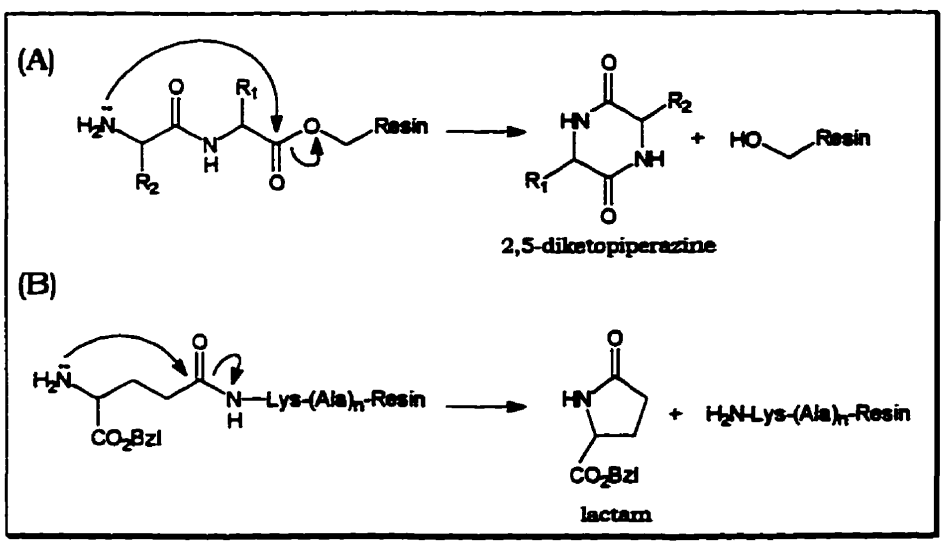


Figure 3.28. Possible side reactions reducing yields during synthesis of peptides.

In each synthesis, the final yields of purified peptide obtained were somewhat disappointing. The low yields were not a result of the loss of product during purification. During purification of the peptides, all fractions collected during chromatography were routinely analyzed by ^{1}H NMR and MALDI MS or ESMS. This analysis revealed that very minimal to no product was lost in fractions other than those pooled containing the product peptides. Comparison of the yield of purified peptides obtained to that expected from Fmoc analysis on the peptidyl-resins (Table 3.12) would suggest that the low recovery of the peptides may have resulted from their inefficient cleavage from the resins and/or inefficient diethyl ether precipitation following their cleavage. The latter may again be significant for peptide **1** in which the Fmoc functionality could increase its diethyl

ether solubility. Another source of the low yields may have resulted during the final coupling of BOC-*L*-Ala to the *iso*-glutamic-peptidyl-resin in which the glutamic acid is cleaved from the peptidyl-resin as the lactam (Fig. 3.28 B). Indeed, the analyses of the crude peptides (i.e. prior to purification) indicated significant presence of peptides that were lacking the Glu residue (see also discussion below for peptide **1**).

Pentapeptide **1** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala) was prepared retaining the Fmoc protecting group on the ε-amino group of Lys. This strategy was undertaken such that peptide **1** could be condensed at its free α-amino group to muramic acid derivatives in our efforts to synthesize a chromogenic substrate for LaL (see 3.3.4). The difficult solution synthesis of a fully protected peptide of the same sequence as peptide **1** was first reported by Tesser & Nivard (1964). However in this synthesis, the peptide bond between the *D*-Glu and *L*-Lys was prepared with the α- and not the γ-carboxyl of Glu. More recent solution syntheses of the peptides BOC-*L*-Ala-*D*-*iso*-Glu(αNH_2)-*L*-Lys(CBz)-*L*-Ala (Lefrancier et al., 1978) and BOC-*L*-Ala-*D*-iso-Glu-*L*-Lys(CBz)-*D*-Ala-*D*-Ala-OBzl (Adam, 1992) have also appeared.

Following its cleavage from the resin and precipitation, peptide **1** could be purified from the crude precipitate by successive Lobar and HPLC RP-C18 chromatography. When analyzed by capillary electrophoresis, the chromatogram of the purified peptide exhibited a single, somewhat broad peak (Fig. 3.29 A). MALDI MS analysis of the same sample revealed a major component with mass of 711.7 Da (M_1+1 ion), consistent with the calculated mass for peptide **1** of 710.78 Da (Fig. 3.30 A). The mass spectrum also indicated the minor presence of components consistent for an Ala deletion (i.e. 2 Ala, 1 Glu, and 1 Lys(Fmoc), M_2+1 640.7; calcd 639.70), a Glu deletion (i.e. 3 Ala, 1 Lys(Fmoc), M_3+1 582.2; calcd 581.66 Da) and both an Ala and Glu deletion (i.e. 2 Ala, 1 Lys(Fmoc), M_4+1 511.2; calcd 510.59). If these minor components are present, then they must co-migrate with peptide **1** by CZE. Their co-migration with peptide **1** over the RP-C18 supports would be expected since their retention on these resins would be dominated by the hydrophobicity of the Fmoc group. The ^{1}H NMR spectrum of peptide **1** (Fig 3.31) strongly corroborates the structure and the integration observed (not shown in Fig. 3.31) suggested that any contaminating deletion peptides are present in only very small amounts. The purity obtained for peptide **1** was more than satisfactory for its intended use towards subsequent synthesis of substrates.

Figure 3.29. Capillary Electrophoresis Chromatograms of Peptides **1**, **2**, and **3**.

(A) Chromatogram of peptide **1** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala).

Sample:	≈ 1 mg/mL dissolved in Buffer diluted 1:10 with MQW
Capillary:	24 cm × 25 μm I.D. coated; 20 °C
Buffer:	BioRad phosphate 0.1 M, pH 2.5
Injection:	20 psi × sec
Running:	14 kV
Polarity:	positive to negative
Detection:	210 nm

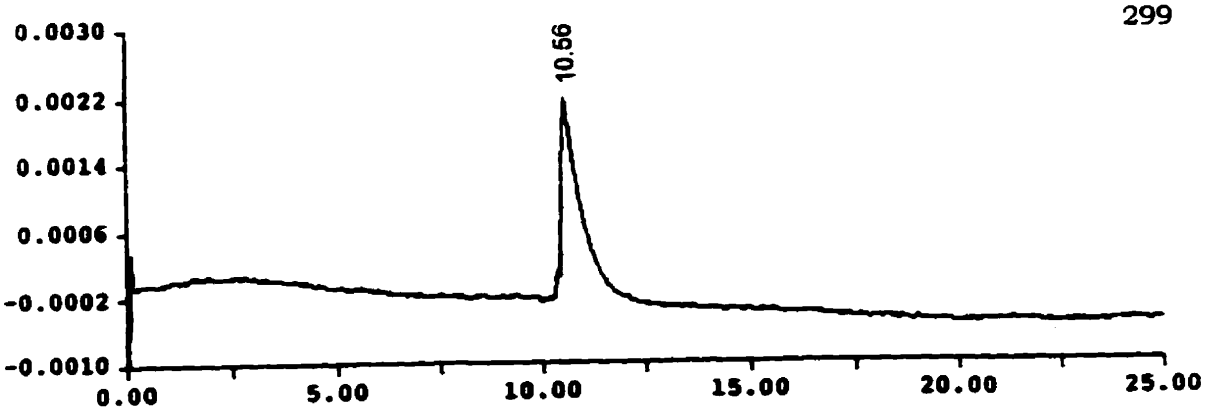
(B) Chromatogram of peptide **2** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala-*D*-Ala).

Sample:	10 mM dissolved in MQW
Capillary:	50 cm × 50 μm I.D. coated; 20 °C
Buffer:	BioRad phosphate 0.1 M, pH 2.5
Injection:	5 psi × sec
Running:	20 kV
Polarity:	positive to negative
Detection:	200 nm

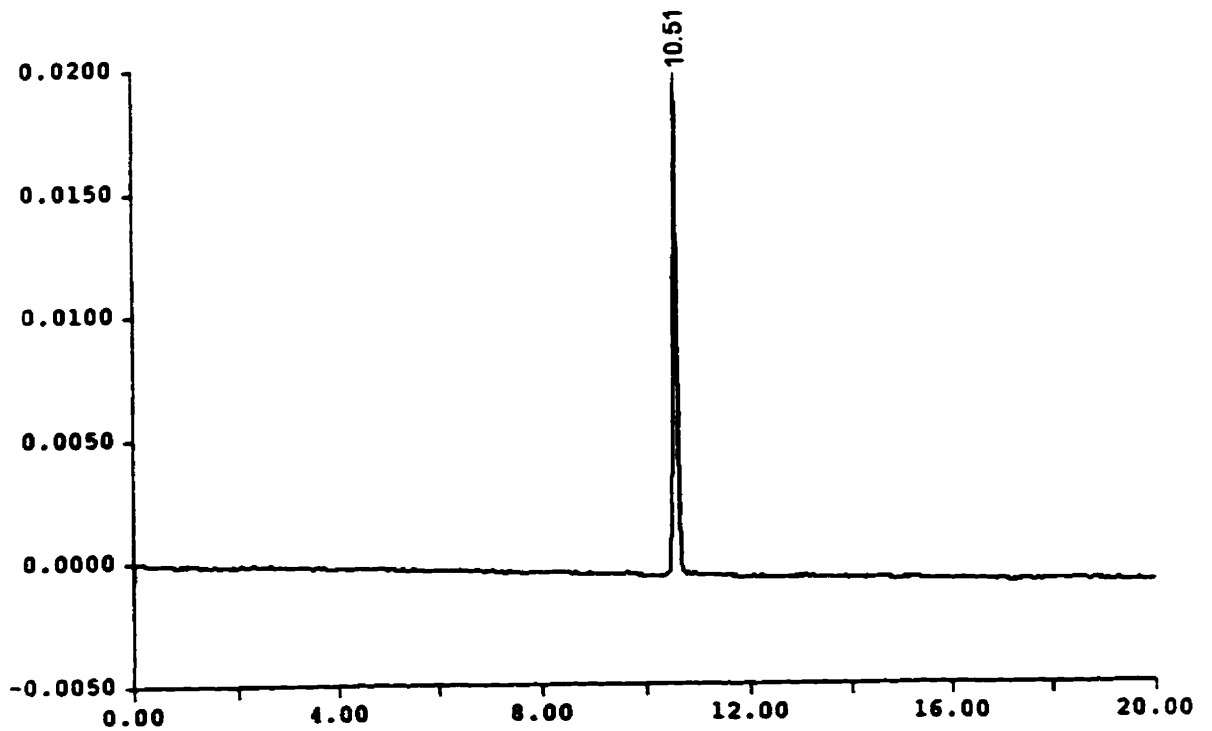
(C) Chromatogram of peptide **3** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala).

Sample:	10 mM dissolved in Buffer diluted 1:10 with MQW
Capillary:	24 cm × 25 μm I.D. coated; 20 °C
Buffer:	BioRad phosphate 0.1 M, pH 2.5
Injection:	5 psi × sec
Running:	5 or 10 kV (as indicated)
Polarity:	positive to negative
Detection:	200 nm

(A)


10.56
0.0030
0.0022
0.0014
0.0006
-0.0002
-0.0010
0.00
5.00
10.00
15.00
20.00
25.00

(B)


10.51
0.0200
0.0150
0.0100
0.0050
0.0000
-0.0050
0.00
4.00
8.00
12.00
16.00
20.00

(C)

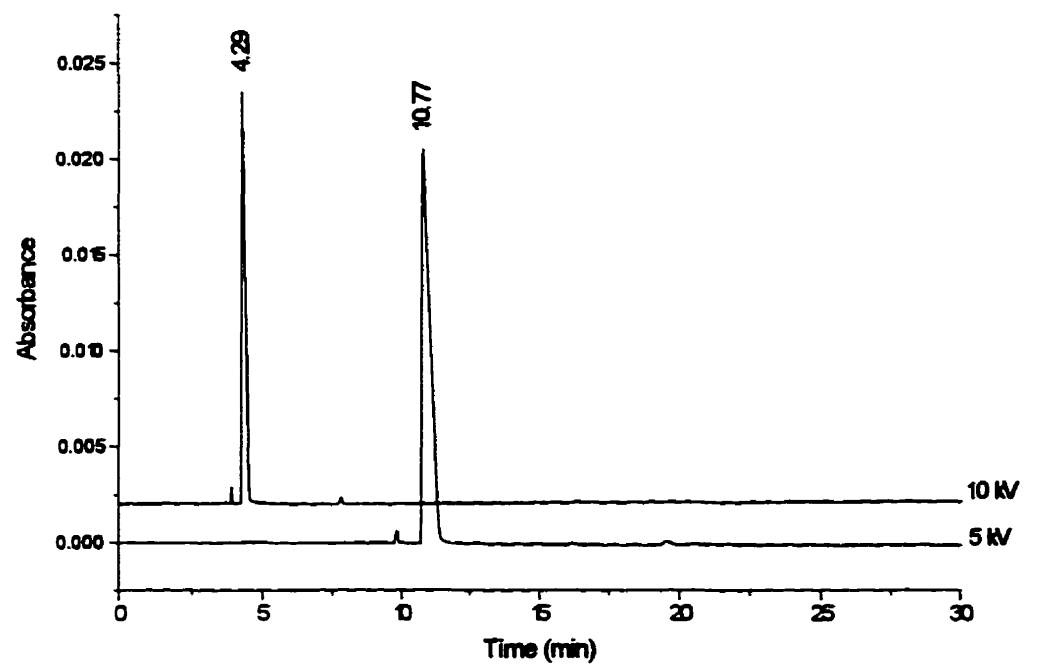

4.29
10.77
0.025
0.020
0.015
0.010
0.005
0.000
Absorbance
10 kV
5 kV
0
5
10
15
20
25
30
Time (min)

Figure 3.30. Mass Spectra of Peptides **1**, **2**, and **3**.

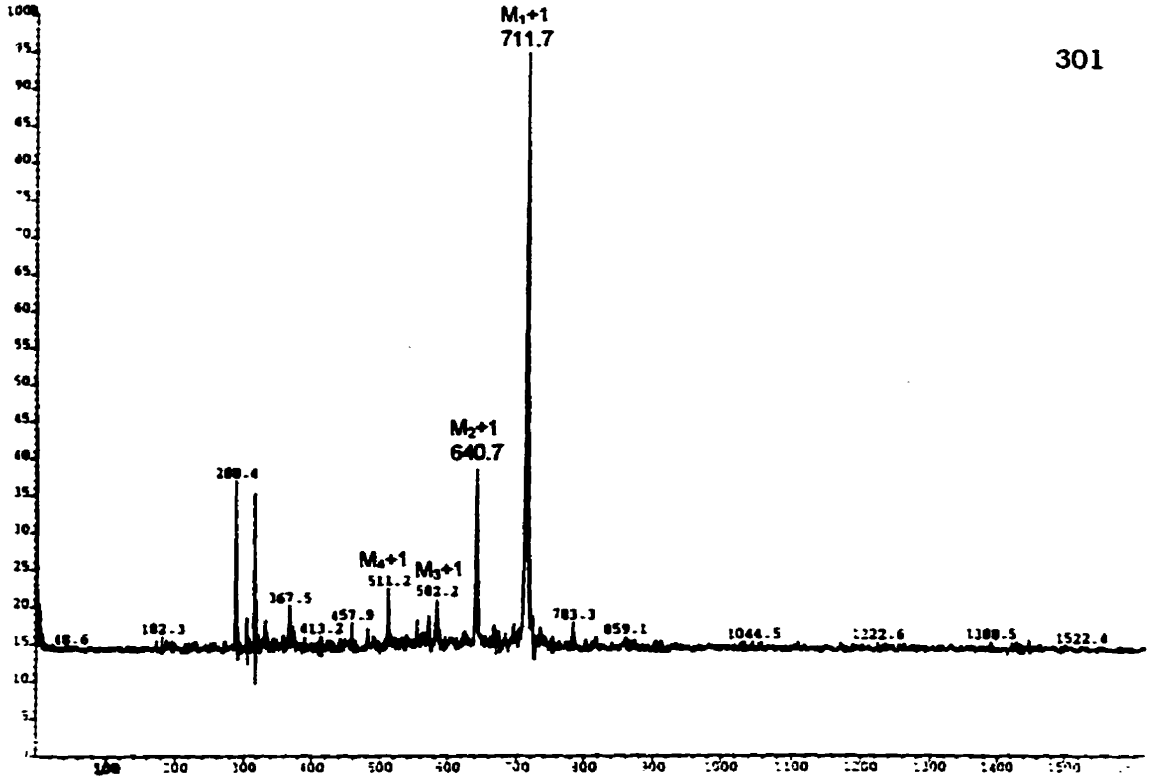
(A) MALDI MS spectrum of peptide **1** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala). The calculated mass is 710.78 Da. The nature of the other peaks (M_2, M_3 and M_4) are described in the text. The spectrum was acquired using a laser energy of 2 (coarse) and 110 (fine), 400 Da matrix suppression and was externally calibrated against the (M+1) and (M+2) molecular ions of Kemptide (MW 771.9 Da; M+1 = 772.9 Da, M+2 = 387.0 Da).

(B) MALDI MS spectrum of peptide **2** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala-*D*-Ala). The calculated mass is 488.54 Da. The spectrum was acquired using a laser energy of 2 (coarse) and 105 (fine), 400 Da matrix suppression and was externally calibrated against Kemptide as described in (A).

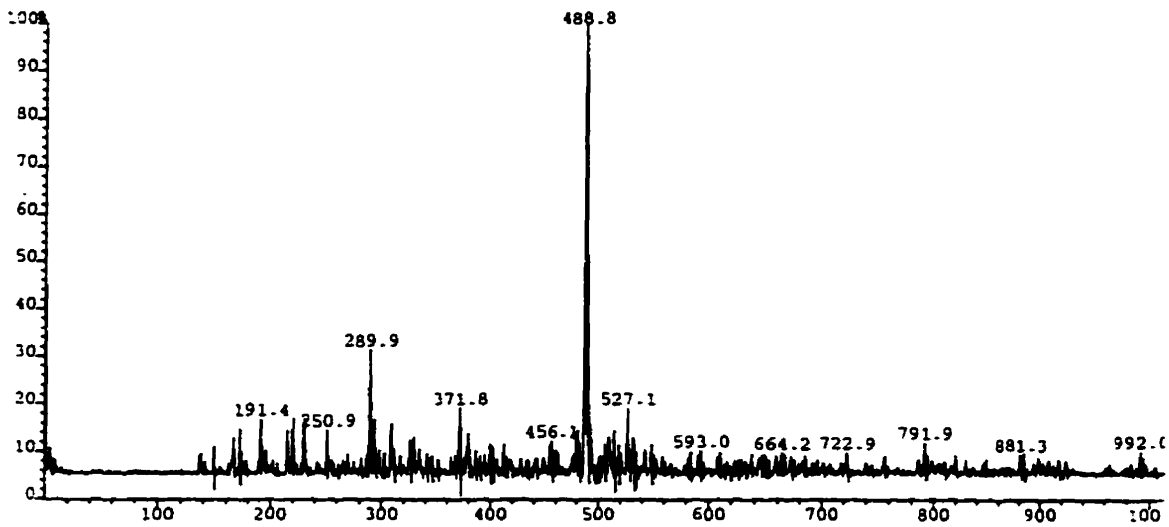
For the MALDI MS spectra, the mass values (in Da) represent the (M+1) molecular ions and therefore, a mass of 1 Da must be subtracted.

C) Transformed ESMS spectrum of peptide **3** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala). The calculated mass is 417.46 Da.

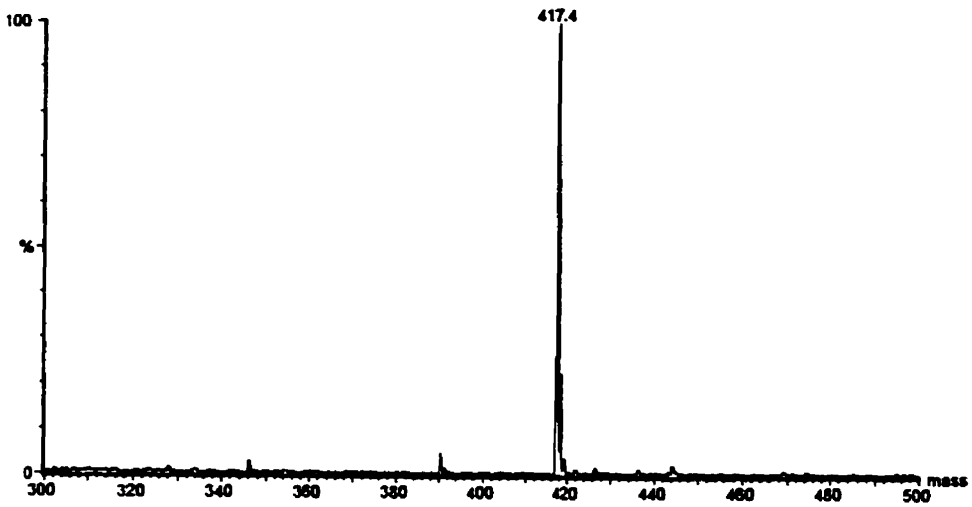
(A)


M1+1
711.7
M2+1
640.7
M4+1
511.2
M3+1
582.2
288.4
367.5
457.9
413.2
182.3
783.3
859.1
1044.5
1222.6
1388.5
1522.4

(B)


488.8
289.9
191.4
250.9
371.8
527.1
456.1
593.0
664.2
722.9
791.9
881.3

(C)


417.4
%
mass

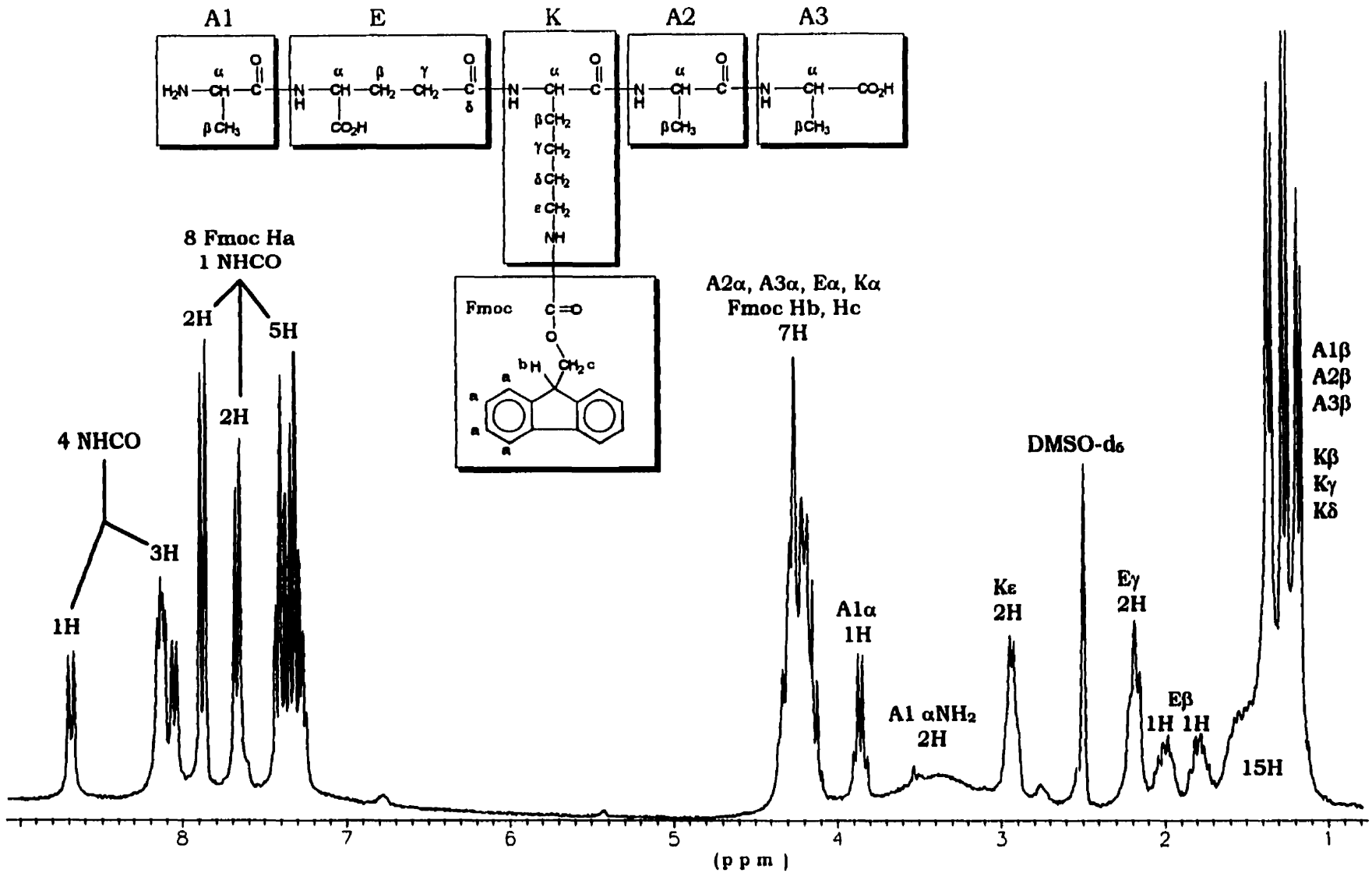


Figure 3.31. 250 MHz ^{1}H NMR spectrum of peptide **1** taken in DMSO-d_6.

Pentapeptide **2** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala-*D*-Ala) was obtained following the removal of the Fmoc group from peptide **1** using 20% piperidine in DMF. Peptide **2** could be purified (with 88% recovery) by successive cation exchange and RP-C18 chromatography. Purification was not possible by only reverse phase chromatography as it appeared that piperidine co-eluted (perhaps as a non-covalent adduct) with peptide **2** under various chromatographic conditions attempted. A high level of purity for peptide **2** is suggested following its analysis by CZE (Fig. 3.29 B). The MALDI MS spectrum (Fig. 3.30 B) provided a mass in excellent agreement with the calculated mass of 488.54.

The 500 MHz ^{1}H NMR spectrum of peptide **2** and the assignment of the individual protons is shown in Figure 3.32 (A and B). The proton assignments were made following the acquisition of two dimensional proton COSY, TOCSY and NOESY NMR spectra of the peptide (results not shown). The COSY and TOCSY spectra allowed the establishment of the connectivities within the individual residues. The N-terminal alanine (A1, Fig. 3.32) was easily identified since, it lacks an amide, no correlation existed between its αH to an amide resonance. The NOESY experiment confirmed the correct sequence. Strong NOE correlations were identified between E_{NH}-$A1_{H\alpha}$ and E_{NH}-$A1_{\beta CH3}$. Similarly, correlations between K_{NH}-$E_{CH2\gamma}$ (which confirms the *iso*-peptide bond linkage between the Glu and Lys), $A2_{NH}$-$K_{H\alpha}$ and $A3_{NH}$-$A2_{H\alpha}$ were sufficient to verify the sequence. The correct sequence and *iso*-Glu peptide bond is also expected from the order of and the protected amino acids utilized in the synthesis. The individual carbon assignments were also determined through the acquisition of [^{1}H-^{13}C]HMQC and [^{1}H-^{13}C]HMBC NMR spectra (results not shown). A summary of the chemical shifts for peptide **2** is given in Table 3.13. Since peptide **2** was obtained from peptide **1**, the NMR structural analysis described for **2** also provides confirmation of the correct structure and sequence for peptide **1**.

Table 3.13. ^{1}H and ^{13}C resonance assignments (in ppm) for peptide **2** in 90% H_2O/10% D_2O. Refer to Fig. 3.32 for the specific assignments of the individual H and C atoms.

	NH	Hα	Cα	Hβ	Cβ	Hγ	Cγ	Hδ	Cδ	Hε	Cε	αC=O
Ala-1	-	4.07	51.7	1.49	19.4							171.7
Ala-2	8.38	4.29	51.5	1.33	19.3							176.2
Ala-3	8.23	4.27	51.8	1.37	18.9							178.7
Glu	8.60	4.29	55.7	1.94 2.13	29.3	2.35	34.0		177.8			177.8
Lys	8.23	4.19	56.3	1.70 1.75	33.0	1.39	24.6	1.63	28.9	2.94	42.0	175.3

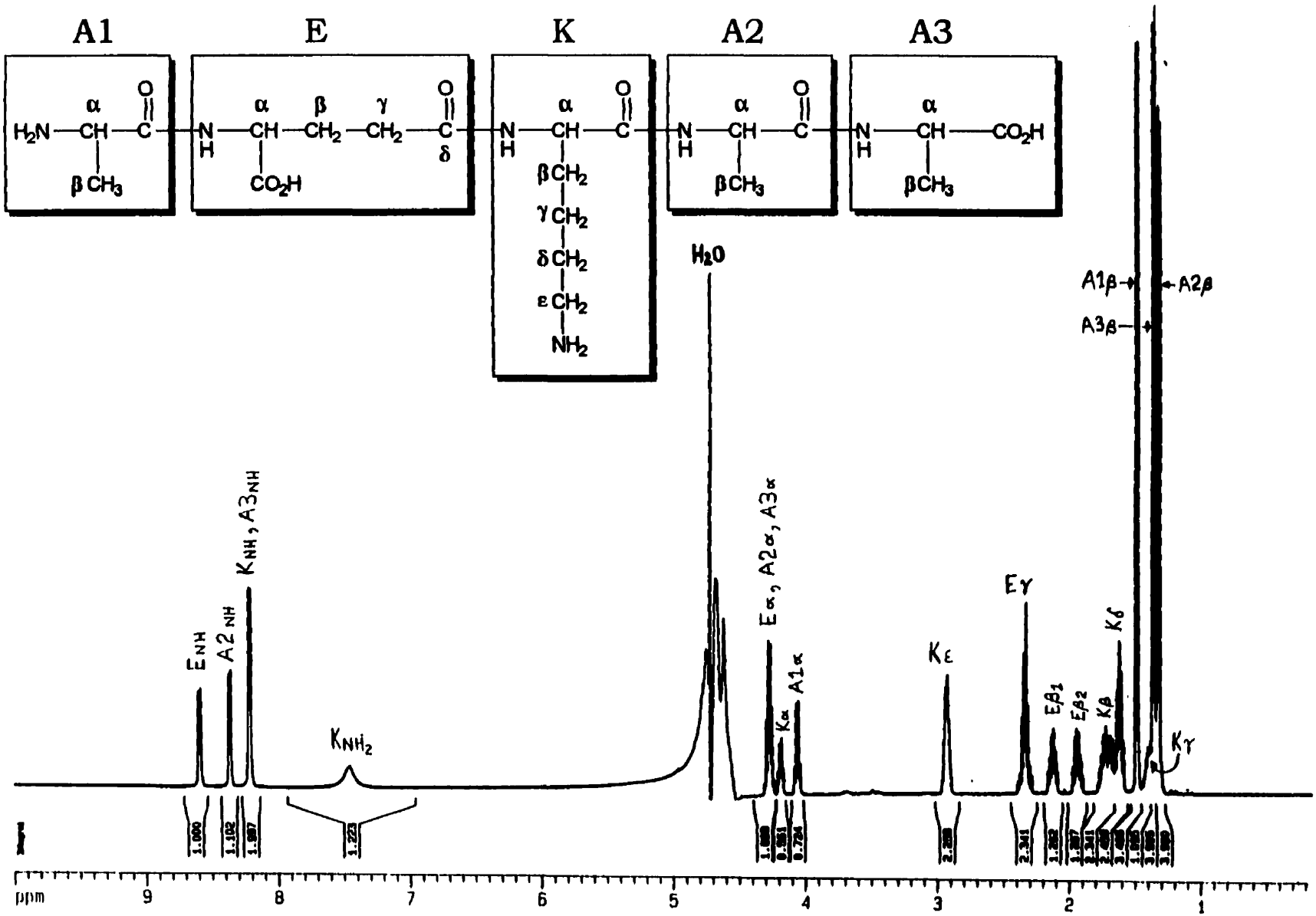


Figure 3.32. (A) 500 MHz ^{1}H NMR spectrum of peptide **2** taken in 90% H_2O/10% D_2O. The spectrum was acquired with suppression of the water signal. An expansion of the spectrum is shown on the following page in Fig. 3.32 (B).

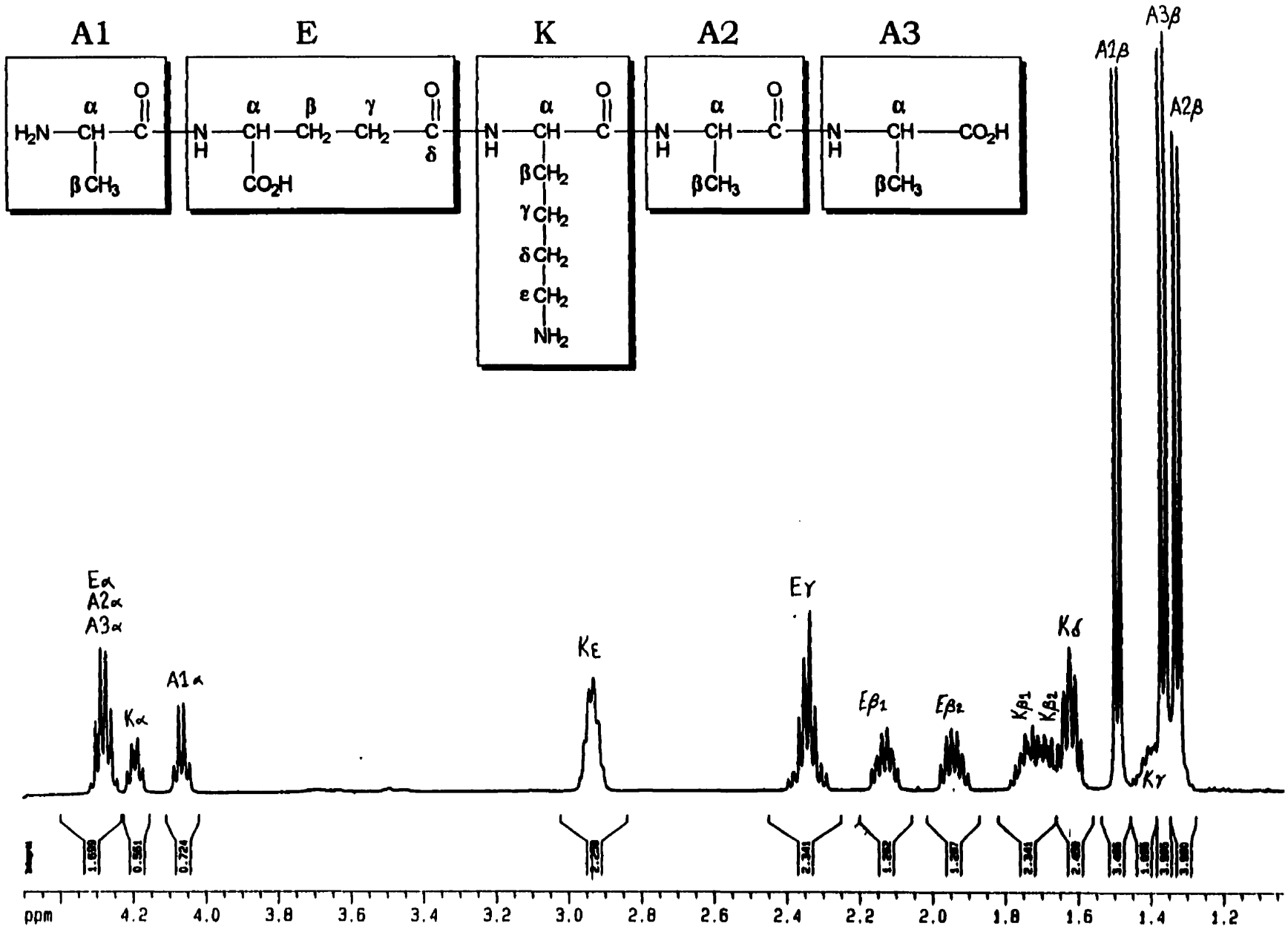


Figure 3.32. (B) Expansion of the spectrum shown in (A).

Tetrapeptide **3** (*L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala) was obtained fully deprotected, removing the Fmoc group while the peptide was still bound to the resin. Following its cleavage from the resin and precipitation, peptide **3** could be purified by RP-C18 chromatography. The small and relatively hydrophilic nature of this peptide resulted in very weak retention on the C18 support. For this reason, the peptide was first chromatographed utilizing eluents containing 0.1% TFA. The peptide eluted in the void fraction separating it from the more hydrophobic contaminants. Reapplication of the void fraction in mobile phases now containing 0.25% TFA, increased the retention of the peptide which permitted its separation from contaminating peptides. A similar strategy (using 0.2% TFA in the eluents) was employed during the RP-C18 chromatography of peptide **2**. The characterization of peptide **3** by CZE (Fig. 3.29 C) and ESMS (Fig. 3.30 C) indicate a high level of purity for the peptide with the expected mass. The ^{1}H and ^{13}C NMR spectra are shown in Fig. 3.33 (A and B). The assignments given in Fig. 3.33 were made based on those established for peptide **2** and are believed to be correct.

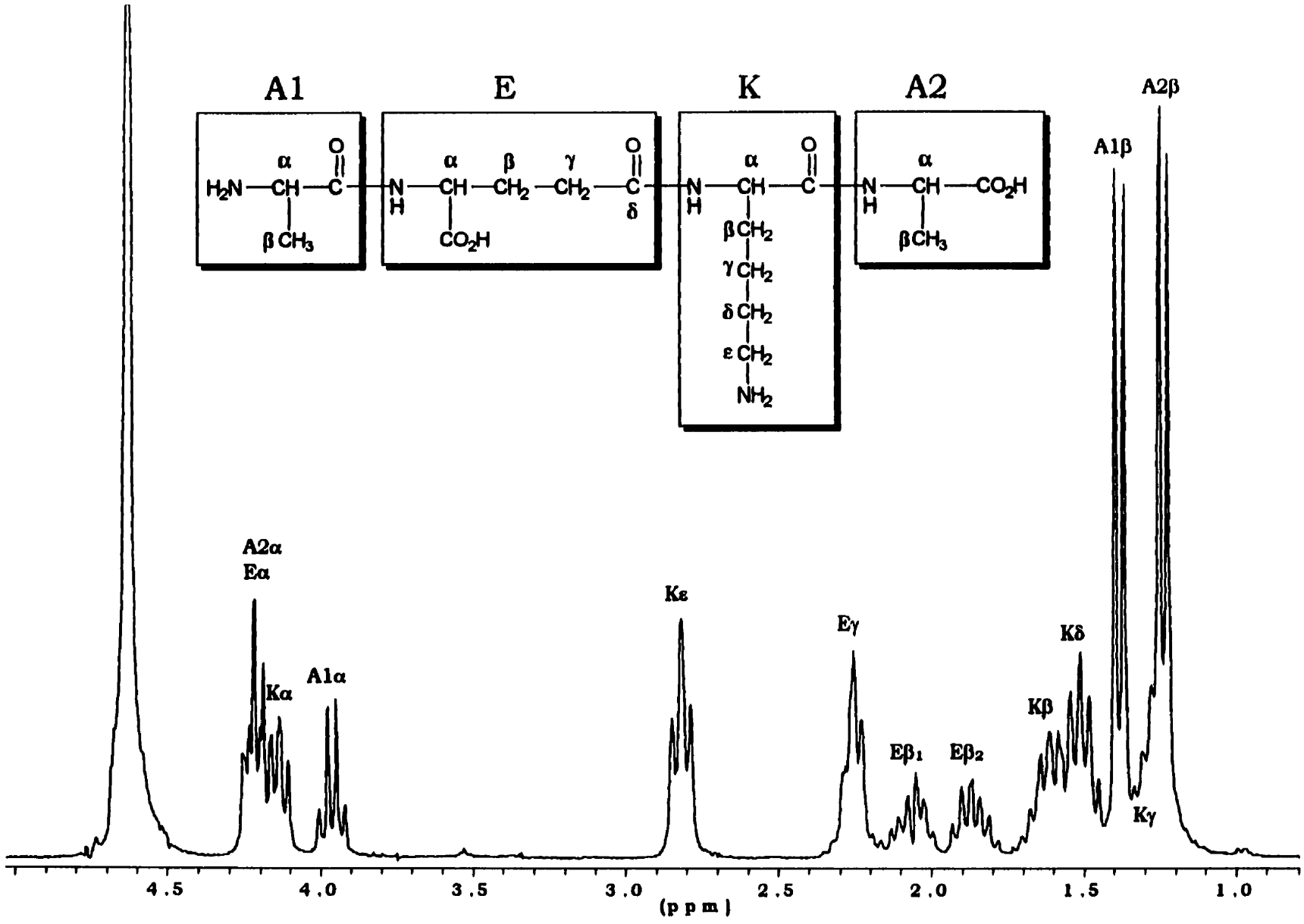


Figure 3.33. (A) 250 MHz ^{1}H NMR spectrum of peptide **3** taken in D_2O. The assignments were made based on those determined for peptide 2 (see Table 3.13)

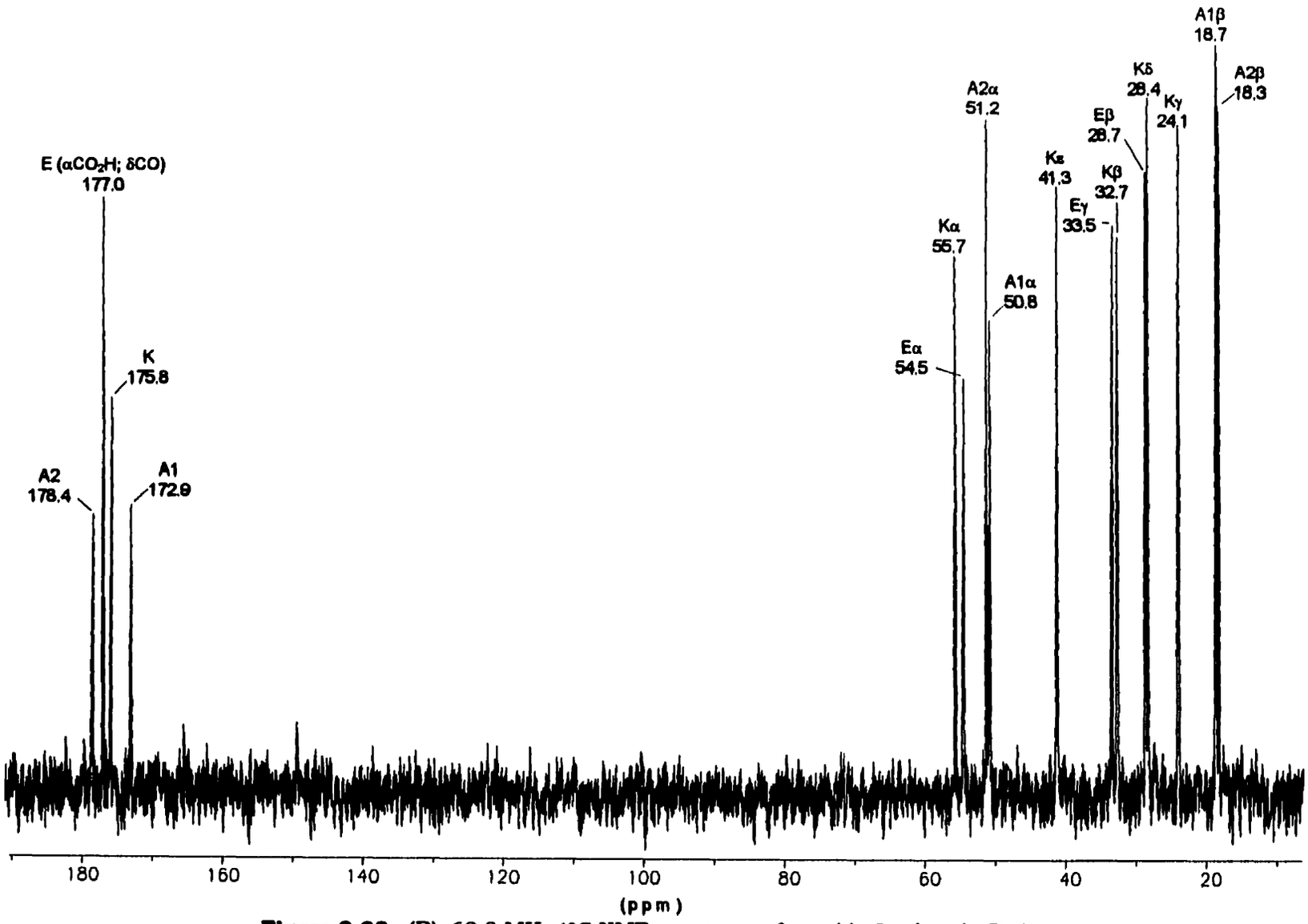


Figure 3.33. (B) 62.2 MHz ^{13}C NMR spectrum of peptide **3** taken in D_2O. The assignments were made based on those determined for peptide 2 (see Table 3.13).

3.3.3.2. Evaluation of the Interactions of Peptides with LaL

Having prepared and purified the *E. coli* cell wall peptidyl mimics, we were able to explore the possible interactions of these peptides with LaL. Very interesting results were obtained when the effect of the peptides on the bacteriolytic activity of LaL were examined. Unlike the inhibitory properties observed for the chitooligosaccharides, both peptide **2** and **3** were found to *increase* the activity of LaL using the turbidimetric assay (Table 3.14). If LaL does have a binding region involved in interactions with the cell wall peptide, then binding of the peptidyl-mimics would be expected to inhibit peptidoglycan binding and reduce the observed activity. It could be argued that possibly peptide **2** and **3** themselves could in some way promote the lysis of the substrate cells. However, peptide **3** was found to decrease the background clearing of the cells by approximately 12 and 20% at 5 mM respectively (see 5 mM #1 and #2) and by 40% at 10 mM while peptide **2** at 2 mM decreased background lysis by 25% when compared to the substrate cells in the absence of the peptide (compare values for a and b, Table 3.14). Since the ionic strength of the assay buffer will increase with increasing concentrations of the peptides, this will render the cells less osmotically sensitive. These observations are consistent with the ionic strength effects on the assay described earlier (see 3.3.1).

Table 3.14. Effect of peptide **2** and **3** on the bacteriolytic activity of LaL.

	Decrease in Turbidity of the Substrate Cells (SC)†			
	3 (5 mM #1)	**3** (5 mM #2)	(**3** 10 mM)	**2** (2mM)
a) SC only	0.0077 ± 0.0009	0.0108 ± 0.0013	0.0090 ± 0.0024	0.0044 ± 0.0015
b) SC + peptide only	0.0068 ± 0.0010	0.0085 ± 0.0009	0.0051 ± 0.0018	0.0033 ± 0.0019
c) SC + LaL	0.0409 ± 0.0027	0.0730 ± 0.0022	0.0699 ± 0.0033	0.0583 ± 0.0052
d) SC + LaL + peptide	0.0463 ± 0.0007	0.0848 ± 0.0051	0.0945 ± 0.0016	0.0641 ± 0.0020
Relative Activity‡	119 ± 11	123 ± 10%	147 ± 11%	113 ± 12%

† Expressed as ΔOD_{600}/min as determined as described in 3.2.3.3.
‡ Calculated by (d-b)/(c-a)×100%.

There is no clear explanation as to the increased activity imparted by the presence of the peptides but there are some possibilities. As mentioned above, the peptides *themselves* do not cause a clearing of the cells in the absence of LaL. However, as the peptidoglycan is degraded by LaL and the integrity of the substrate cells is weakened, new modes of action for the peptide-induced lysis may become available which could lead to an

apparent increase in activity. On the other hand, the direct binding of exogenous peptides to LaL may occur which in some way, causes an increase in the rate of LaL mediated peptidoglycan degradation. Conceivably, the peptide may bind to LaL resulting in (subtle) conformational alterations in the enzyme which could favour its association for the peptidoglycan.

As was detailed in the introduction of Chapter 1, productive complexes between HEWL and $(GlcNAc)_n$, n = 4-6, form through the collective steps of the β and γ processes (refer to Fig. 1.9). Since we have established that LaL binds $(GlcNAc)_n$ presumably across the active site yet does not cleave them, this suggests that productive "type" complexes do form which are however, unable to continue through to a reactive complex. Interactions between LaL, the glycan as well as the peptide portion of the peptidoglycan may need to occur synergistically in order for a reactive complex and cleavage to occur. The extensive contacts made between T4L and the peptidoglycan fragment (Fig. 3.4; p.196) and the fact that T4L does not hydrolyse $(GlcNAc)_n$ but does hydrolyse the peptide substituted tetrasaccharide (Fig. 3.1, compound **I.C**; p. 183) strongly suggests this to be true for T4L.

If one considers the cell wall as a substrate comprising two components, where component 1 (C1) is the glycan and component 2 (C2) is the peptide, then either (i) C1 binds before C2; (ii) C2 binds before C1, or (iii) C1 and C2 bind simultaneously to LaL. With the further assumption that the exogenous peptides occupy the same enzyme binding region for C2, then the binding of exogenous peptide to LaL would be expected to prevent peptidoglycan binding under scenario (ii) or (iii). However, it is also possible that a ternary complex between LaL, exogenous peptide and peptidoglycan can be made if scenario (i) is operative, i.e. the simultaneous binding of exogenous peptide and only C1. The same three situations regarding the order of the formation of this ternary complex is again possible. In some unknown manner, this ternary complex could enhance the formation of a reactive complex between LaL and the peptidoglycan which is observed in an apparent increase in bacteriolytic activity. For this to occur, the exogenous peptide must be displaced by C2 since it is highly unlikely that the enzyme can accommodate each of C1, C2 and exogenous peptide during a productive, reactive complex. Formation of a non-reactive complex between LaL, C1 and exogenous peptide can be speculated if one envisions C2 extending away from the enzyme and exogenous peptide occupying the C2 binding region.

We do have some supporting evidence that the synthetic peptides are able to complex with LaL. The interaction of peptide **3** with LaL was explored by affinity capillary electrophoresis (ACE) and the results are illustrated in Figure 3.34. Affinity capillary

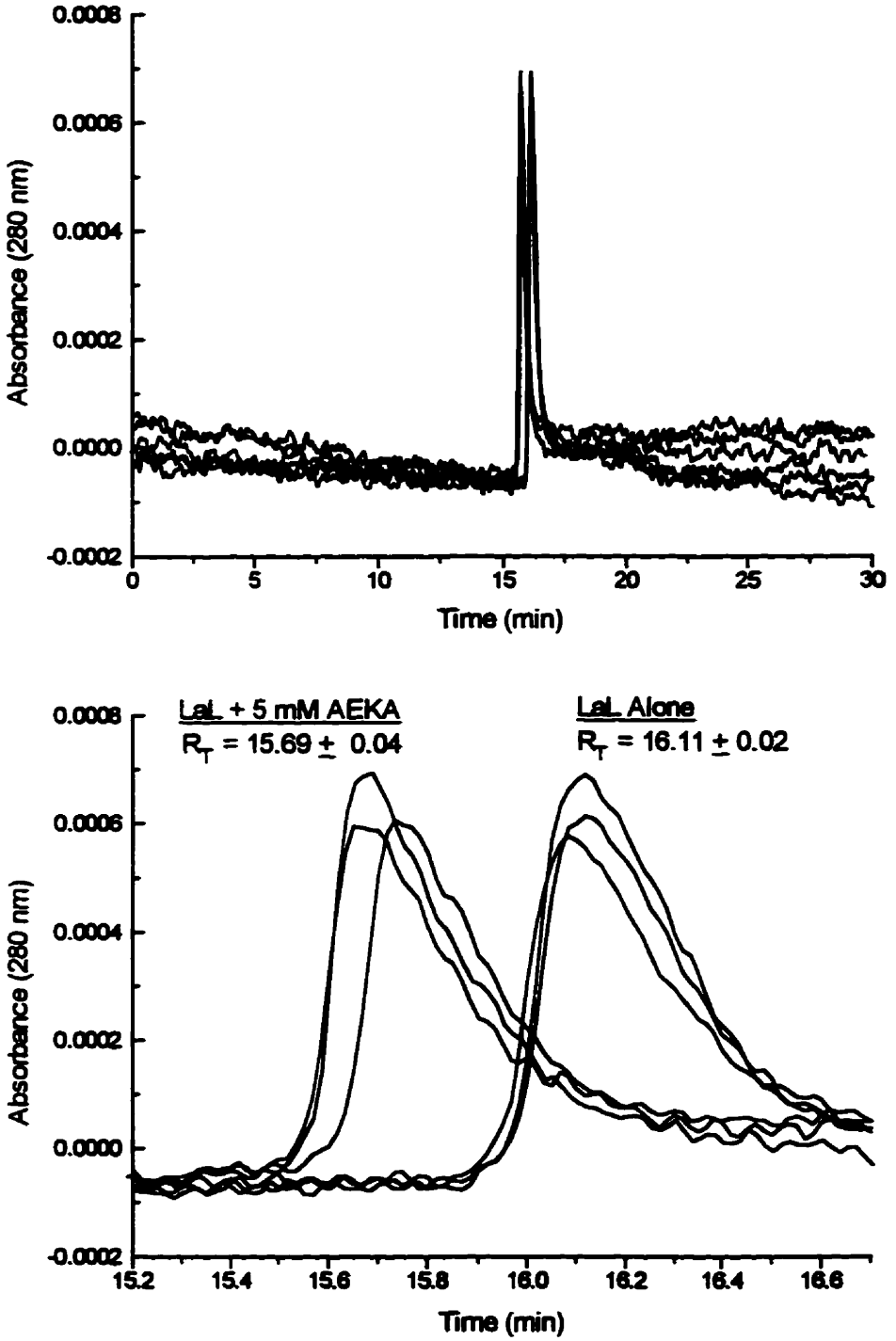


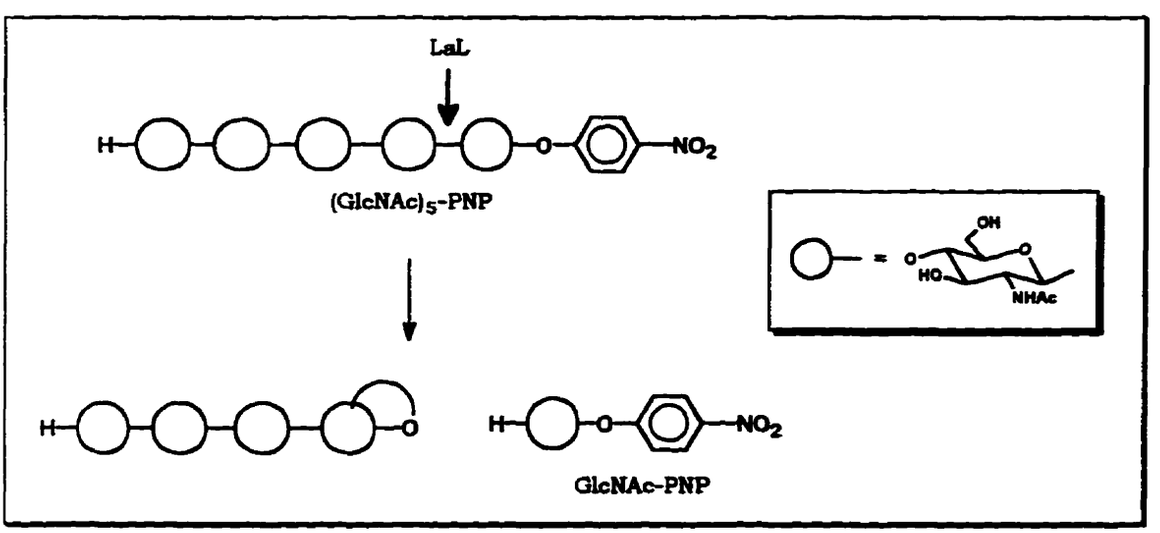
Figure 3.34. Affinity capillary electrophoresis of LaL with peptide **3**. The running buffer was 0.2 M borate, pH 8.5 containing none or 5 mM of peptide **3**. Samples were subjected to CZE with the 24 cm x 25 μM coated cartridge at a voltage of 10 kV and positive to negative polarity. The bottom panel is the expansion of the region of interest.

electrophoresis has emerged in recent years as a useful method for demonstrating binding of ligands to proteins (Chu et al., 1992; Kraak et al., 1992; Chu & Whitesides, 1992; Avila et al., 1993; Chu et al., 1994; Gomez et al., 1994). The electrophoretic mobility of a protein is dependent on both its mass and charge. In principle, if a protein binds a ligand and the overall charge and mass characteristics of the complex differ from the native protein, then the protein-ligand complex will migrate at a different rate than the uncomplexed protein. As shown in Fig. 3.34, the presence of 5 mM peptide **3** resulted in a slightly accelerated migration for LaL than that of the free enzyme by approximately 0.4 min, suggesting some type of interaction between the peptide and LaL.

At pH 8, LaL (pI ≈ 9.5-9.8) will be positively charged while peptide **3** will acquire an overall charge of approximately -0.5 (assuming pKa's for: the N-terminal amino, 8; the Lys-ε-amino, 10.6; the C-terminal and Glu carboxyls, ≈ 3). As such, binding of peptide **3** to LaL is expected to result in decreasing the overall positive charge of the complex and hence, slow the migration of the complex as compared to the free enzyme (i.e. a more positively charged complex will migrate faster under a positive to negative polarity while a less positively charged species will migrate more slowly). However, it is possible that the individual charge contributions of LaL and/or peptide **3** *when complexed* differ from their non-liganded states and the resulting complex may become more positive and hence migrate more quickly as observed. An increase in overall positive charge can be realized if the binding of peptide results in alterations in the global protein environment which either raises the pK_a of an acidic or basic residue(s). The same argument can be applied to the ionizable groups of the enzyme-bound peptide. The liganded enzyme will also be of greater mass than the free enzyme. This would also contribute to changing the electrophoretic mobility of the complex since this mobility is inversely proportional to mass.

The ACE results presented above are preliminary and do not preclude any non-specific effects that the presence of the peptide in the running buffer may have on the migration of LaL. This topic could be addressed by investigating the effect of related peptides differing in the amino acid sequence and composition (i.e. the *L* and *D* isomers could be interchanged). Fluorescence studies with LaL and the pentapeptide **2** were performed but this peptide did not influence the fluorescent properties of LaL as did the chitooligosaccharides. However, DSC analysis indicated that peptide **2** (at 5 mM) did raise the T_m of LaL by 0.6 °C, a similar effect as noted for $(GlcNAc)_3$. This result is again preliminary since only a single experiment was performed and the non-specific effects of the peptide on the T_m of LaL have not been addressed.

More intriguing and convincing evidence implicating the involvement of peptide in the mechanism of LaL was established upon co-incubation of LaL with peptide **3** and $(GlcNAc)_5$-PNP. The rationale for this study was that if LaL could be saturated with peptide, the important interactions between the peptide and enzyme would be realized endowing LaL with the ability to cleave a carbohydrate substrate. The choice of $(GlcNAc)_5$-PNP as the substrate are as follows. It is presently the largest commercially available (yet very costly) PNP glycoside of a chitosaccharide. Since it is known that the action of LaL on *E. coli* peptidoglycan produces a tetrasaccharide (Taylor et al., 1975), then LaL may effectively cleave between the fourth and fifth sugar residue from the non-reducing end of an oligosaccharide. If so, then monitoring the reactions by HPLC at 300 nm will permit a more sensitive and unique window to detect only those products containing PNP. As such, the reaction can be followed for the disappearance of $(GlcNAc)_5$-PNP and for the appearance of GlcNAc-PNP (Scheme 3.4).



Scheme 3.4.

The results of this study are shown in Figure 3.35. The reactions were followed by HPLC and quantitated from the ratio of the areas corresponding to GlcNAc-PNP and $(GlcNAc)_5$-PNP. It is expected that over the duration of the incubation (6 days) non-specific spontaneous hydrolysis of the substrate would occur. When LaL was incubated with only $(GlcNAc)_5$-PNP (curve B), no appreciable difference in the measured ratio was observed over that of the enzyme control (curve A) although a very slight difference may be

Figure 3.35. Analysis of the co-incubation of LaL with (GlcNAc)5-PNP and peptide 3.

Incubations were performed in 50 mM KPB, pH 7.0 and LaL was present at a concentration of 1 mg/mL, (GlcNAc)5-PNP at 1 mM, peptide 3 (*L*-Ala-*D*-iso-Glu-*L*-Lys-*D*-Ala) at 20 mM and GlcNAc at 20 mM. Samples were incubated at 37 °C and over a period of 6 days, aliquots (30 μL for sample **A**; 50 μL for samples **B-D**) were chromatographed over the Showdex® OHpak Q-801 (8 mm × 50 cm) gel permeation column heated at 55 °C using MQW and a flow rate of 0.5 mL/min with detection at 300 nm.

The samples are:

(A) Enzyme control: (GlcNAc)5-PNP + peptide 3 + GlcNAc

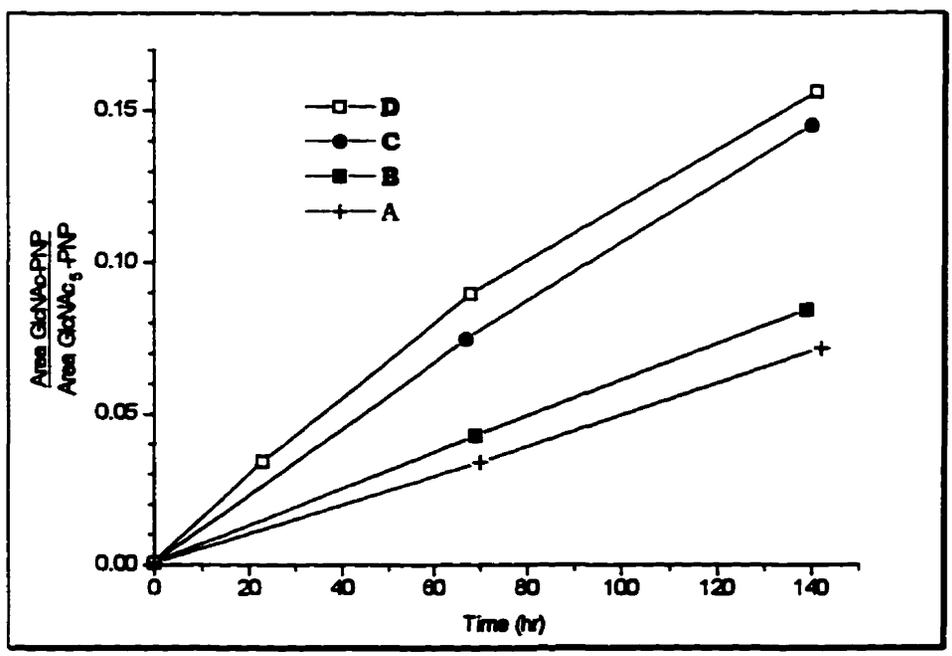
(B) LaL + (GlcNAc)5-PNP

(C) LaL + (GlcNAc)5-PNP + peptide 3

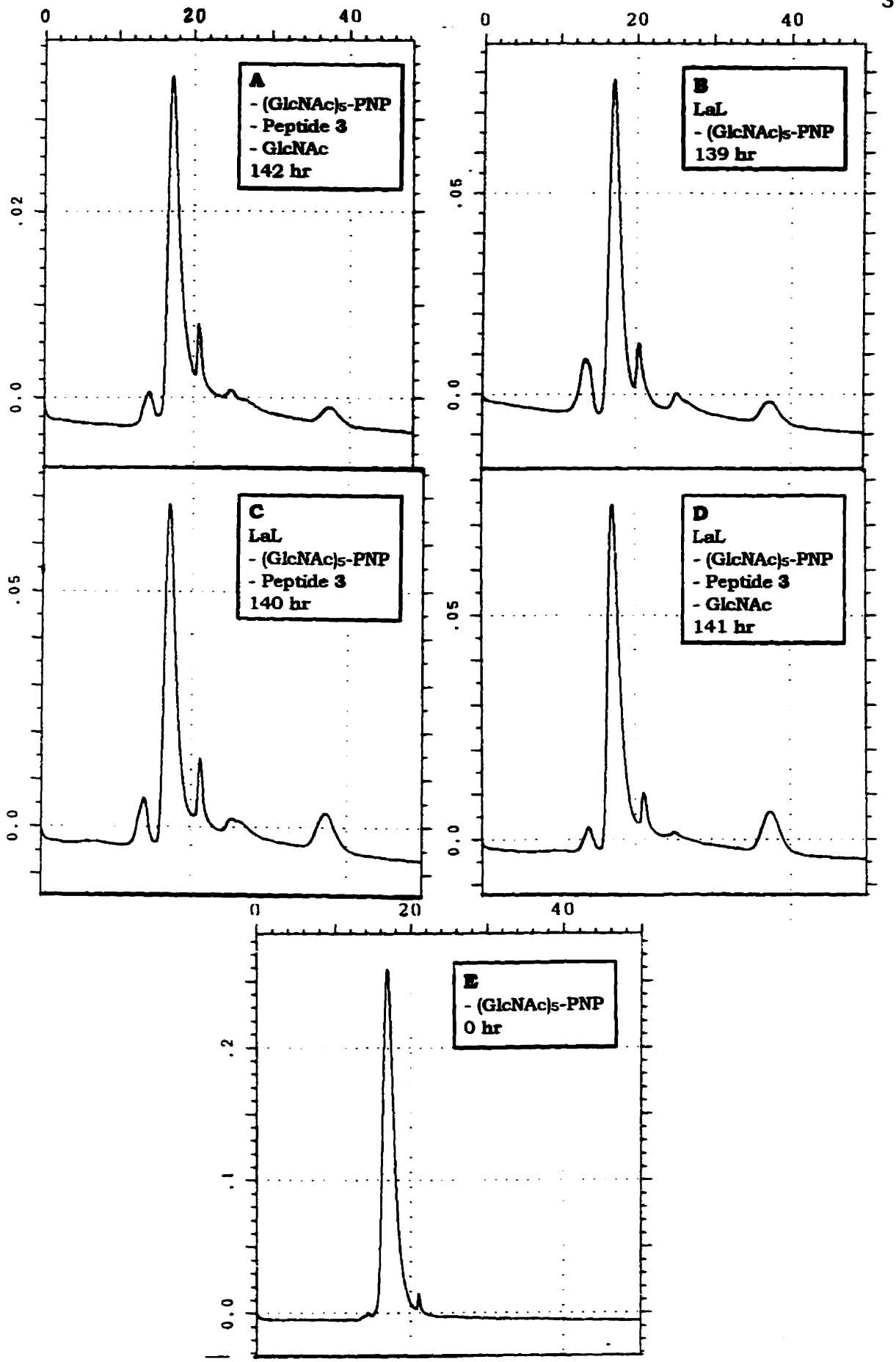
(D) LaL + (GlcNAc)5-PNP + peptide 3 + GlcNAc

(E) 1 mM (GlcNAc)5-PNP in 50 mM KPB, pH 7.0

The areas under the curve resulting from GlcNAc-PNP (R_T 37.4 min) and (GlcNAc)5-PNP (R_T 17.1 min) were obtained and their ratio plotted vs. time (see below). The zero time point ratio was obtained from chromatography of 100 μL of sample **E**.



Representative chromatograms of the samples are shown on the following page. The ordinate and abscissa axes are absorbance (300 nm) and time (min) respectively.


A
- (GlcNAc)5-PNP
- Peptide 3
- GlcNAc
142 hr
B
LaL
- (GlcNAc)5-PNP
139 hr
C
LaL
- (GlcNAc)5-PNP
- Peptide 3
140 hr
D
LaL
- (GlcNAc)5-PNP
- Peptide 3
- GlcNAc
141 hr
E
- (GlcNAc)5-PNP
0 hr
0
20
40
.02
.05
.2
.1
0.0

suggested from the data. However, when LaL was incubated with both the substrate and peptide **3** (curve C), the ratio of GlcNAc-PNP to $(GlcNAc)_5$-PNP was nearly twice that of the control. This observation suggests that in the presence of peptide **3**, LaL catalyzed cleavage of the substrate has taken place. LaL was also incubated with the substrate and peptide in addition with GlcNAc. In the case of HEWL, GlcNAc acts as an acceptor some 2000 times better than water (Rupley & Gates, 1967; Rupley et al., 1968). However, the presence of GlcNAc resulted in essentially the same measured ratio (curve D) as occurred in its absence (curve C). This might suggest that the mechanistic features of LaL resulting in 1,6-anhydro formation are not conducive for transglycosylation reactions with another carbohydrate acceptor. Where the carbocation intermediate can be trapped by water or more effectively, by GlcNAc with HEWL, these reactants may be precluded in the active site of LaL where the C6-hydroxyl acts as the favoured acceptor.

After 140 hr of incubation, the ratio for the enzyme control (A) and samples including LaL and peptide **3** (C and D) are approximately 0.07 and 0.15 respectively. Examination of the chromatograms in Fig. 3.35 does reveal that other PNP-containing products have been generated (compare A-D with E, the substrate only). Under the same chromatographic conditions, $(GlcNAc)_2$-PNP and $(GlcNAc)_3$-PNP have retention times of 26.3 and 21.1 min respectively and are apparently present. Assuming however, that the only product of the enzyme catalyzed reaction is GlcNAc-PNP and that equivalent non-specific hydroylsis of $(GlcNAc)_5$-PNP has occurred in each sample, the values noted above can be taken to indicate that 93% of $(GlcNAc)_5$-PNP remained in the control sample and 87% remained in the presence of LaL and peptide **3**. Although this difference is small, indicating a rather weak activity, the results nonetheless ascribe LaL with the ability to cleave a $(GlcNAc)_n$ substrate and the peptide may be viewed as a positive modulator for this activity. To our knowledge, this is the first account of a measured activity for a phage lysozyme with a non-peptidoglycan substrate. It will be imperative to isolate and characterize the carbohydrate products of this reaction. This will hopefully identify products containing the 1,6-anhydro linkage and substantiate involvement of LaL. Taken together with the apparent ability of the peptides to enhance the bacteriolytic activity of LaL, there does appear to exist a peptide-specific function for the activity of LaL.

It is important to consider the possibility that the peptidyl-mimics assist in catalysis without specific binding to LaL; i.e. similar to the substrate-assisted catalysis model proposed by Matsumura & Kirsch (1996) (refer to Fig. 3.5). In their model, substrates containing a covalently attached carboxylate group which can be appropriately positioned can serve to augment catalysis in lysozymes that *are lacking a counterpart* to

Asp52 of HEWL. The peptides **2** and **3** that we have investigated contain two free carboxyl groups at the *D*-Glu and terminal *D*-Ala positions. In order for the peptides to enhance catalysis: i) LaL must lack the Asp52 counterpart such that the peptides provide a carboxylate that is normally absent; and ii) the peptides must interact with a LaL/substrate complex in such a manner that one of the carboxyl groups will be closely positioned near the cleaved sugar. The discussion which follows will argue against the possibility that the peptides assist in catalysis through the involvement of one of the two peptide carboxylate groups.

Through both sequence alignment comparisons and mutagenesis, it has been proposed that Asp34 in LaL may serve as the catalytic acid counterpart to Asp52 (Jespers et al., 1992). The mutations of Asp34 to Asn and Ala reduced the activity of LaL by factors of 17 and 250 respectively (Jespers et al., 1992) which are comparable to the reductions in activity reported for mutants of Asp52 in HEWL (Malcolm et al., 1989; Lumb et al., 1992; Matsumura & Kirsch, 1996). Therefore, there is supporting evidence suggesting that LaL does possess the complement to Asp52. Substrate-assisted catalysis did not occur with the wild type HEWL (Matsumura & Kirsch, 1996) indicating that Asp52 acts as the preferred "catalytic" carboxylate even when the substrates contained potential (and covalently attached) "catalytic" groups. From this argument, the increase in the bacteriolytic activity observed for LaL in the presence of the peptides cannot be attributed to the carboxylates of the peptide assisting in catalysis.

In order to stabilize the glycosyl transition state, the carboxylate must be correctly situated to functionally serve this role. Mutation of the catalytic Asp53 in human lysoyzme to Glu produced an enzyme with diminished activity (<5%) on both bacterial cells and $(GlcNAc)_5$-PNP (Muraki et al., 1991). No significant differences in the binding affinities for $(GlcNAc)_{3,6}$ was observed for the mutant and *wt* enzymes nor was there any difference in the structure of the D53E enzyme except at position 53. Therefore, the movement of the carboxylate group invoked by the additional methylene unit in the D53E mutant resulted in a reduction of activity and establishes the importance of the precise positioning of the carboxylate group (Muraki et al., 1991). If there is no specific binding of the peptides to LaL, then a random encounter between peptide and a LaL/substrate complex that is actively engaged in cleavage and further, an encounter producing a precisely positioned peptide-carboxylate, seems unlikely.

Furthermore, no phage lysozyme has demonstrated activity on a non-peptidoglycan substrate. The apparent ability of LaL to cleave $(GlcNAc)_5$-PNP in the presence of peptide **3** again seems unlikely to have occurred by simple participation of the peptide's carboxylate groups. Like LaL, GEWL is ineffective against chitosaccharide substrates and

is believed to utilize additional binding energy from its association with the peptide portion of the peptidoglycan in order to strain the D site MurNAc into the reactive conformation (Weaver et al., 1995). If LaL must also develop additional energy to cleave its substrates, then the observed action on $(GlcNAc)_5$-PNP could only be accounted for if peptide **3** had specifically bound to LaL and supplied the required additional interactions.

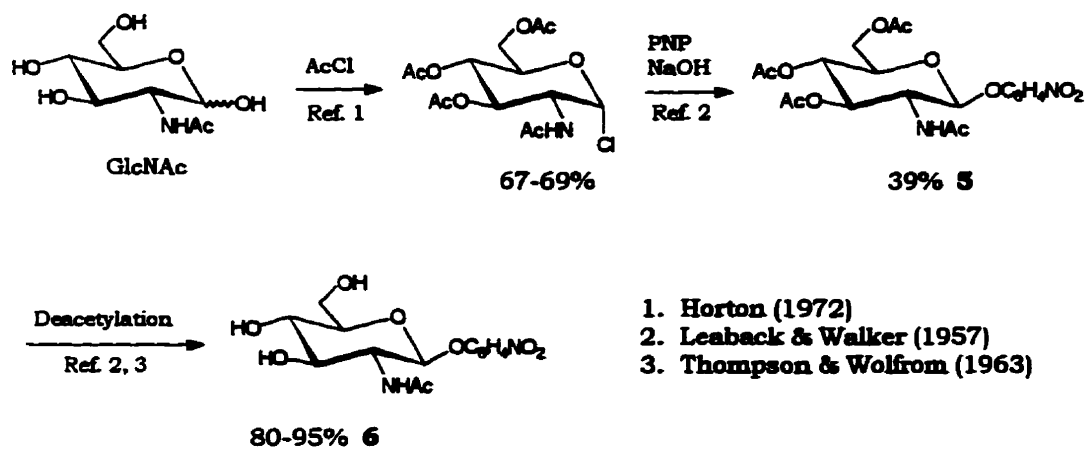
If one considers that the hypothetical binding of the peptidyl mimics to LaL does in fact occur, then some issues should be addressed comparing the binding of the peptides and the binding of the peptide portion of *E. coli* peptidoglycan. Firstly, in the peptide analogues prepared, *L*-Lys replaces the *L-meso*-DAP residue in the natural substrate. The analogues therefore lack the ε-carboxylate and as well, any potential interactions between LaL and this group. The crystal structure of the mutant T4L cell wall adduct showed that the peptide moiety extended across an open groove on the surface of the enzyme between 2 α helices making both extensive main chain and side chain interactions between enzyme and peptide (Kuroki et al., 1993; see also Fig. 3.4). The amide and hydroxyl of S117 form hydrogen bonds with the ε-carbonyl of DAP. Mutations at this position in T4L (as well as other residues which are in proximity to the peptide binding site, eg. Q105, M106, F114, S136) are relatively intolerant to substitution (Rennell et al., 1991). By analogy, it is possible that like T4L, LaL does also interact with ε-carboxyl of DAP and the absence of this moiety in peptide **2** and **3** may reduce their affinity with LaL. However, if the analogy between LaL and T4L can be extended further, then the presence of Lys instead of DAP may in fact not be consequential. It was discussed previously that the inhibition of T4L by muropeptides containing either of these two diamino acids are similar (refer to section 3.1.3 and Table 3.5).

Secondly, the peptide mimics possess a free N-terminal amino group that will carry a positive charge. This charge is absent in the natural substrate since this amine forms an amide with the lactyl ether of muramic acid. What the effect of this unnatural charge would have on the binding of the peptides has not been established but could be explored by preparing the N-acetyl or N-formyl derivatives of the peptides. This positive charge however, may in fact be accommodated in the active site of LaL. Calculations of the electrostatic fields in three lysoyzmes (HEWL, HuL and T4L) have suggested that in each, there exists an asymmetric distribution of charge such that there exists a positive charge on the upper lobe and negative charge on the lower lobe across the active site around the catalytic residues (Dao-Pin et al., 1989). If a similar distribution exists in LaL, then the negative charge of the lower lobe may be favorable towards binding of a positively charged amino group.

3.3.4. Synthesis of β-*p*-Nitrophenyl-Muramic Acid Analogues

To further probe the structural requirements of compounds that might serve as substrates for LaL, we undertook the synthesis of PNP glycosides of muramic acid and some substituted derivatives thereof. Unlike the β-PNP glycosides of N-acetylglucosamine oligosaccharides whose chemical (Leaback & Walker, 1957; Osawa, 1966; Osawa & Nakazawa, 1966; Ballardie et al., 1977; Nanjo et al., 1988a) and chemoenzymatic (Usui et al, 1988) syntheses have been described in the literature, a comparable description of the synthesis of β-PNP-muramic acid analogues is currently lacking. To date, only the preparation of β-PNP-MurNAc has been reported (Jeanloz et al., 1968). Although others have utilized this methodology to prepare β-*p*-aminophenyl-MurNAc-*L*-alanyl-*D*-isoglutamine (Lefrancier et al., 1978) we sought to improve upon the synthesis of β-PNP-MurNAc as well as to prepare some novel substituted analogues.

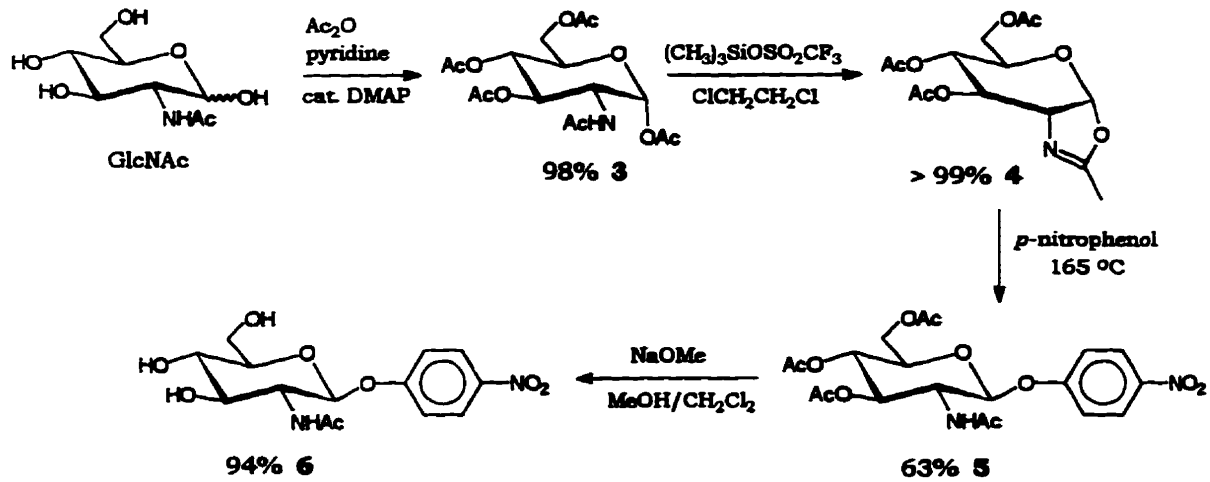
Our efforts towards the synthesis of β-PNP-MurNAc required the synthesis of significant quantities of the chosen starting material, β-PNP-GlcNAc **6** (*p*-nitrophenyl-2-acetamido-2-deoxy-β-D-glucopyranoside), which can be prepared from GlcNAc. From a review of the literature, the optimal conditions to afford **6** would involve the 3 step conversion below (Scheme 3.5) resulting in an overall yield of 21-26% from GlcNAc.



Scheme 3.5.

Because of the low reported synthetic yields and high commercial costs for β-PNP-GlcNAc (available from Sigma), a new synthetic strategy for its preparation was required and successfully developed. We achieved the synthesis of **6** from commercially available GlcNAc (Scheme 3.6). The critical step towards **6** was the condensation of PNP with a

suitably activated peracetate of GlcNAc to give the peracetate of β-PNP-GlcNAc **5**. Various attempts were made to improve upon the Koenigs-Knorr reaction of PNP with tri-*O*-acetyl-1-chloro-α-GlcNAc (Scheme 3.5), the latter being prepared by the method of Horton (1972), but no improvement in yields were obtained. However we found that the oxazoline **4** reacts with PNP under autocatalytic fusion conditions to give **5**.



Scheme 3.6.

The autocatalytic fusion reaction was first reported for the reaction of a sugar peracetate with a protic nucleophile (Ishido et al., 1965; Ishido et al., 1967) and has been applied to the preparation of PNP glycosides of chitooligosaccharide peracetates (Nanjo et al., 1988a). However, only the β-anomer of a sugar peracetate is susceptible to the autocatalytic fusion reaction yielding the β-glycoside, whereas the α-anomer is not (Nanjo et al., 1988a). Therefore, conditions must be established to preferentially prepare the β-anomer in order to obtain high fusion yields with a sugar peracetate. Although glucose can be peracetylated with Ac_2O in the presence of a basic catalyst (eg. NaOAc) to give predominantly the β-anomer, or in the presence of an acidic catalyst (eg. $ZnCl_2$) to give predominantly the α-anomer, the anomeric configuration obtained with any catalyst also depends on the nature of the carbohydrate (Wolfrom & Thompson, 1963). When acetylation of GlcNAc was attempted with both NaOAc and $ZnCl_2$, no preferential anomer of GlcNAc peracetate was obtained.

In our synthesis (Scheme 3.6), GlcNAc was converted into its peracetate **3** with Ac_2O in the presence of pyridine and DMAP. Under these conditions, an anomeric ratio of $\alpha:\beta \approx 15:1$ was observed, and therefore **3** would be unreactive towards fusion with PNP.

However, Nakabayashi et al. (1986) have reported that both the α and β anomers of GlcNAc peracetate readily react with TMS triflate to generate the oxazoline **4** in near quantitative yields. We have applied this methodlogy to prepare the oxazoline **4** and have further demonstrated that under controlled conditions, **4** underwent autocatalytic fusion with PNP yielding **5** in a respectable 63% yield. The peracetate **5** could be purified using solvent extraction. Confirmation of the structure for **5** was also achieved by peracetylation of authentic β-PNP-GlcNAc whose ^{1}H NMR was identical to **5**. Deacetylation of **5** was achieved with sodium methoxide, and by including CH_2Cl_2 in the solvent, **6** could be isolated in high purity through its precipitation from the solvent. The methodology developed is highly reproducible and results in a 57% overall yield of β-PNP-GlcNAc from GlcNAc, improving substantially upon the previous reported methods.

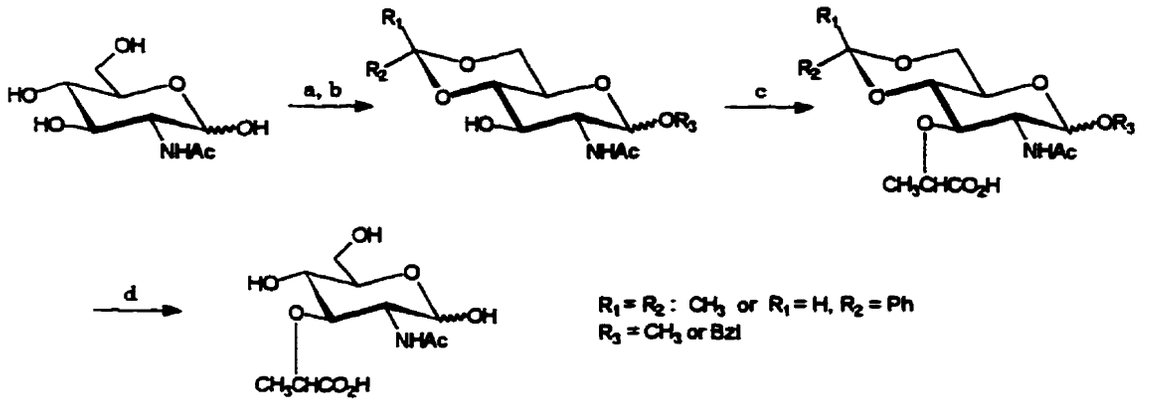
The first synthesis of muramic acid (MurNAc, 2-acetamido-2-deoxy-3-*O*-[*D*-1-carboxyethyl]-*D*-glucose) was reported by Kent (Kent, 1957; Strange & Kent, 1959) and since that time, numerous other syntheses of MurNAc and related compounds have been reported[3.3] . The more recent interests regarding MurNAc syntheses have been directed towards the synthesis of muramylpeptides to investigate their biological properties and reviews on this subject have appeared (Lefrancier & Lederer, 1981; Baschang, 1989). The synthesis of MurNAc generally follows a scheme in which modifications include the choice of protecting groups and etherification method (Scheme 3.7). *N*-Acetylglucosamine is protected at C1 as either the methyl or benzyl glycoside, and either a 4,6-*O*-isopropylidene or 4,6-*O*-benzylidene protects the C4 and C6 hydroxyl groups. The isopropylidene group is more labile than the benzylidene group yet confers better solubility. However, the most frequently used protected muramyl derivative is the 1-*O*-benzyl-4,6-*O*-benzylidene derivative as the simultaneous elimination of the benzyl and benzylidene group is possible through hydrogenolysis. The lactyl ether is introduced onto the C3 hydroxyl typically through condensation with a 2-halo-propanoic acid (typically *DL*-2-chloropropanoic acid has been used because of its availability) or the tosylate of lactic acid. Separation of the diastereomeric (*D,L*)-lactate ethers is required to obtain the *D*-lactyl isomer (the *L* isomer is *iso*-muramic acid) unless optically pure 2-(S)-halo-propanoic acid or *L*-lactic acid is used. Deprotection of the hydroxyl protecting groups then affords muramic acid.

Having adequate access to β-PNP-GlcNAc, the synthesis of β-PNP-MurNAc and derivatives was undertaken. The first key intermediate required was the protected β-PNP-

---

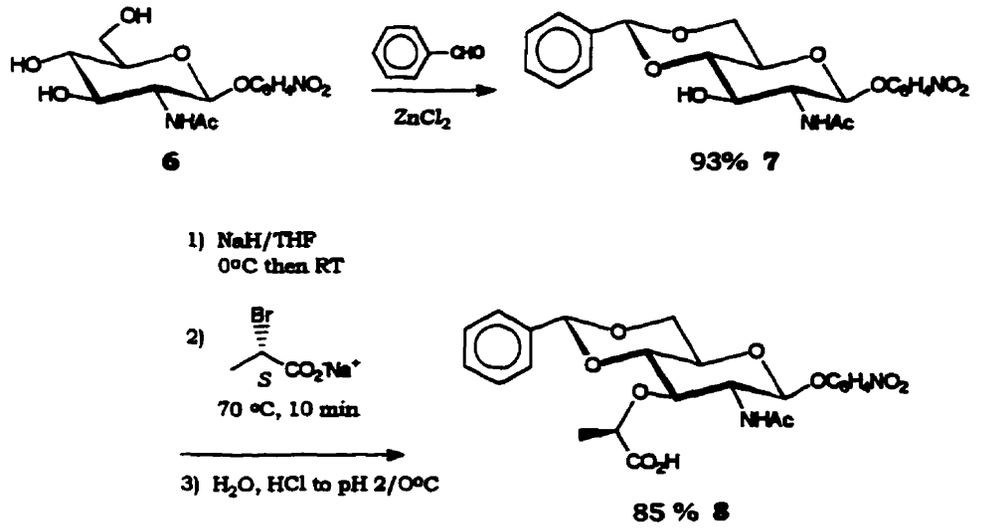
[3.3] For examples see Matsushima & Park (1962), Flowers & Jeanloz (1963), Gigg et al. (1965), Osawa & Jeanloz (1965), Merten & Brossmer (1989).

MurNAc analogue **8** (Scheme 3.8). Previously, the methyl ester of **8** had been prepared from **6** in 20% overall yield (Jeanloz et al., 1968). In this report, *DL*-2-chloropropanoic acid was used to introduce the lactyl ether to give **8** which was subsequently esterified with diazomethane by these researchers, presumably to enhance the chromatographic isolation of the ester over silica gel. In our synthesis (Scheme 3.8), **8** was prepared from **6** in an overall yield of 79%, improving upon the published method.



a: $C_6H_5CH_2OH$ or CH_3OH, H^+; b: PhCHO or 2,2-dimethoxypropane, H^+
c: $CH_3CH(X)CO_2H$ or $CH_3CH(OTs)CO_2H$, base; d: H^+ or H_2/cat

Scheme 3.7.



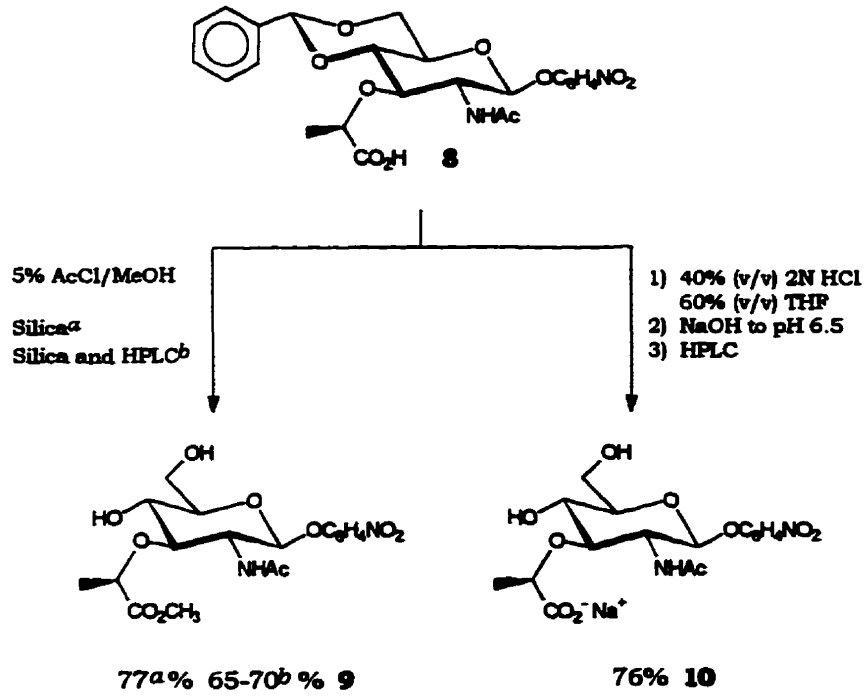
Scheme 3.8.

Conversion of **6** to the benzylidene **7** (Scheme 3.8) went efficiently and it is recommended that freshly purified benzaldehyde be used (Fletcher, 1963). During the course of this work, it was fortunate that (*S*)-(-)-2-bromopropanoic acid became commercially available, whose enhanced reactivity over the chloro derivative and optical purity resulted in an improvement of the introduction of the *D*-lactate ether. Modifications to the synthesis of Jeanloz et al. (1968), which followed more closely the traditional approaches, were developed for conversion of the alchohol precursor **7** to the ether **8** which are believed to have led to improved yields and are discussed below.

The traditional solvent used in the condensation is 1,4-dioxane, however the solubility of **7**, although still poor, was much greater in THF. Secondly, in the synthesis detailed by Jeanloz et al. (1968) and in other previous syntheses (see references given in footnote 3.3), the precursor alcohol is maintained at elevated temperatures (60-70 °C) in the presence of NaH for some time (15 min to 1 hr) prior to the addition of the 2-halo-propanoic acid. When this approach was attempted with **7**, the mixture became very dark and the final yields of **8** obtained after reaction with the 2-bromopropanoate were reduced by 15-20%. Instead, the reaction was heated only briefly (10 min) in the presence of the 2-bromopropanoate and then immediately cooled, which resulted in higher yields and no discolourization of the mixture. Finally, Jeanloz et al. (1968) recovered **8** after acidification and organic extraction using high proportions of $CHCl_3$:H_2O, followed with methyl esterification of the acid and silica gel chromatography. In our method, **8** was precipitated and collected by filtration following acidification in solvent containing an excess proportion of water. Isolation of **8** as the free acid was also advantageous and desired since the acid is available to be coupled with peptides towards the preparation β-PNP-muramylpeptide analogues.

Two methods were employed to remove the benzylidene group of **8** giving the methyl ester **9** and the sodium salt **10** of β-PNP-MurNAc respectively (Scheme 3.9). Since hydrogenolysis would result in reduction of the nitro group to the amine, acid catalyzed deprotection was chosen. We prepared **9** in 65-70% yield by treatment of **8** in a 5% solution of AcCl in anhydrous MeOH, which generates HCl *in situ* as the acid source. Application of this method was favourable resulting in both the simultaneous methanolysis of the benzylidene group and protection of the acid as the methyl ester. The analogue **9** is well suited to act as an acceptor in glycoside synthesis. The C6-hydroxyl could be protected (eg. as a silyl ether) leaving the C4-hydroxyl accessible for condensation with an appropriate donor towards the preparation of chitooligosaccharide

analogues with β-PNP-MurNAc as the terminal unit. Purification of **9** was achieved by sequential normal and reverse phase chromatography. It is suggested that if **9** is to be used as a starting material in subsequent syntheses, a sufficient level of purity of **9** (>99%) is obtained for this purpose and in higher yields following only its chromatography over silica gel and the reverse phase chromatography need not be performed. The latter purification step was employed to obtain **9** for purposes of enzymic assays.



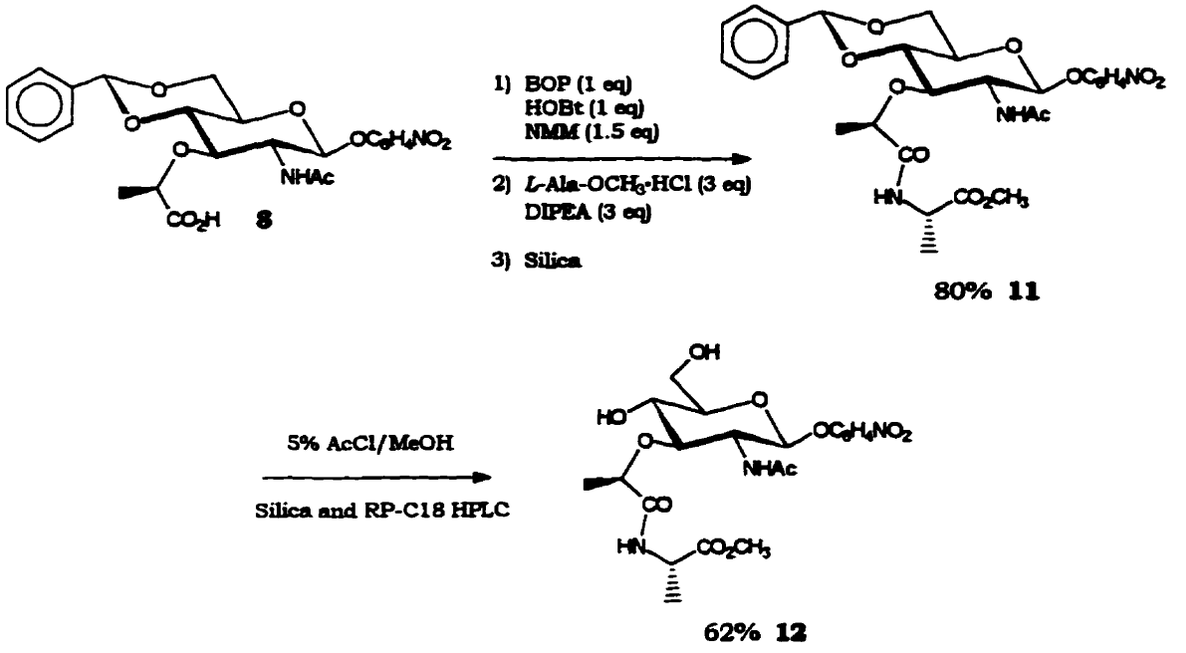
Scheme 3.9.

The removal of the benzylidene group from **8** was also achieved using a co-solvent of 2N HCl and THF. Following basification, the sodium salt of **10** was obtained. Isolation of the sodium salt instead of the free acid was initially considered to be favourable since i) the salt exhibited superiour aqueous solubility and ii) it was feared that the acid may be prone to undergo lactonization between the C4-hydroxyl and the carboxylic acid. Although in one instance ^{1}H NMR suggested that the lactone had been generated, it is now believed that the lactonization resulted from inappropriate workup conditions. It is believed that the free acid can be isolated using identical methodology as to obtain the salt except that TFA must be included in the mobile phases during RP-C18. This was realized by chromatography of the salt (which elutes in a mobile phase of 100% MQW) in

mobile phases containing 0.1% TFA in which the acid is recovered after application of a gradient elution to CH_3CN (potential chromatographic conditions to obtain the acid are given with the experimental details for **10**).

There is some uncertainty regarding the level of purity obtained for **10**. Elemental anlaysis was not consistent with its composition and the results obtained could best be rationalized, but not in full agreement, if **10** was purified as the dihydrate with 1 mol equivalent of NaCl (the yield of 76% given is based on this composition). Both the ^{1}H and ^{13}C NMR spectra of the material purified revealed that no organic contaminants were present and the chemical shift of the lactyl carboxylate (180.7 ppm) suggested that the salt was indeed obtained. The ESMS spectrum of **10** provided correct molecular weights for the salt, the free acid (which is generated under the acidic ESMS conditions) and for a hydrochloride adduct of the free acid in which both isotopes of chlorine (^{35}Cl and ^{37}Cl) were observed. For this reason, contamination of **10** with NaCl is assumed but was unexpected. Considering that **10** was retained on the RP-C18 support used for its purification, it was expected that it would be separated from any NaCl. The evaluation of the preparation of **10** described above as a potential substrate was performed prior to the ESMS and elemental analyses. Since **10** did not serve as a substrate, further attempts towards its purification were not warranted as the preparation was considered sufficiently pure to be evaluated as a substrate.

The preparation of substituted β-PNP-MurNAc analogues was initiated through model reactions to couple the protected derivative **8** with the methyl ester of *L*-alanine. Carbodiimide mediated coupling methodologies were investigated for their efficacy to promote coupling, each using an excess of the amino acid and included DCC/*N*-hydroxysuccinimide (Vega-Pérez et al., 1992), DCC/HOBt and DIPC/HOBt. Although each method was successful, superior results were obtained employing BOP/HOBt. We prepared Bzi-β-PNP-MurNAc-*L*-Ala-CO_2Me **11** in a reproducible 80% yield following activation of **8** in the presence of BOP/HOBt and coupling with the alanine derivative (Scheme 3.10) and **11** could be purified by chromatography over silica gel. Deprotection of the benzylidene group of **11** using a 5% solution of AcCl in anhydrous methanol provided the analogue β-PNP-MurNAc-*L*-Ala-CO_2Me **12** in a modest 62% yield. As was discussed for the analogue **9**, further protection of the C6-hydroxyl of **12** could be performed to produce an analogue suitable for subsequent condensation through the C4-hydroxyl with a chitooligosaccharide donor.



Scheme 3.10.

The analogue **12** required purification utilizing both normal and reverse phase chromatography and it is believed that this requirement lended to a reduction in the yield obtained. Although any chromatographic procedure is expected to proceed without quantitative recovery of sample, most of the PNP containing analogues exhibited peculiar and poor solubility characteristics[3.4] which presented certain difficulties with respect to column purification. Fortunately, compounds **5** through **8** could be isolated and in high purity without chromatography, but this was not possible for the analogues **9** through **12**. Dissolution of the crude products from the reactions producing **9**, **11** and **12** for silica gel chromatography required relatively large amounts of organic solvent (or solvent mixture) for the proportion of material. This, in turn, necessitated the use of larger amounts of silica gel (than is typically required) to ensure adequate resolution which will lower the overall recovery of product from the column.

The reverse phase chromatgraphy of **9** and **11** was itself not straightforward. Under the conditions utilized for their RP-C18 chromatography, **9** and **11** were eluted at CH_3CN percentages of approximately 24 and 30% respectively. However, it was necessary

---

[3.4] Solubility difficulties were notably prevalent upon sample preparation for NMR characterization. Only DMSO-d_6 could serve as an adequate solvent and even then providing only limited solubility. In some cases, it was necessary to use a combination of deuterated solvents (eg. the NMR of **9** and **12** were obtained in a mixture of acetone-d_6 and $CDCl_3$ while **6** was obtained using acetone-d_6 and D_2O) in order to achieve dissolution of the compounds.

to equilibrate the column with 20% CH_3CN in order to enhance their solubility in the initial mobile phase and prevent their precipitation onto the column. As such, initial adsorption to the column would be compromised which again could have lowered the overall recovery of product. A strategy was also adopted to load limited amounts of material (no more than 20-50 mg) even though the μBondapak™ C18 Prep-Pak® cartridge employed in their purification is designed for purification of up to 1 g of material. This in turn required multiple runs to process the samples. The factors outlined above are believed to contribute to a reduction in the final yields of these products.

Our ultimate intention was to prepare a peptide substituted β-PNP-muramyl analogue (see compound **15**, Scheme 3.11) of sufficient homology to mimic the structure of the *E. coli* peptidolgycan. To this end, the peptide *L*-Ala-*D*-*iso*-Glu-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala (peptide **1**) was prepared so that it could be coupled through its N-terminus with Bzi-β-PNP-MurNAc **8**. The success achieved in preparation of the *L*-alanyl derivative **11** prompted initial exploitation of coupling peptide **1** with **8** employing BOP/HOBt activation methodology.

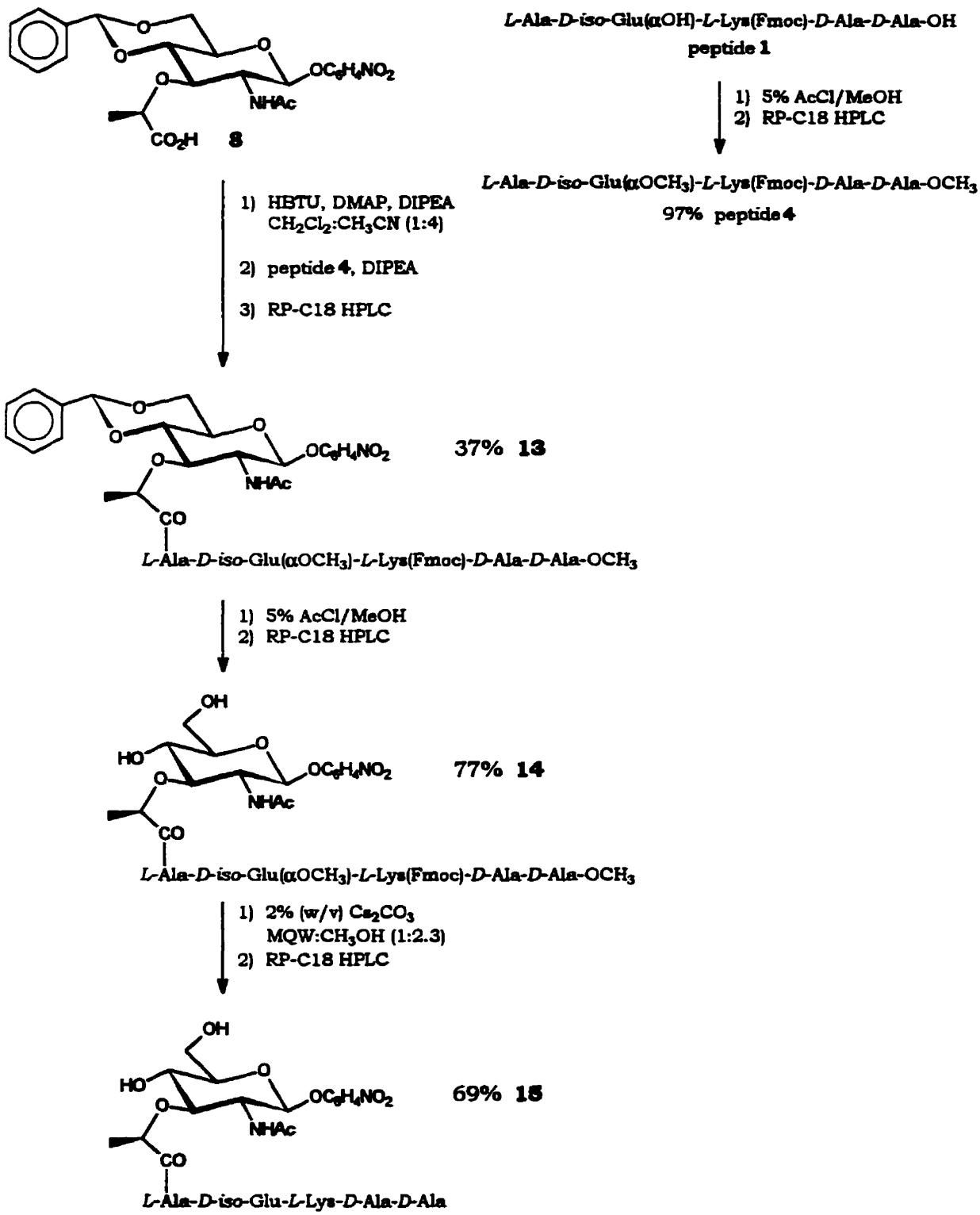
A number of attempts at coupling **8** with peptide **1** were performed under similar conditions[3.5] as used to prepare **11** with the exception that peptide **1** was the limiting reactant. The monitoring of these reactions by analytical RP-C18 HPLC indicated that coupling had indeed been achieved. However, preparative purification of the coupled product by RP-C18 HPLC appeared to be the only feasible method for purification and proved to be very difficult. In these attempts, it is believed that the by-product of the BOP activation, namely hexamethylphosphoric triamide (HMPT), presented certain problems during purification. In addition, cross-coupling of peptide **1** (i.e. through the N-terminal amino and the α-carboxyl of Glu or the C-terminal carboxyl) with itself may have also resulted. Partial pre-column purification of the coupled product from the reactions by extraction or trituration was ineffective for lack of suitable solvents. Although analytical chromatography was possible, preparative scale chromatography always resulted in precipitation of one or several of the reaction components on the column and the separation (resolution) obtained for the components was poor.

[3.5] Typical coupling conditions were: to anhydrous DMF (0.5-1 mL) was added **8** (25 mg, 1 equiv.), BOP and HOBt·H_2O (1 equiv. each) and NMM (1.5 equiv). The reaction was stirred (15-30 min) and peptide **1** (0.5-0.9 equiv.) and DIPEA (3 equiv. wrt to peptide **1**) were added and the reaction stirred at 25 °C.

Despite these drawbacks, on two occassions, minimal amounts (< 5 mg) of partially purified Bzi-β-PNP-MurNAc-*L*-Ala-*D*-*iso*-Glu-*L*-Lys(Fmoc)-*D*-Ala-*D*-Ala was isolated when prepared as described above. It was hoped that the products of the subsequent deprotection of the benzylidene and Fmoc group would facilitate their isolation and attempts were made in which either the benzylidene or Fmoc group was removed first. The fully deprotected derivative **15** (< 1mg) was obtained (the structure confirmed only by ^{1}H NMR) after sequential benzylidene and Fmoc deprotection and RP-C18 HPLC purification.

Because of the purification difficulties encountered and low recovery of the coupled product of **8** and peptide **1**, the BOP/HOBt methodology, although extremely successful for the preparation of **11**, was abandoned and other reaction conditions were explored. Small peptides have been successfully coupled to muramic acid derivatives using N-ethyl-5-phenylisoxazolium-3'-sulfonate (Merser et al., 1975), isobutylchloroformate (Lefrancier et al., 1978; Durette et al., 1979) and DCC/*N*-hydroxysuccinimide (Kusumoto et al., 1986; Vega-Pérez et al., 1992). A report by Adam (1992) describing the utilization of HBTU (*O*-benzotriazolyl-N,N,N',N'-tetramethyluronium hexafluorophosphate) as a mild activating reagent of muramic acid for its coupling with peptides prompted use of this methodology for our intended synthesis.

It was considered judicious to protect the carboxylic acids of peptide **1** as their methyl esters to prevent possible cross-coupling of the peptide with itself. The dimethyl ester of peptide **1** (peptide **4**) was prepared in 97% yield by treatment of peptide **1** with a solution of AcCl in anhydrous methanol and RP-C18 HPLC purification. Using slight modifications to the conditions described by Adam (1992), the coupling of peptide **4** with **8** was attained (Scheme 3.11). During the coupling reaction, the progression of the reaction was monitored by analytical RP-C18 HPLC which indicated complete disappearance with time of the peak corresponding to peptide **4**, which was used as the limiting reactant. This suggested that coupling went efficiently and quantitatively. Unfortunately, we were only able to isolate the coupled product **13** in 37% yield following preparative RP-C18 HPLC purification. As was discusssed for **9** and **11**, similar solubility problems and the requirement of multiple HPLC runs to purify **13** is thought to be responsible for the low recovery. Effective removal of the benzylidene was again achieved using a solution of AcCl in anhydrous methanol and **14** was isolated in 77% yield. Treatment of **14** by the method of Kaestle et al. (1991), using Cs_2CO_3 in a cosolvent of water and methanol, was found to be effective for cleavage and hydrolysis of the Fmoc and methyl esters respectively affording **15** in 69% yield.


L-Ala-D-iso-Glu(αOH)-L-Lys(Fmoc)-D-Ala-D-Ala-OH
peptide 1
1) 5% AcCl/MeOH
2) RP-C18 HPLC
L-Ala-D-iso-Glu(αOCH3)-L-Lys(Fmoc)-D-Ala-D-Ala-OCH3
97% peptide 4
8
1) HBTU, DMAP, DIPEA
CH2Cl2:CH3CN (1:4)
2) peptide 4, DIPEA
3) RP-C18 HPLC
37% 13
L-Ala-D-iso-Glu(αOCH3)-L-Lys(Fmoc)-D-Ala-D-Ala-OCH3
1) 5% AcCl/MeOH
2) RP-C18 HPLC
77% 14
L-Ala-D-iso-Glu(αOCH3)-L-Lys(Fmoc)-D-Ala-D-Ala-OCH3
1) 2% (w/v) Cs2CO3
MQW:CH3OH (1:2.3)
2) RP-C18 HPLC
69% 15
L-Ala-D-iso-Glu-L-Lys-D-Ala-D-Ala

Scheme 3.11.

Therefore, the peptide substitued β-PNP-muramyl analogue **15** was prepared in an overall yield of 20% from **8** and peptide **4**. The ^{1}H NMR of the intermediates **13** and **14** and the final product **15** are shown in Fig. 3.36 (A-C). Comparison of the spectra for **13** and **14** with that of peptide **1** (refer to Fig. 3.31) facilitates identification of the resonances arising from the peptide portion of these compounds. Removal of the benzylidene from **13** is accompanied with loss of the benzylidene resonances (i.e. Hc and Hd) and the appearance of the C4 and C6 hydroxyl resonances (i.e. H_3; compare the spectra for **13** and **14**). Conversion of **14** to **15** results in removal of the resonances due to the Fmoc group (i.e. He, Hf and Hg) and the 2 methyl ester signals (i.e. Hj). The spectra are very encouraging regarding the purity obtained for these compounds.

The analogue **15** was also characterized by MALDI MS. The calculated molecular weight for **15** is 884.85 Da, and an M_1+1 ion peak at 885.3 Da (therefore, M_1 = 884.3 Da) was detected (Fig. 3.37). In addition, an M_2+1 ion of mass 748.1 Da (and M_2 = 747.1 Da) was observed. The mass of M_2 suggests that fragmentation at the glycosidic bond of the parent ion, M_1, had resulted in which *p*-nitrophenoxide ($C_6H_4NO_3$, MW 138.10 Da) was lost through fragmentation (and 884.3 (M_1) - 138.10 = 746.2 Da). If one considers that the possible acquisition of a proton by the glycon accompanied fragmentation, then the observed mass for M_2 can be rationalized (i.e. 746.2 + H = 747.2 Da ≈ M_2). Ionization of samples during MALDI MS is accomplished using a 337 nm nitrogen laser source. The absorption maxima for a *p*-nitrophenyl glycoside is 300 nm. Thererfore, energy provided by the laser and absorbed by the *p*-nitrophenyl group may have effected the fragmentation. Similar fragmentation was observed during ESMS of analogues **9** and **12**. In these situations, the degree of fragmentation was found to be dependent on the acquisition parameters (detailed in the experimental section for these analogues). It should be noted that the M_2 species in the MALDI MS spectrum of **15** does not reflect contamination by the hydrolysis product of the glycoside. This product, in which the R-OPNP aglycon is replaced with R-OH has a calculated molecular weight of 763.80 Da.

In the case of the purification of peptide **1**, MALDI MS analysis was able to detect the presence of certain deletion peptides (refer to Fig. 3.30 A). These deletetion peptides would be expected to couple with **8** during the reaction to prepare **13**. However, similar deletion analogues of **15** were not found in its mass spectrum suggesting a homogeneous composition for **15**.

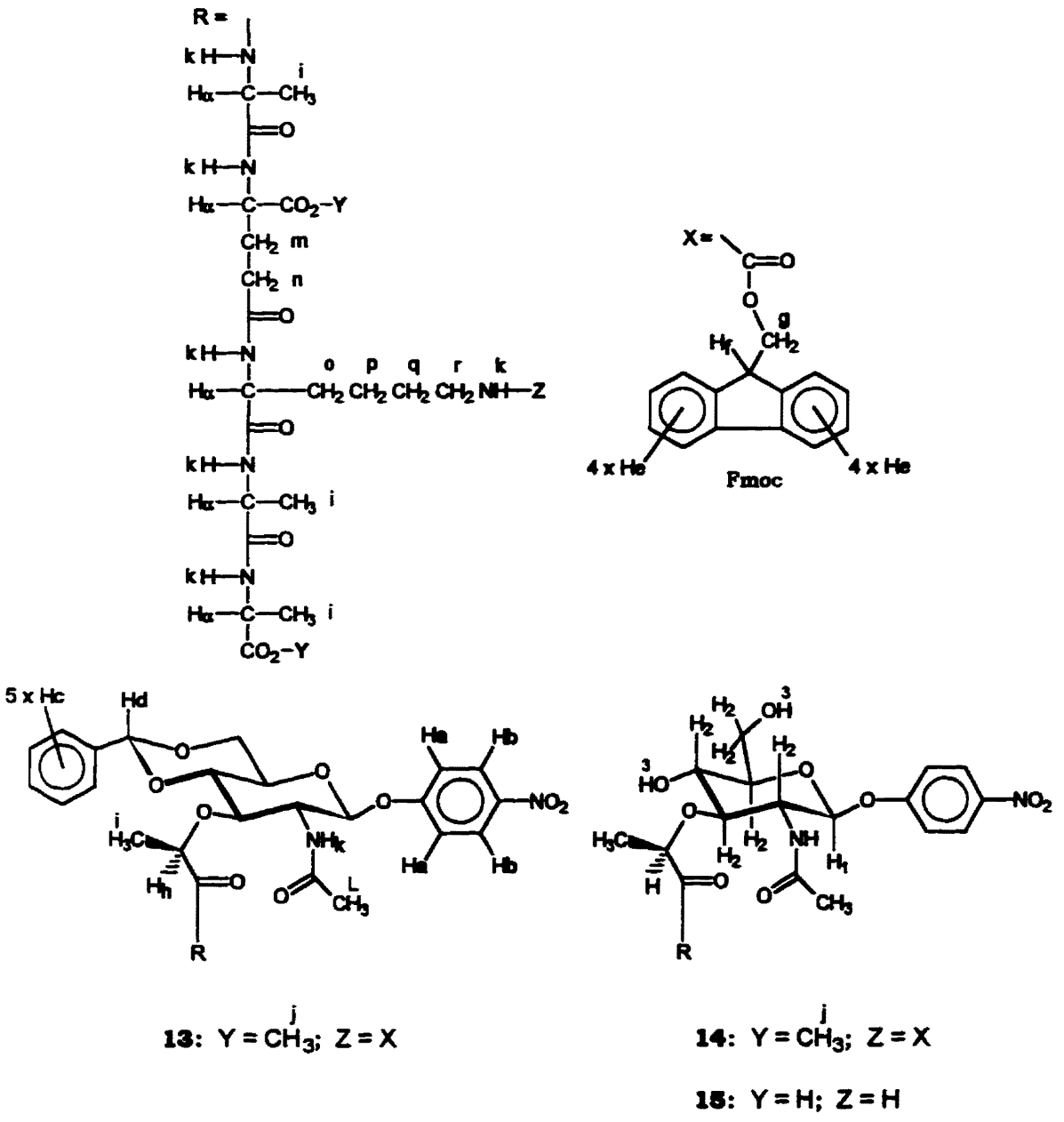
Figure 3.36. ^1H NMR of the β-PNP-MurNAc analogues **13**, **14** and **15**. The spectra appear on the three pages which follow.

(A) 500 MHz spectrum of compound **13** taken in DMSO-d_6.

(B) 500 MHz spectrum of compound **14** taken in DMSO-d_6.

(C) 200 MHz spectrum of compound **15** taken in D_2O.

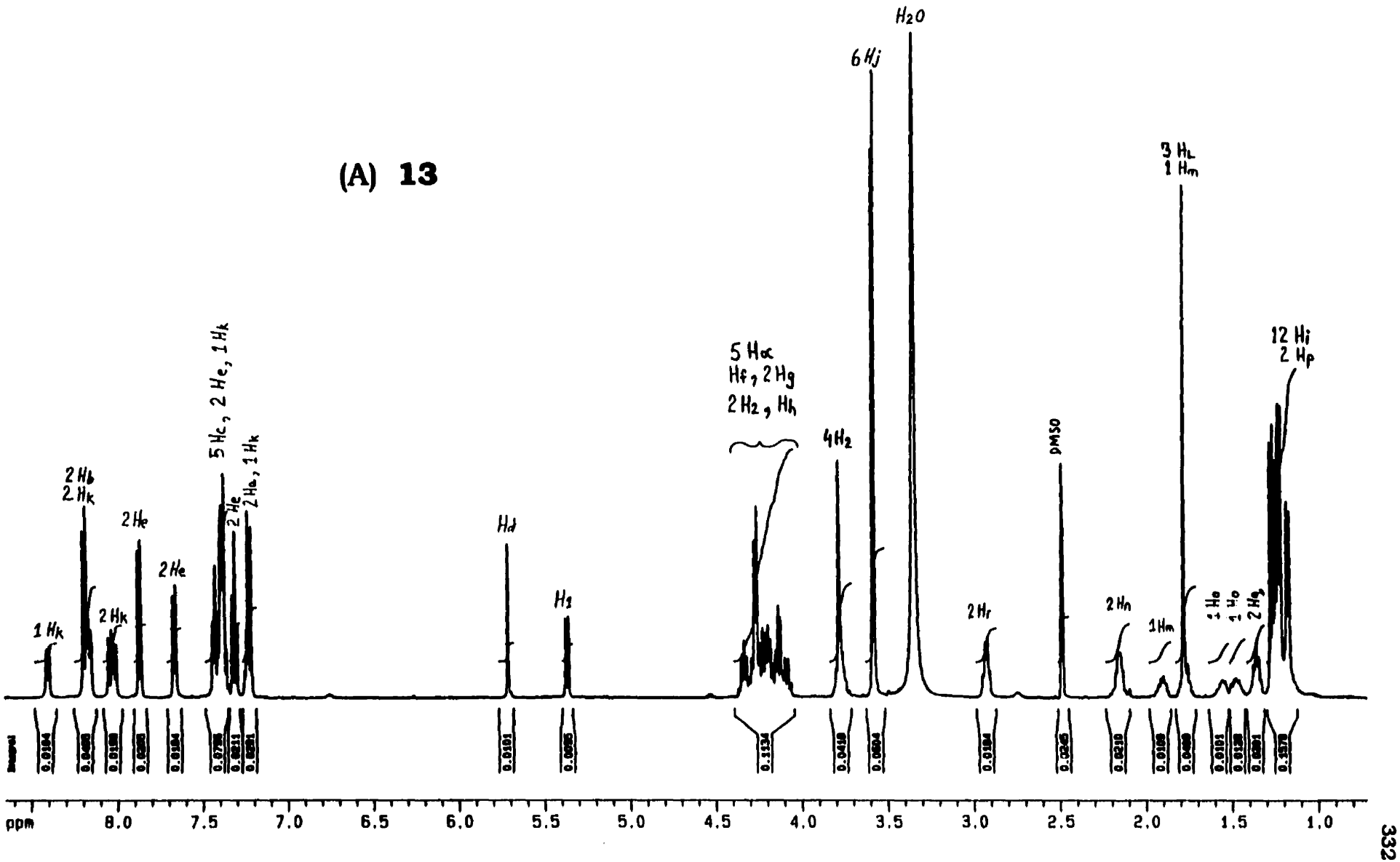
The letter or symbol designations given below represent individual or groups of protons and are used in the assignment of the resonances on the respective spectra.

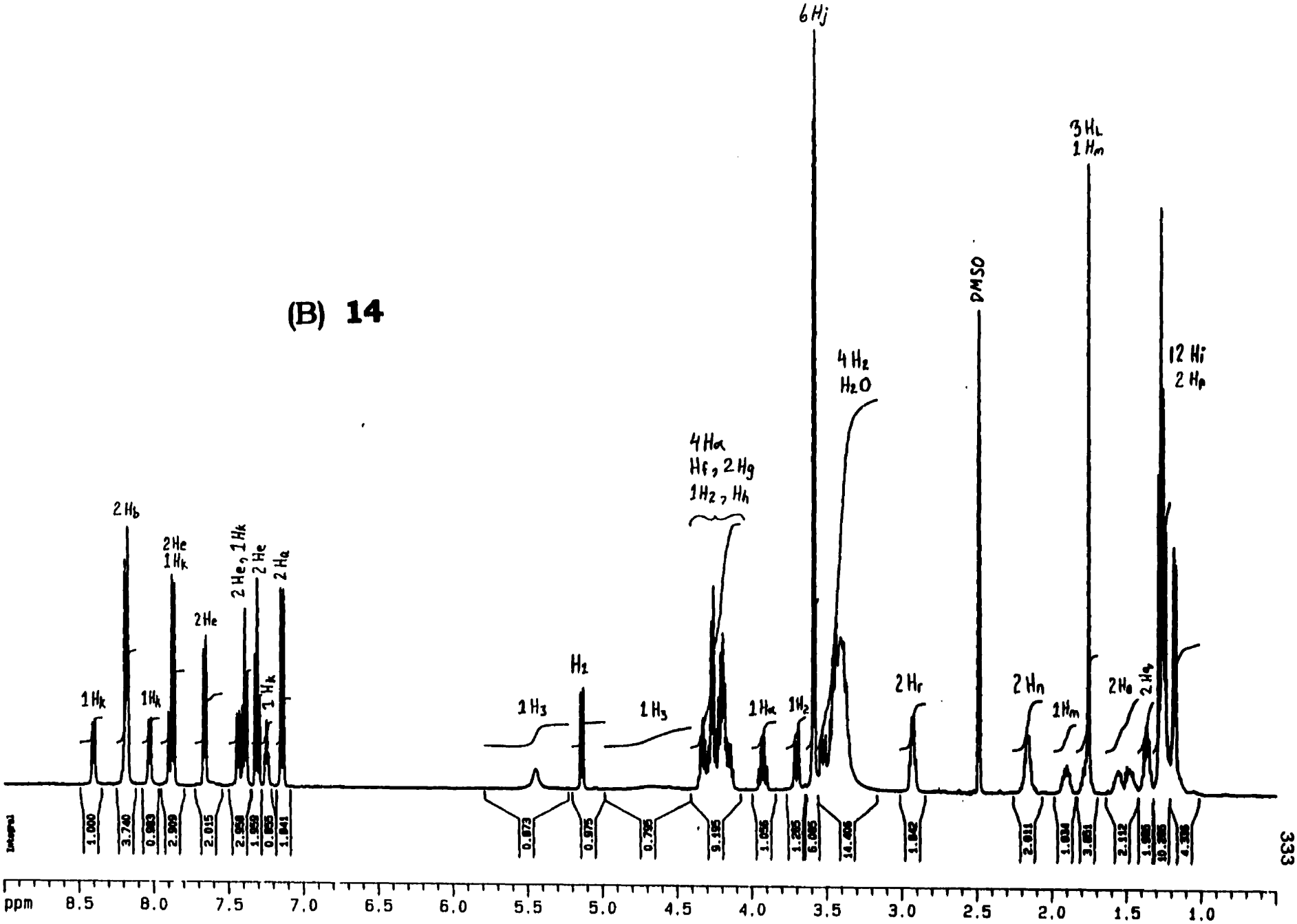


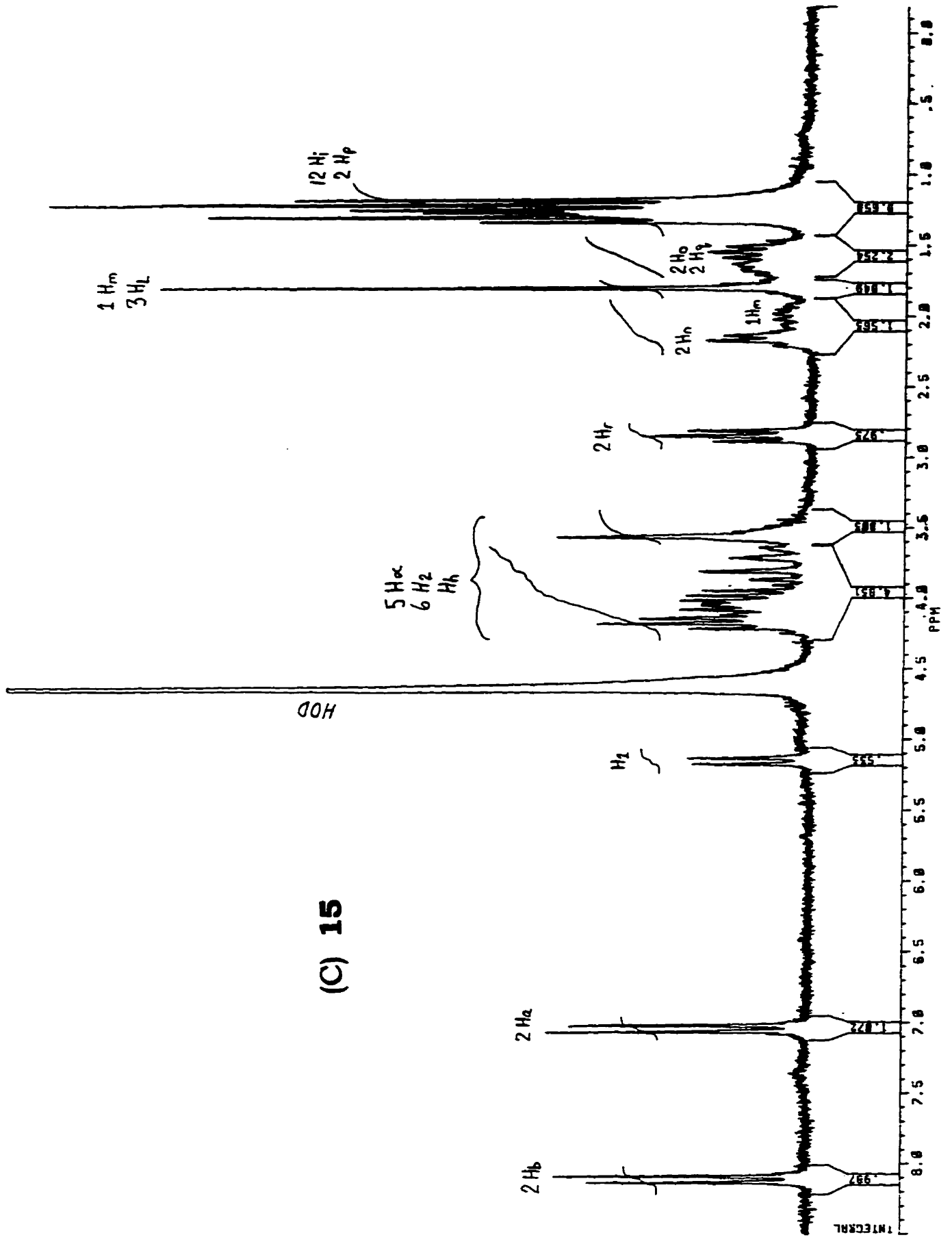
13: Y = CH_3^j; Z = X

14: Y = CH_3^j; Z = X

15: Y = H; Z = H


(A) 13
1 Hk
2 Hb
2 Hk
2 Hk
2 Hc
2 He
5 Hc, 2 He, 1 Hk
2 He
2 Ha, 1 Hk
Hd
H1
5 Ha
Hf, 2 Hg
2 H2, Hh
4 H2
6 Hj
H2O
DMSO
2 Hr
2 Hn
1 Hm
3 HL
1 Hm
1 Ho
1 Hp
2 Hq
12 Hi
2 Hp
ppm
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0


(B) 14
6 Hj
3 HL
1 Hm
DMSO
4 H2
H2O
12 Hi
2 Hp
4 Ha
Hf, 2 Hg
1 H2, Hh
2 Hb
2 He
1 Hk
2 He
2 He, 1 Hk
2 He
1 Hk
2 Ha
1 Hk
1 Hk
1 H3
H1
1 H3
1 Ha
1 H2
2 Hr
2 Hn
1 Hm
2 Ho
2 Hq
ppm
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0


HOD
12 Hi
2 Hp
1 Hm
3 HL
2 Ho
2 Hg
1 Hm
2 Hn
2 Hr
5 Hα
6 H2
Hh
H1
2 Ha
2 Hb
PPM
INTEGRAL

(C) **15**

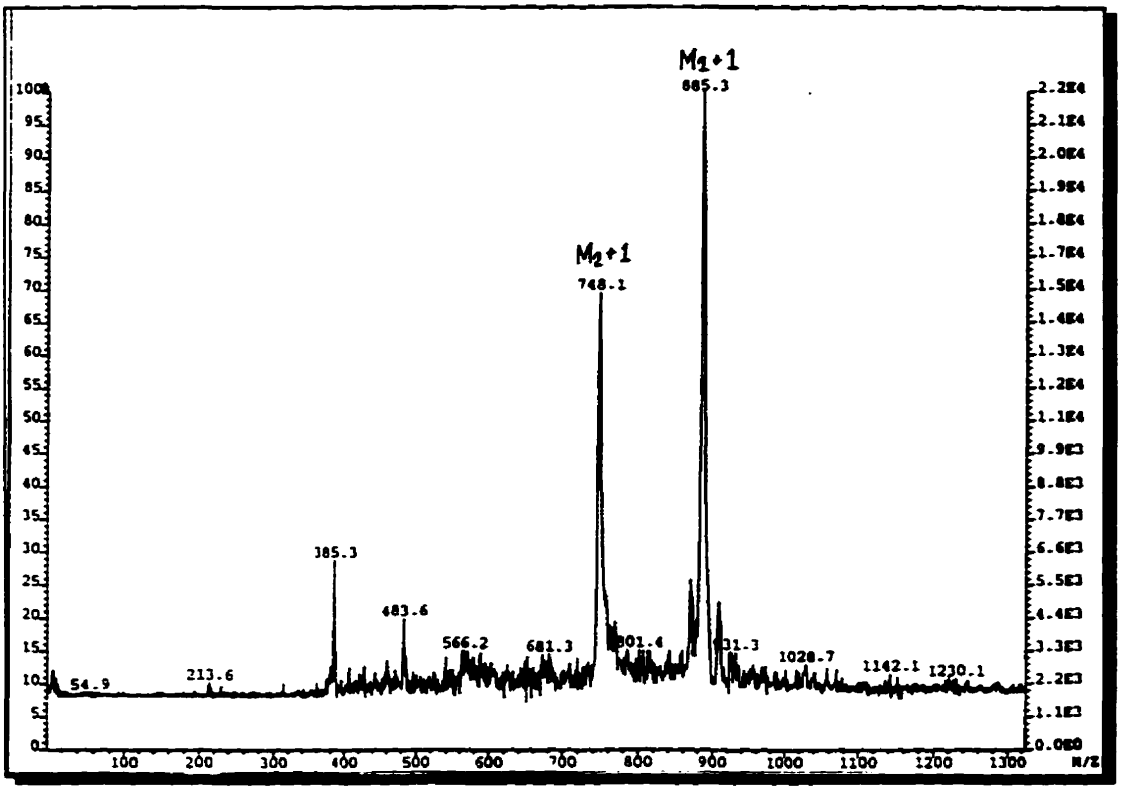


Figure 3.37. MALDI MS spectrum of compound **15**. The spectrum was acquired at a laser energy of 2 (coarse) and 110 (fine), 600 Da matrix suppression and was externally calibrated using Gramicidin S (1141.7 Da) and Kemptide (771.9 Da).

When **15** was initially purified by RP-C18 HPLC, it was determined by CZE that PNP was present as a contaminant (Fig. 3.38 A), which was generated either from possible decomposition or instability. For this reason, a second RP-C18 HPLC purification was required in order to purify **15** from the contaminating PNP. Analysis of the re-purified material by CZE indicated no contamination by PNP (Fig. 3.38 B). In addition, analysis of **15** after prolonged storage (5 months) by CZE revealed no decomposition. No other species were detected when CZE of **15** was extended for 30 min under the same conditions as given in Fig. 3.38 with the exception that the absorbance at 200 nm was monitored throughout electrophoresis (chromatogram not shown). In conjunction with the MALDI MS results, a high level of purity for **15** is suggested from the CZE analyses.

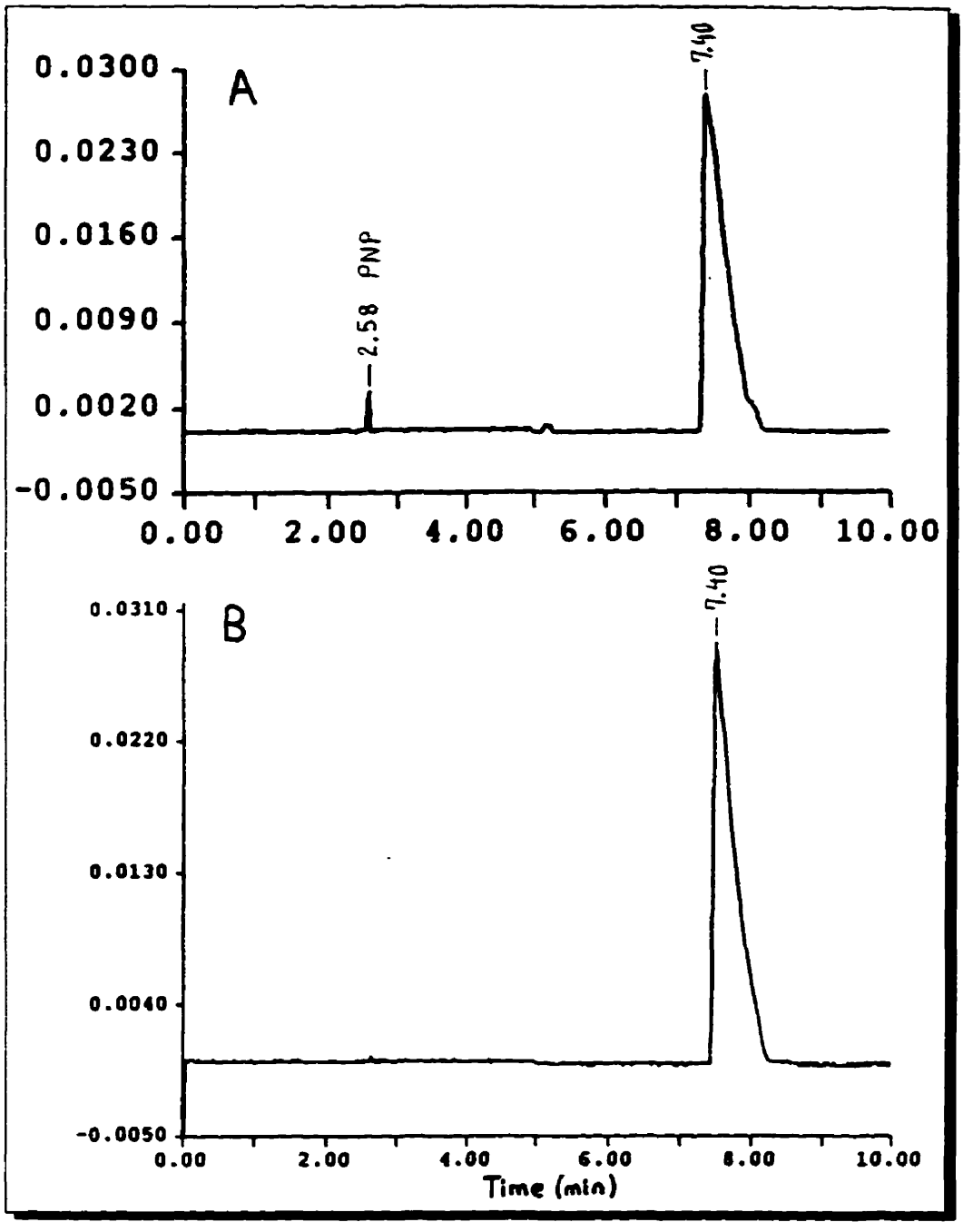


Figure 3.38. Capillary electrophoresis chromatograms of compound **15**. Chromatogram (A) demonstrates contamination by PNP while chromatogram (B) demonstrates purification of **15** from PNP.

Sample:	**15** (10 mM) in 50 mM KPB, pH 7.0
Capillary:	24 cm x 25 μM I.D. coated, 20 °C
Buffer:	BioRad 0.3 M sodium borate, pH 8.5
Injection:	(A) 15 psi x sec; (B) 20 psi x sec
Running:	12 kV
Polarity:	negative to positive
Detection:	0 to 5 min, 400 nm; 5 to 10 min, 300 nm

From experience obtained over the course of the preparation and purification of the β-PNP-MurNAc analogues, some pertinent details are offered for future reference. The PNP glycosidic bond is prone to hydrolysis especially under acidic conditions. For those analogues which were purified by RP-C18 HPLC using mobile phases containing TFA, it is suggested that the temperature of the bath used to accelerate evaporation of solvent during sample concentration (i.e. by use of a rotary evaporator) be monitored and not be permitted to exceed ≈ 40 °C. The combination of excessive heat and the presence of acid will promote hydrolysis. Secondly, individual peaks on chromatograms resulting from HPLC chromatgraphy were analyzed for their spectral properties (a feature available with the Waters 994 photodiode detector). This permitted the identification of compounds and of their retention times during chromatography by the characteristics of their UV spectra. Representative spectra of selected compounds are illustrated in Fig. 3.39. In all cases, PNP glycosides demonstrate an absorption maxima at 300 nm. Of interest, note how the spectra of **13** and **14** are essentially a combination of the spectra of a PNP-glycoside (eg. see spectra of **9**) and of peptide **4** (peptide **4** possesses the Fmoc functionality which absorbs at 265, 289 and 300 nm. The spectrum of peptide **1** is very similiar to that of peptide **4**). Finally, several of the analogues and peptides were purified by RP-C18 HPLC using the μBondapak™ C18 Prep-Pak® cartridge. Retention times of compounds were not absolute and showed some dependency on the amount of material applied to the column. In addition, over the course of this work, a new cartridge was purchased and it was found that compounds demonstrated slightly stronger retentions using the new cartridge as compared to an older cartridge under identical conditions. Therefore, the details given in the experimental section (3.2.5 and 3.2.6) concerning elution times of peptides and MurNAc analogues should be viewed as accurate but slight variations can be expected if the purifications are to be repeated by others.

3.3.5. Evaluation of the β-PNP-MurNAc Analogues as Substrates

We had envisioned that by preparing the β-PNP-muramyl analogues, the greater structural similarity of these compounds to the peptidoglycan would direct them to the position in the active site of LaL analogous to the D site of HEWL. Then, if there existed sufficient interactions between the enzyme and the substrates that are required for catalysis, cleavage might occur releasing PNP as a method for measuring enzyme activity. It would have then been possible to vary the peptide-substituent and obtain more descriptive kinetic measures on the importance of the peptide and the kinetic parameters

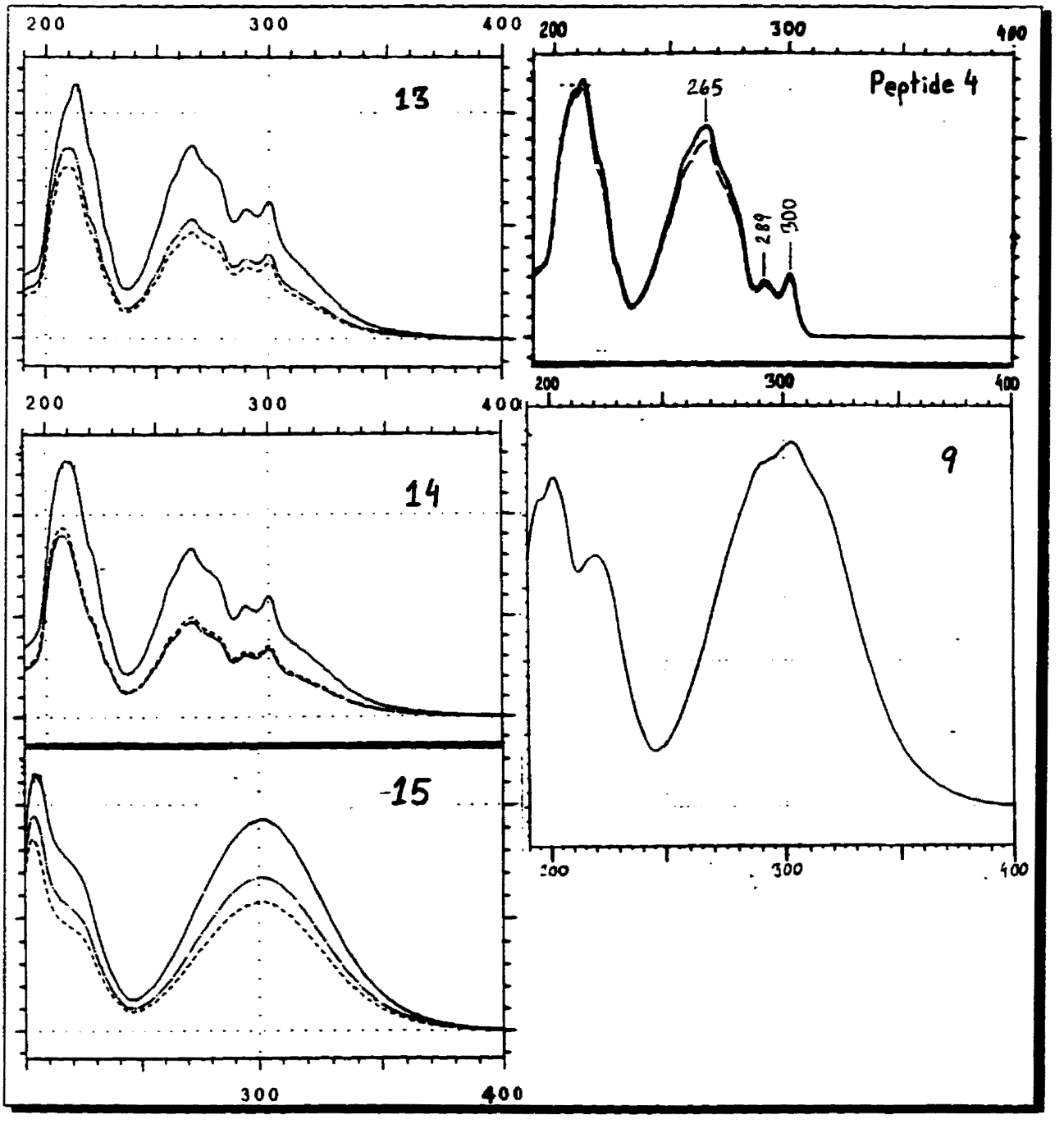
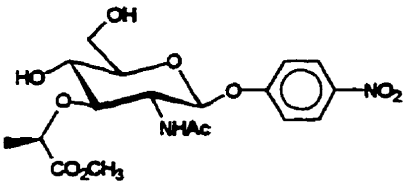
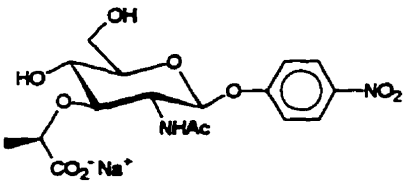
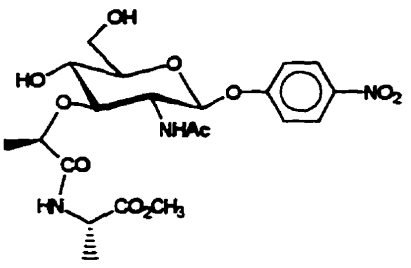


Figure 3.39. UV spectra of selected compounds. The abscissa is wavelength (nm).

for LaL with these compounds. Secondly, it was thought that the peptide-substituted analogues could lend themselves to the chemoenzymatic preparation of some novel 1,6-anhydro muramyl peptides for biological investigations.

The results pertaining to the β-PNP-muramic acid analogues **9**, **10** and **12** are shown in Table 3.15. As can be seen neither LaL nor HEWL catalyzed cleavage or

Table 3.15. Evaluation of compounds **9**, **10** and **12** with LaL and HEWL.

Compound [a]	LaL [b] Absorbance		HEWL [c] Absorbance	
	Blank	With LaL	Blank	With HEWL
9 	0.293 ± 0.003	0.296 ± 0.008	0.250 ± 0.001	0.254 ± 0.001
10 	0.177 ± 0.005	0.172 ± 0.005	0.174 ± 0.006	0.167 ± 0.003
12 	0.041 ± 0.002 0.113[f] ± 0.009	0.041 ± 0.001 0.118[f] ± 0.006	0.039 ± 0.002	0.039 ± 0.000

Enzyme (1 mg/mL) and enzyme blanks were incubated at 37 °C for 47-48 hr with each compound. Aliquots of the samples (200 µL or 500 µL[f]) were diluted with 0.1 N NaOH (to 1.0 mL) and the absorbance at 400 nm recorded (performed in triplicate).

[a] Included at concentrations of 5 mM.

[b] Performed in 50 mM KPB, pH 7.0. [c] Performed in 25 mM NaOAc, pH 5.0.

hydrolysis of PNP from these compounds that was detectable above the enzyme control. Although the activity of HEWL conducted with **10** had been previously reported with negative results (Jeanloz et al., 1968), the evaluation of **9** and **12** with HEWL is presented here for the first time.

Unfortunately, the peptide-substituted derivative β-PNP-MurNAc-*L*-Ala-*D*-*iso*-Glu-*L*-Lys-*D*-Ala-*D*-Ala **15** was also found not to be susceptible to the action of LaL or HEWL.

Because of the limited quantities of **15** prepared, the evaluation of **15** as a substrate was performed by CZE. It was estimated that release of PNP from 0.02% of **15** was required in order for its detection under the experimental conditions employed (see 3.2.7.2). After incubation of **15** for 24 and 48 hr with LaL and HEWL respectively, no PNP could be detected.

Although the negative results obtained with the prepared analogues were quite discouraging considering the extensive effort dedicated towards their syntheses, the negative results in themselves present some insight into the minimal requirements of a substrate for LaL. It is not surprising that the non-substituted or only alanine substituted analogues **9**, **10** and **12** were unproductive. Monosaccharide-based compounds, such as the β-phenyl (Berger & Weiser, 1957) and β-*p*-nitrophenyl glycosides of GlcNAc (Osawa, 1966) or the β-phenyl glycoside of MurNAc (Yamamoto et al., 1963) are not substrates for HEWL. The fact that our analogues are not substrates for LaL suggests that, like HEWL and the simple monosaccharide substrates, there exist insufficient interactions with the enzyme either for binding or absence of developed binding energy that are channeled toward or are obligatory for catalysis (for example to promote distortion of the substrate). Even in the case of the peptide substituted analogue **15**, where, if the assumption is made that the peptide directs the compound to the appropriate position in the active site and fulfils the peptide binding requirements, it appears that more than one saccharide is required to effect catalysis. The more complete occupation of binding sites (i.e. analogous to filling sites A-D in HEWL) might anchor the subtrate in the active site of and provide crucial physical restraints to the terminal saccharide that is involved in the bond breaking and forming processes. The presence of only one additional saccharide is sufficient to render $(GlcNAc)_2$-PNP as a substrate for HEWL (Osawa, 1966).

3.3.6. Summary and Future Work

The work presented in Chapter 3 provided a prelimary as well as, in several aspects, some detailed examinations into the interactions of saccharides and peptides with LaL. To initiate these studies, a dependable and accurate method for the measurement of LaL activity was required. We have developed and documented methods for the preparation and characterization of *E. coli* substrate cells in considerable detail for use in measuring the bacteriolytic activity of LaL turbidimetrically. Although the assay does not lend itself to the measurement of specific activity because of "batch-to-batch" variations in the substrate cells, the assay was sufficiently standardized to permit reliable and reproducible descriptions of relative enzyme activity. We have demonstrated that the lysis of the substrate cells, and hence, the assay, is influenced by ionic strength which must be taken into consideration for any studies employing compounds which could affect ionic strength.

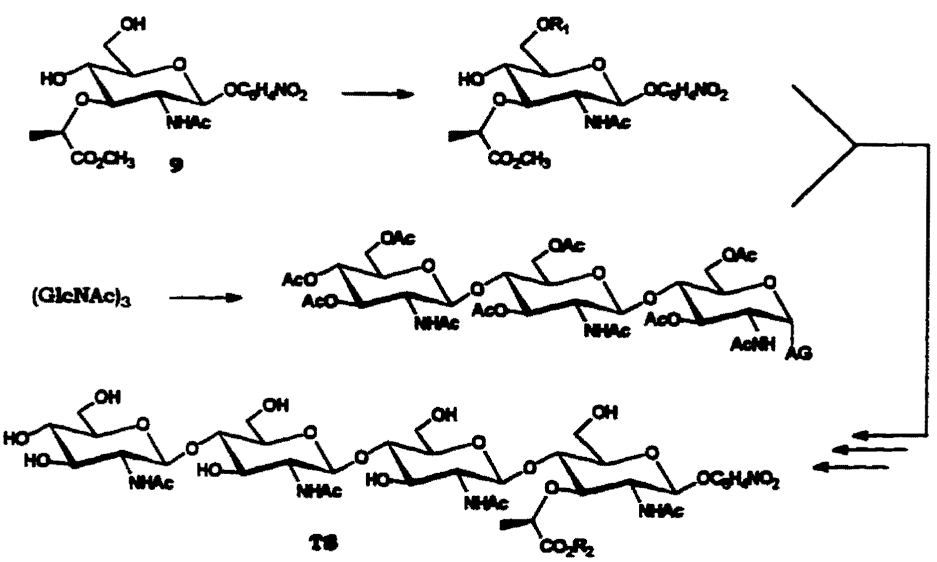
Taken together, our findings offer substantial evidence that LaL is indeed capable of binding N-acetylglucosamine oligosaccharides but is incapable of cleaving these sugars. Not only have we demonstrated that these saccharides are capable of inhibiting the bacteriolytic activity of LaL and do so more efficiently with increasing size, but specific interactions with LaL are suggested from the NMR studies on the ^{13}C-methionine labelled enzyme as well as the fluorescence and calorimetric investigations with the wild type enzyme. The ability of these saccharides to protect LaL from inactivation by the carboxyl specific reagents EDC and EAC offers strong indications that these saccharides bind to the active site of LaL and support involvement of acidic residues in the peptidoglycan cleavage activity of LaL. Further work could centre on the determination of the important acidic residues by labelling with a radioactive carbodiimide (followed by proteolysis and peptide sequencing or ESMS) as well as investigating other group specific modifying reagents.

The data obtained from the inhibition, fluorescence and protection studies also suggest that a marked difference exists in the mode of binding for the two groups of $(GlcNAc)_{n\leq 3}$ and $(GlcNAc)_{n\geq 4}$. The results could indicate that more elaborate interactions with LaL are satisfied for the latter group that guide them to a critical region in the active site. Surprisingly, $(GlcNAc)_n$ are not capable of inhibiting the bacteriolytic action of T4L or of GEWL nor are they substrates for these lysozymes (Mirelman et al., 1975; Arnheim et al., 1973b) while these oligomers are well established as inhibitors and substrates for HEWL (Rupley, 1967; Sharon, 1967). Our findings that these oligomers do bind and inhibit, but are not substrates for LaL support the recent implication of an evolutionary link between the lysozymes from hen egg white and phage T4 (Jespers et al., 1992).

The preparation of the *E. coli* cell wall peptidyl mimics has provided the opportunity to explore the possible role of the peptide portion of the peptidoglycan in the action of LaL. The apparent abilities of peptides **2** and **3** to enhance and not to inhibit the bacteriolytic activity of LaL demonstrates a curious and active participation of peptide. This possible role as a positive modulator on activity was more directly established from the observation that LaL could only cleave $(GlcNAc)_5$-PNP in the presence of peptide **3**. These results must be confirmed with the identification of a 1,6-anhydro containing product which would not only establish the involvement of LaL but would also verify its atypical mechanistic property resulting in 1,6-anhydro production.

The effects demonstrated by the peptides are sufficiently convincing to warrant futher synthesis of peptide-containing substrates. A soluble, chemically-defined substrate for LaL will be indispensable for obtaining kinetic information leading to a better insight into the specificity and mode of action of this lysozyme. Our efforts towards this objective have provided improved and convenient methods for the synthesis of β-PNP-GlcNAc and β-PNP-MurNAc. Although the β-PNP-MurNAc analogues prepared were ineffective as substrates for LaL, we believe that a successful substrate can be prepared by extending the carbohydrate chain. It is suggested that the synthesis of a tetrasaccharide (see **T8**, Scheme 3.12) should be primarily attempted since it appears, as mentioned above, that a tetrasaccharide might provide the obligatory enzyme-carbohydrate interactions required.

A possible synthetic strategy is suggested in Scheme 3.12. The analogue **9** could act as the acceptor following protection of the C6-hydroxyl (R_1). The glycosylation donor should centre on the per-acetate of $(GlcNAc)_3$ although the choice of activating group (AG) at the anomeric position will require exploration.



Scheme 3.12.

The desired glycoside synthesis may prove difficut due to the known low reactivity of the C4-hydroxyl group of a glucosamine in the pyranose form (Kusumoto et al., 1986). However, other methods to achieve this coupling and which may be applicable to the proposed synthesis have appeared and the literature on this topic should be consulted (Durette et al., 1979; Ogawa et al., 1981; Paulsen, 1982; Schmidt, 1986; Termin & Schmidt, 1992; Banoub et al., 1992). After the glycoside is made, the methyl ester can be removed and the acid coupled with different peptides to afford variation at the R_2 position in **TS**. A recent report on the synthesis of a differentially protected form of *meso*-diaminopimelic acid and of the peptide *iso-D*-Glu-*L-meso*-DAP (Williams & Yuan, 1994) may be applied to the preparation of peptides that would then be of greater homology to the peptide portion of the *E. coli* peptidoglycan. With good fortune, a derivative of **TS** will be found that will serve as an effective substrate for LaL. The development of such a substrate would present many new avenues to explore with LaL (primarily kinetic investigations) and would also be of considerable interest to others who are involved in research with other phage lysozymes or bacterial muramidases.

Our current research interests with LaL are centering on the enzyme in which we have incorporated trifluoromethionine and which has presented a need for the preparation of site-specific mutants of LaL (Chapter 4). More elaborate descriptions on the kinetics of the labelled and mutant enzymes are required to fully establish the significance of the methionine substitutions and of the mutations respectively, which further demonstrates the need for a defined substrate.

CHAPTER 4

Incorporation of Trifluoromethionine into Lambda Lysozyme at High and Low Levels: Characterization, ^{19}F NMR and Cyanogen Bromide Reaction of the Labelled Proteins

4.1. INTRODUCTION

4.1.1. Methionine in Proteins

Over the past several years our appreciation of the important, although sometimes subtle, roles that the amino acid methionine plays in protein structure and function has dramatically increased (Gellman, 1991; Viguera & Serrano, 1995). Although methionine is the third least-abundant amino acid found in proteins (Doolittle, 1989), its presence serves a variety of roles in biological systems. Methionine or N-formylmethionine is present as the initiating amino acid in translation events (RajBhandary, 1994). The thioether moiety of methionine functions as a ligand to metal centers. Methionine has been found to act as a ligand in iron centres such as in cytochromes c (Wallace & Clark-Lewis, 1992), *c*-550 (Ubbink et al., 1994) and *c*-553 (Senn et al., 1983) and as well to copper centers such as is found in plastocyanin (Collyer et al., 1990; Guckert et al., 1995), several azurins (Murphy et al., 1993; Salgado et al., 1996), nitrite reductase (Godden et al., 1991), ascorbate oxidase (Messerschmidt et al., 1992), dopamine β-hydroxylase (Reedy & Blackburn, 1994) and peptidylglycine α-hydroxylating monooxygenase (Eipper et al., 1995).

Methionine has the distinction of displaying superior flexibility in its side chain than the other hydrophobic amino acids. This flexibility depends not only on the lack of branching but more importantly, on the properties of the thioether (Gellman, 1991). The longer length of a C-S (1.8 Å) than a C-C (1.5 Å) bond leads to a decrease in the steric demands on the methyl group, than in for example, leucine, isoleucine or norleucine. This diminution of steric repulsion allows the χ_3 torsion angle (Fig. 4.1 C) to adopt a greater proportion of available angles. For *n*-butane, there is almost a 1 kcal/mol energy barrier between the anti and gauche conformations leading to a preference for the anti>gauche>>eclipsed conformations (Fig. 4.1 A). However, for ethyl methyl sulfide (the side chain of methionine) this rotational barrier is small (Fig. 4.1 B) and there are no

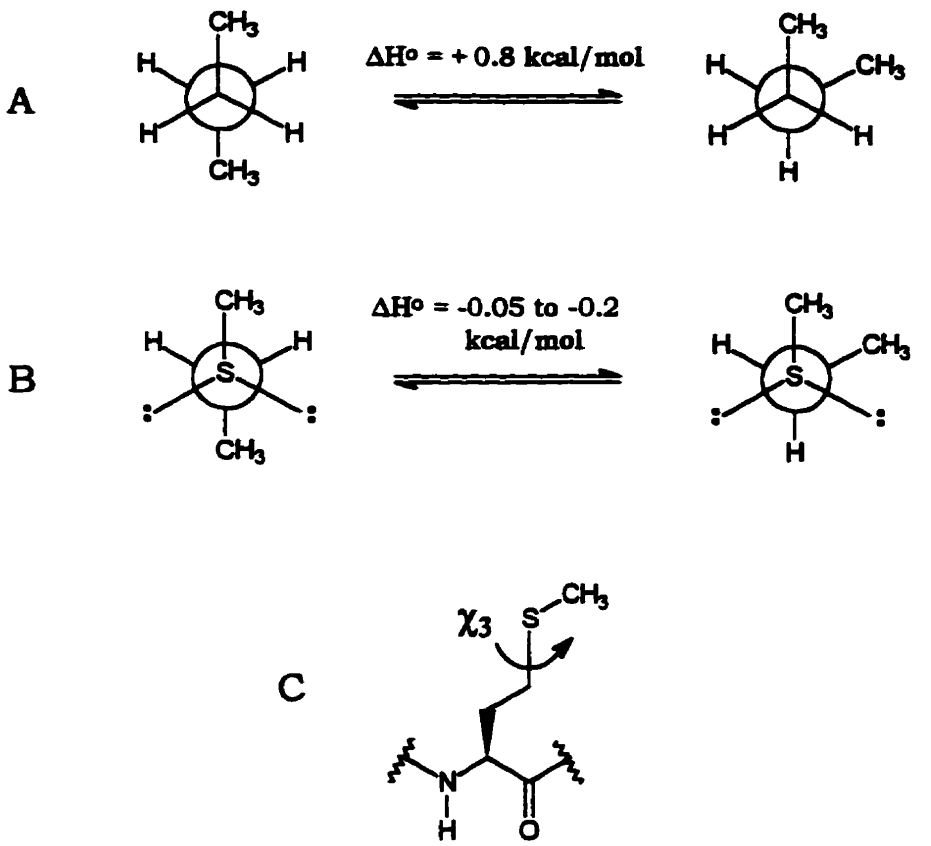


Figure 4.1. Anti and gauche conformations of (A) *n*-butane ($CH_3CH_2CH_2CH_3$) and (B) ethyl methyl sulfide ($CH_3SCH_2CH_3$). (C) A methionine residue indicating the χ_3 torsion angle. Adapted from Gellman (1991).

significant restrictions on a particular χ_3 angle (except for χ_3 values near 0° with Cε and Cβ eclipsed). In fact, a statistical survey of crystalline proteins has revealed that essentially any angle is sampled equally well (Janin et al., 1978). This is in contratst to the χ_3 torsion angles of leucine and isoleucine which demonstrate pronounced periodicities (Janin et al., 1978). Futhermore, the presence of the sulfur atom provides a centre of much greater polarizability than a methylene equivalent. Therefore, methionine has the potential to interact by both polar and nonpolar interactions. In the latter case, the polarizable sulfur atom can lead to stronger interactions from the increase in dispersion forces between methionine and another hydrophobic fragment.

Nature has apparently taken advantage of the unique properties of methionine in a variety of substrate recognition processes. A number of proteins appear to utilize methionine's flexibility as well as its polarizable sulfur atom for use in recognition and

binding events. Methionine residues are now known to act as important structural elements in the sequence-independent recognition of non-polar peptide and protein surfaces (Gellman, 1991). Signal recognition particle 54 (SRP54) is responsible for the recognition and binding of a variety of signal sequences found in nascent secretory and membrane proteins. The sequence-independent recognition of these signal peptides by SRP54 is believed to be due, in part, to the presence in SRP54 of three amphiphilic helices containing hydrophobic faces rich in methionine residues. These methionine rich regions, or M-regions, are believed to form the basis of a binding pocket for these signal sequences (Bernstein et al., 1989; Lutcke et al., 1992). In the case of calmodulin, which binds several different peptides unrelated in sequence, recent X-ray (Meador et al., 1992) and NMR (Ikura et al., 1992; Zhang & Vogel, 1994) studies of calmodulin-peptide complexes indicate that all nine methionine residues in the protein interact directly with the peptide ligand. A similar interaction is believed to occur in the case of a related protein, troponin C (Lin et al., 1994). In addition, a recent report has implicated methionine residues in the collagen/gelatin binding of plasma fibronectin (Miles & Smith, 1993). On the basis of these general features it was suggested that protein regions rich in methionine can provide a malleable nonpolar surface adaptable for binding peptide ligands of varying dimensions (Gellman, 1991).

As important for recognition and binding but not perhaps as dramatic are the proteins that appear to utilize single methionine residues in their active sites for the recognition of carbohydrate substrates and ligands. Methionine 108 (Fig. 4.2 A) in the arabinose-binding protein from *E. coli* has been shown by X-ray crystallography and site-directed mutagenesis to play a direct role in binding of this monosaccharide (Vermersch et al., 1991). The preference of the protein for arabinose (arabinose > galactose >> fucose) is altered in the M108L mutant (galactose >> fucose > arabinose). It is believed that this selectivity is governed by both a "nonpolar-nonpolar" as well as a "polar-nonpolar" interaction with the sugar and the methionine residue. When arabinose is bound, the methyl group of M108 "swings" toward the sugar to optimize nonpolar contact with the sugar (Vermerch, 1991). Other enzymes which are involved in saccharide binding and have a methionine in their active sites have been found. The crystal structure of the *E. coli* exomuramidase Slt70 (Thunnissen et al., 1994, 1995a), which is mechanistically related to LaL with respect to 1,6-anhydro bond formation, showed Met498 in close proximity to the active site directed inhibitor bulgecin (Fig. 4.2 B). Located near Trp108 and the C and D subsites is found Met105 of HEWL, which experiences small changes in

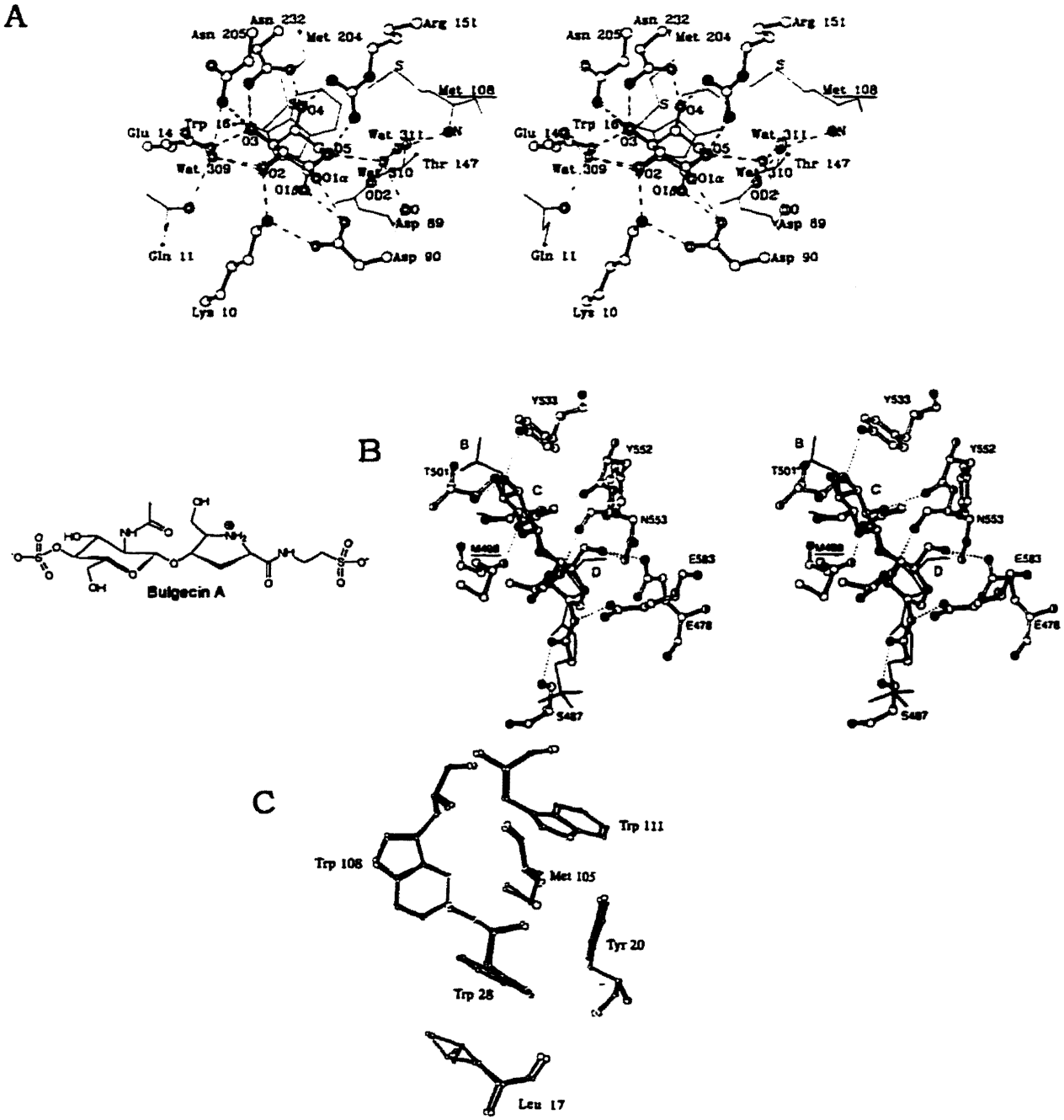


Figure 4.2. Methionines found in the carbohydrate binding sites of some proteins.

(A) Arabinose binding protein with bound arabinose showing M108. Reproduced from Quicho et al. (1989).

(B) Slt70 with bound bulgecin A (drawn in bonds only) with superposition of GlcNAc-MurNAc structure (drawn in ball and stick) showing M498. Reproduced from Thunnissen et al. (1995a).

(C) Conformational changes of M105 of HEWL upon $(GlcNAc)_3$ binding. The unbound enzyme is shown in bold and the $(GlcNAc)_3$ bound enzyme is shown in outline. Reproduced from Lumb et al. (1994).

its side chain conformation (Fig. 4.2 C) and in the chemical shift of the εCH_3 protons on binding $(GlcNAc)_3$ (Lumb et al., 1994). In addition, M106 is found in the peptidoglycan-binding region of the active site of T4 lysozyme (refer to Fig. 3.4, p. 196).

The ability of methionine residues to undergo facile oxidation to the sulfoxide level may play a role in the control of protein function. Several proteins are known to have altered biological activities when one or more methionine residues have been oxidized (Brot & Weissbach, 1991). The α-galactosidase from *Trichoderma reesei* shows a 12-fold increase in activity as a result of specific oxidation of one of its five methionines (Kachurin et al., 1995). The effect of oxidation, such as in the case of the lung elastase α-1-proteinase inhibitor, can become clinically relevant (Travis & Salvesen, 1983). A protein-methionine-S-oxide reductase (EC 1.8.4.6) appears to mediate thioredoxin-dependent reduction of methionine sulfoxide residues in proteins (Moskovitz et al., 1995). Although the effects of methionine oxidation are well documented, the extent of the contribution of methionine oxidation-reduction in the control of enzyme activity *in vivo* is currently unknown (Vogt, 1995). A recent report on the possible control of calmodulin activity by *in vivo* oxidation of one of its methionine residues would appear to favor this novel mechanism in the modulation of protein activity (Yao et al., 1996).

4.1.2. ^{19}F NMR of Proteins

The application of solution ^{19}F NMR has become an influential and versatile approach to the study of protein structure and function. In recent years, the application of ^{19}F NMR studies are increasing and have been highlighted in timely reviews on this topic (Sykes & Hull, 1978; Gerig, 1989, 1994; Danielson & Falke, 1996). Although the number of protein crystal structures is continuously expanding, solution studies on proteins by multidimensioal NMR studies are restricted to those of limited size (< 30 kDa) and requires at least, a partial to complete assignment of the protein resonances. The introduction of fluorine labels into a select number of protein sites and application of ^{19}F NMR can be useful to detect conformational changes in proteins, especially of use for those proteins that are not amenable to multidimensional NMR techiques.

There are several features of incorporating fluorine labels into proteins which are desirable in application with ^{19}F NMR (Gerig, 1994; Luck & Falke, 1991a; Danielson & Falke, 1996). These include:

1. The spin ½ ^{19}F nucleus occurs at a natural abundance of 100% and has a sensitivity to NMR detection of 83% that of ^{1}H.

2. Because ^{19}F is absent in solvent and, unless introduced by design, in proteins, there are no background signals to contend or interfere with the spectroscopy.

3. The van der Waals radius of covalent fluorine (0.14 nm) is not that different from ^{1}H (0.12 nm). Therefore, the single replacement of a fluorine for a proton is not expected to be sterically too demanding on a protein (see also Gregory & Gerig, 1991). However, a functionality containing a group of fluorines (eg. CF_3, C_6F_5) is collectively larger than the corresponding proton system and may disrupt protein structure for reasons of steric bulk.

4. The ^{19}F chemical shift range is 100-fold that of ^{1}H and approaches 1000 ppm. This wide range is accompanied with a resolution that generally yields well-resolved ^{19}F resonances in one-dimensional spectra.

5. Controlled largely by the fluorine lone-pair electrons, which provide a large paramagnetic term in the shielding formula, the ^{19}F chemical shift is extremely sensitive to its non-bonded environment providing a highly sensitive indicator of local structural changes.

6. Although it may be difficult to obtain large amounts of a fluorine-labelled protein, one-dimensional ^{19}F NMR studies generally require lower protein concentrations and shorter spectral acquisition times than do multi-dimensional NMR experiments.

Novel fluorinated amino acid incorporation into peptides and proteins can be achieved by several methods both chemically and biosynthetically. Chemical approaches can include synthesis of polypeptides containing the analogue (Koul et al. 1979; Taylor et al., 1985), by the use of selectively acylated suppressor tRNA's for *in vitro* protein expression (Noren et al., 1989), and by direct labelling of amino acid side chains with specific organic fluorinating reagents (reviewed by Gerig, 1994). However, limitations in chemical approaches have rendered the bioincorporation of fluorinated amino acids into proteins as the method most widely used. If low levels of incorporation are adequate, this can be achieved by simply including the analogue in the diet of the organism. For example, sufficient levels of incorporation of fluorinated amino acids into avian lysozymes for ^{19}F NMR studies has been achieved by inclusion of the requisite fluorinated amino acid in the diet of the birds (Lian et al., 1994). However, the availability of amino acid auxotrophic strains of microorganisms (Sykes & Weiner, 1980) and modern recombinant expression systems have provided a more regulated approach to the introduction of fluorinated amino acids into a wide variety of proteins (Gerig, 1994).

The chemical synthesis of fluorinated derivatives of most of the common amino acids have appeared (see Gerig, 1994). Several of these analogues have been shown to be useful in the study of protein structure/function relationships. Collagen biosynthesis has been examined with fluorinated analogues of lysine and proline (Shirota et al., 1977a,b). A semi-synthetic ribonuclease-S containing a 4-F-histidine residue was catalytically inactive yet was found by crystallography, to be structurally similar to the native, active protein (Taylor et al., 1981).

Current interests with fluorine-labelled proteins are primarily associated with ^{19}F NMR investigations. However, it is interesting that a greater variety of the amino acids have not received application in these studies. The predominant ^{19}F NMR studies reported to date have centred on proteins containing fluorinated analogues of the aromatic amino acids tryptophan (4-, 5- and 6-fluoro), phenylalanine (2-, 3- and 4-fluoro) and tyrosine (3-fluoro) (Gerig, 1994; Danielson & Falke, 1996). A paucity of studies dealing with the application of other fluorinated amino acids to ^{19}F protein NMR currently exists. Only very recently have ^{19}F NMR investigations on dihydrofolate reductase containing 5-F-leucine been reported (Feeney et al., 1996).

The usage of only the ^{19}F-labelled aromatic amino acids has nonetheless demonstrated their versatility as ^{19}F NMR markers. The high sensitivity of fluorine resonances to environment is exemplified in the complex of $(GlcNAc)_3$ with 4-F-Trp labelled HEWL, in which a 16.8 ppm chemical shift range of the fluorine resonances was observed (Lian et al., 1994). This study also compared the ^{19}F NMR spectra from several avian lysozymes and the interspecies similarities, especially for the 4-F-Trp labelled enzymes, was startling. In a series of papers on the 3-F-Phe and 5-F-Trp labelled D-galactose/D-glucose chemosensory receptor, the application of ^{19}F NMR to probe conformational changes has been elegantly described (Luck & Falke, 1991a,b,c). The crystal structure of the receptor has been determined for the conformation that contains bound glucose, but the structure of the apo-protein is not known. The positions of the 12 targeted probe locations (i.e. five Trp and seven Phe) are well distributed throughout the protein and therefore serve as markers for individual regions of the protein. The ^{19}F NMR studies on the labelled protein provided clear indications as to both global and more localized conformational changes in the protein that occurred upon ligand binding.

The preferential use of the labelled aromatic amino acids may be largely due to practical considerations, such as the commercial availability of each of the aromatic analogues. A more important consideration may pertain to the permissiveness of the cellular biosythetic machinery to incorporate the fluorine containing amino acid into a protein. Most notable are the questions of how efficient are the transport of the analogue into the cell and the charging of the analogue onto the respective tRNA by amino-acyl tRNA sythetases. If these systems are compromised, low levels with respect to both incorporation and protein expression may result. Although some difficulties have been expressed with use of non-auxotrophic strains (Kimber et al., 1977), an auxotroph grown in medium containing controlled concentrations of both the natural and the ^{19}F-labelled aromatic amino will typically express the desired protein in modest to good yields and incorporation levels. Incorporation of 3-F-Phe has been achieved at a level of 20% (Luck

& Falke, 1991a) and appears to be the least well incorporated of the aromatic amino acid analogues. In contrast, 5-F-Trp has been incorporated at 65% (Luck & Falke, 1991a), 6-F-Trp at 93% (Li et al., 1989), while an impressive level of 95-97% has been reported for the incorporation of 3-F-Tyr (Parsons & Armstrong, 1996).

4.1.3. Scope and Objectives

Our laboratory has been involved with the investigation of the biochemical importance of the amino acid methionine. A previous member of this laboratory, Dr. Michael Houston, had prepared a variety of fluorinated methionine and methionine salvage pathway analogues and some small analogue-containing peptides (Houston & Honek, 1989; Houston et al., 1991; Houston, 1992). To continue these efforts and because of our current interest in investigating the subtleties methionine plays in protein structure and function, we have investigated the incorporation of the analogue *L*-S-trifluoromethylhomocysteine (*L*-trifluoromethionine; *L*-TFM) into LaL.

These efforts were intended to address some fundamental questions. Can a microorganism synthesize a protein containing TFM and will there be preferential incorporation at a given position? Is it possible to selectively control the level in which the analogue is incorporated into a protein? Will the TFM-labelled protein fold into an active conformation and are there any effects on enzyme activity or ligand interactions? Does replacement of a CH_3 group with CF_3 bear any consequences on protein structure? Can incorporated TFM serve as a new marker for the study of ligand-protein interactions by ^{19}F NMR?

LaL contains 3 methionines in its primary structure, located at positions 1, 14 and 107. With the availability of the efficient LaL expression system *E. coli* BL21/pLR102 and the methodology to prepare the novel fluorinated methionine analogue, the opportunity existed to examine these questions. Our investigations demonstrate that the approaches described in this Chapter may prove valuable for the specific study of the function of methionine residues in biological systems.

4.2. EXPERIMENTAL

4.2.1. General Experimental and Materials

Materials. Trifluoromethyl iodide (CF_3I) and N-acetyl-*D,L*-homocysteine thiolactone were purchased from Aldrich (Milwaukee, WI). Porcine kidney Acylase I (Grade I), *L*-amino acid oxidase (Type I, from *Crotalus adamateus* venom), horseradish peroxidase (Type II), cyanogen bromide (CNBr), peptide Met-Gly-Met-Met, 3,3'-dimethoxybenzidine dihydrochloride (*o*-dianisidine·2HCl), and the *L* isomers of methionine, serine, threonine and tyrosine were obtained from Sigma (Mississauga, ON). The methionine auxotroph *E. coli* B834(DE3) is a product of Novagen (Madison, WI). *E. coli* CM89 harboring pSYC1174 (#53245) was purchased from the American Type Culture Collection (Rockville, MD).

The preparation and characterization of plasmid pLR102 was described previously (section 2.2.5.5) and was used to transform *E. coli* B834(DE3) under standard conditions (section 2.2.4). General protein purification (2.2.9.1), analytical (3.2.2) and synthetic (3.2.6.1) techniques have been described in the sections indicated.

4.2.2. Synthesis of *L*-S-Trifluoromethylhomocysteine (*L*-TFM)

The procedure is a slight modification of that previously developed (Houston & Honek, 1989; Houston, 1992).

N-Acetyl-*D,L*-S-trifluoromethylhomocysteine Methyl Ester (16)

Sodium metal (1.80 g, 78.3 mmol) was added over a 10 min period in small portions to anhydrous methanol (155 mL) cooled on ice under argon. When all the sodium had dissolved, N-acetyl-*D,L*-homocysteine thiolactone (5.0 g, 31.4 mmol) was added at once, the ice bath was removed and the solution was stirred for 40 min. A dry ice/acetone condenser was attached to the flask and trifluoromethyl iodide (18.5 g, 94.4 mmol) was added to the solution over 15 min and the solution was irradiated by three long wave ultraviolet lamps. After 1 h, an additional amount of trifluoromethyl iodide (6.5 g, 33.2 mmol) was introduced into the flask over 5 min and the solution was irradiated for an additional 1 h. The solvent was removed *in vacuo* and the oily residue was taken up in CH_2Cl_2 (300 ml) and washed successively with 5% $NaHCO_3$ (100 ml) and brine (125 ml). The organic layer was dried over $MgSO_4$, filtered and the filtrate was evaporated *in vacuo*. The residue was chromatographed on silica gel using ethyl acetate (EtOAc)/hexanes (3:1) yielding 6.46 g (24.9 mmol, 79% yield) of a clear light amber oil.

R_f (silica TLC_2, EtOAc/hexanes 3:1;) 0.38; ^{1}H NMR (250 MHz, $CDCl_3$) δ 6.69 (d, 1H, $J_{NH,\alpha}$ = 7.8 Hz, NHAc), 4.64 (dt, 1H, $J_{\alpha,NH}$ = 7.8 Hz, $J_{\alpha,\beta}$ = 5.0 Hz, Hα), 3.69 (s, 3H, CO_2CH_3), 2.85 (t,

2H, $J_{\gamma,\beta}$ = 7.5 Hz, **H**γ × 2), 2.28-1.90 (m, 2H, **H**β × 2), 1.97 (s, 3H, NHCOC**H**$_3$); ^{13}C NMR (62.2 MHz, $CDCl_3$) δ 171.9, 170.5 (**N**H**C**OCH$_3$, **C**O$_2$CH$_3$), 130.8 (q, J_{CF} = 306.3 Hz, **C**F$_3$), 52.5 (**C**α), 51.0 (CO$_2$**C**H$_3$), 32.7 (**C**β), 25.9 (**C**γ), 22.8 (NHCOC**H**$_3$).

***L*-S-Trifluoromethylhomocysteine (17)**

N-Acetyl-*D,L*-S-trifluoromethylhomocysteine methyl ester **16** (6.46 g, 24.9 mmol) was dissolved in methanol (75 mL) and then Milli-Q water (75 mL) was added. The solution was cooled in an ice bath and powdered NaOH (1.10 g, 27.5 mmol) was added. The solution was stirred on ice for 20 min and additionally at room temperature for 1 h. The solvent was removed *in vacuo* and the resulting sodium salt of N-acetyl-*D,L*-S-trifluoromethylhomocysteine was dissolved in degassed MQW (150 mL) and the pH was brought to 7.4 with dilute HCl. Enantioselective hydrolysis of the N-acetyl-*L*-S-trifluoromethylhomocysteine was performed with slight modification to the method of Chenault et al. (1989). Acylase I (10 mg, 19900 U) was added and the solution was stirred gently under nitrogen at 25 °C for 8 h. Following 2 and 4 h, the solution was again degassed and the pH was adjusted to 7.4 with 0.1 N NaOH. The pH of the solution was brought to 5.0 with conc. HCl. Decolourizing charcoal was added and the solution was heated at 60 °C for 5 min and then filtered. The filtrate was adjusted to pH 1.4 with conc. HCl and extracted with EtOAc (160 mL). The aqueous layer was applied to a Dowex® 50W-X8 (H^+ form) column and eluted with MQW water (pH 5-6) until the effluent was at pH 5-6. The column was then eluted with 1 M NH_4OH and fractions containing the product (monitored by TLC_2) were pooled and applied to a Lobar reverse phase C-18 column (Lichroprep® RP-18, Merck). Elution was achieved sequentially with MQW, 10% and then 20% CH_3CN and fractions containing the product were pooled and lyophilized giving 2.11 g (10.4 mmol, 42% yield) of a white powder.

mp 227-229 °C (lit: 229-230 °C dec, Houston & Honek, 1989; 230 °C, Dannley & Taborsky, 1957); R_f (silica TLC_2, butanol-H_2O-acetic acid 4:1:1) 0.46; ^{1}H NMR (250 MHz, D_2O) δ 3.54 (t, 1H, $J_{\alpha,\beta}$ = 6.5 Hz, **H**α), 2.90 (t, 2H, $J_{\gamma,\beta}$ = 7.7 Hz, **H**γ × 2), 2.14-1.91 (m, 2H, **H**β × 2); ^{13}C NMR (62.2 MHz, D_2O) δ 172.2 (**C**O$_2$H), 131.1 (q, J_{CF} = 306.0 Hz, **C**F$_3$), 52.3 (**C**α), 30.8 (**C**β), 25.5 (**C**γ); ^{19}F NMR (376 MHz, 50 mM potassium phosphate pH 7.0, $CFCl_3$ external reference) δ -40.05 (C**F**$_3$); ESMS: exact mass calcd for $C_5H_8F_3SNO_2$ 203.0228, found 203.1.

4.2.3. Expression and Purification of TFM-LaL

Unless otherwise stated, cells were grown in M9 minimal medium (Appendix A) supplemented with 0.4% glucose and 40 mg/L ampicillin ($M9_{Glu,amp}$). It was determined previously (section 3.2.4.3.2) that a *L*-Met supplement of 0.2 mM sustained normal growth of the auxotroph *E. coli* B834(DE3). However, to reduce the residual amounts of *L*-Met in the medium during protein induction, a 0.1 mM *L*-Met supplement was chosen which thoroughly supported the exponential growth phase of the auxotroph (refer to Fig. 4.3 in section 4.3.2, p. 365).

E. coli B834(DE3) harboring pLR102 was grown at 37 °C in $M9_{Glu,amp}$ supplemented with 0.1 mM *L*-Met to an absorbance of 0.65 at 600 nm. Cells were centrifuged (9950 × g), washed with $M9_{Glu,amp}$ and then resuspended in one-half the original culture volume of $M9_{Glu,amp}$ containing 0.75 mM IPTG and either 1.0 mM *L*-TFM for high levels of incorporation or 1.0 mM *L*-TFM and 20 μM *L*-Met for low levels of incorporation and grown at 37 °C for 9.5 h.

Purification of the labelled lysozymes was performed essentially as described in section 2.2.9.4 with the following exceptions:

1) cells harvested following protein induction were frozen in liquid N_2, thawed and disrupted by sonication (refer to sections 2.2.9.1 and 2.2.9.5).
2) When purified from a volume of <1 L of labelling/induction medium, only single chromatographic steps over both Mono-S HR 10/10 and Phenyl-Superose HR 10/10 were performed.

The final fraction obtained from the Phenyl-Superose chromatography (i.e. the protein concentration step) was dialysed extensively and lyophilized as described in section 2.2.10.2. Protein yields of high level TFM-LaL were determined by A_{280} measurements using the ε_{280} value of 31712 M^{-1} cm^{-1}. Two purifications of high level TFM-LaL were performed and the respective yields (from 1 L of labelling/induction medium) were 1.4 and 1.7 mg. Low level TFM-LaL was quantitated by both A_{280} measurements and by mass following buffer salt correction and for the single purification performed, yielded 15.1 mg and 15.2 mg by the two methods respectively per 1 L of labelling/induction medium.

The level of incorporation of the respective TFM-labelled proteins was determined by ESMS analysis (examples are given in Appendix D) following their desalting by chromatography over the Delta Pak™ C_{18} column under the conditions given in section 4.2.4.

4.2.4. Optimization of Conditions Yielding Maximal TFM-LaL Expression and Desired Incorporation Levels

A culture of B834(DE3)/pLR102 was grown in $M9_{Glu,amp}$ supplemented with 0.1 mM *L*-Met, washed and resuspended in $M9_{Glu,amp}$ as described in section 4.2.3 and aliquoted into separate samples.

For High Levels of Incorporation. Sample aliquots were supplemented with *L*-TFM (0, 0.01, 0.1, 1.0, 2.0 or 5.0 mM) and 0.75 mM IPTG and incubated at 37 °C. At post-induction times of 4, 9.5 and 18 h, 0.5 mL aliquots of each were removed, centrifuged and the pellets taken up in 50 µL protein loading buffer. Samples were analyzed by SDS-PAGE and the intensity of the LaL band (at approx. 18 kDa) was visually compared to determine which conditions afforded maximal expression.

For Low Levels of Incorporation. Aliquot samples (typically 10 to 50 mL) were supplemented with 1.0 mM *L*-TFM, 0.75 mM IPTG and *L*-Met ranging in concentration from 0.1 nM to 1.0 mM in 10-fold increments and grown for 9.5 h at 37 °C. Following removal of a 0.5 mL aliquot for SDS-PAGE analysis, cells were collected, resuspended in 3-6 mL 50 mM KPB (pH 7.0) and disrupted by sonication. The lysates were filtered and applied to a S-Sepharose Fast Flow HR 5/5 column equilibrated with Buffer A (see below) using a flow rate of 0.5-0.7 mL/min and detection at 280 nm. After washing the column (0.5-0.7 mL/min) with Buffer A until elution of unadsorbed material was complete, the column was eluted with the linear gradient given below:

Time (min)	Flow Rate (mL/min)	% Buffer A	% Buffer B
0	0.5	100	0
2	0.5	95	5
22	0.5	75	25
22.1	1.5	75	25
25	1.5	0	100
32	1.5	0	100
35	1.5	100	0

Buffer A: 50 mM KPB, pH 7.0
Buffer B: 1.0 M (w/v) KCl in Buffer A

For each sample, the eluent between approximately 8-10 min and 15-18 min into the gradient was collected and lyophilized (NOTE: this retention time is expected to vary slightly depending on how well the column is packed with the resin).

The lyophilized samples were suspended in 100 μL MQW, filtered and further purified and desalted by chromatography over the 2.0 × 150 mm Delta Pak™ C_{18} column using the linear gradient given below and detection at 215 nm:

Time (min)	Flow Rate (mL/min)	% MQW†	% CH_3CN†
0	0.2	80	20
10	0.2	60	40
25	0.2	45	55
30	0.4	5	95
35	0.4	5	95
40	0.4	80	20

† containing 0.1% TFA

The peak eluting at approximately 18-20 min was collected baseline to baseline and an aliquot was subjected to ESMS analysis to determine the level of incorporation.

4.2.5. Effect of TFM on Cell Growth

An overnight culture (4 mL) of *E. coli* B834(DE3) harboring pLR102 was grown in $M9_{Gln,amp}$ supplemented with 0.5 mM *L*-Met. The cells were collected and washed with $M9_{Gln,amp}$ and resuspended in 2 ml of the same media. A 40 μL aliquot of the cell suspension was used to inoculate 4 mL samples of $M9_{Gln,amp}$ supplemented with *L*-Met, *L*-TFM or both as indicated in Table 4.1 (performed in triplicate) and cultures were incubated at 37 °C. Cell growth was monitored over time by following the absorbance at 600 nm for each culture. Following the initial 12 hr incubation period, cells from 0.5 mL of each sample culture were collected, washed and resuspended in 0.5 mL of $M9_{Gln}$ and 10 μL of each then subcultured into 1.0 mL $M9_{Gln}$ or $M9_{Gln,amp}$ each supplemented with 0.5 mM *L*-Met and grown at 37 °C. Cultures were examined with time and culture growth was assessed according to the turbidity of the culture.

4.2.6. ^{19}F NMR Spectroscopy

NMR data were obtained at 376.3 MHz on a Varian Unity spectrometer fitted with a 5 mm dual broadband probehead with the proton coil tuned to fluorine. Standard undecoupled spectral parameters were 27740 Hz spectral width, 32K data points zero filled to 64K, 10 μs pulse width, 0.577 s acquisition time with a 0.5 s relaxation delay. The first 6 points of the fids were corrected by linear prediction before fourier transformation

with a line broadening of 0.5 Hz. Unless otherwise stated, all spectra were recorded at 22 ± 0.5 °C. The ^{19}F spin-lattice relaxation times (T_1) were estimated by the inversion recovery method. Phase-sensitive 2D exchange spectra (EXSY) were accumulated (512 data points in t_2 for each of 64 t_1 values, 180 scans per increment) with mixing times from 0.1 to 0.9 sec (Jeener et al., 1979). Samples were measured using a coaxial insert containing CD_3OD (as the lock solvent) and $CFCl_3$ as an external frequency standard (referenced to 0.00 ppm). The insert was placed into a standard 5 mM NMR tube containing the sample. Buffer composition and TFM-LaL concentrations for NMR studies are given in figure and table legends. Chitopentaose ((GlcNAc)$_5$), used for sugar binding experiments, was introduced directly into samples in the NMR tube to the desired concentrations (with the assumption that the volume did not change), and the tube gently agitated to obtain complete and uniform dissolution. The error for the integrations of individual resonances was estimated by measuring the ratio of each resonance to the other three from the same spectrum and comparing these ratios to the corresponding ones from other spectra (four spectra each of high and low level TFM-LaL were compared). In this manner, the average error was estimated to be 9 ± 5% (a range of errors between 1.2-19.4% was observed for the ratios measured). The standard deviation in ^{19}F chemical shift measurements was ± 0.02 ppm based on 3 independent measurements of the same sample of high level TFM-LaL.

4.2.7. Methionine Aminopeptidase Activity on LaL

We were interested in ascertaining whether methionine aminopeptidase (MAP) could remove the N-terminal methionine residue in both *wt* LaL and TFM-labelled LaL. This could be useful in aiding in the partial assignment of the ^{19}F NMR signals for the TFM incorporated proteins. Initially, we obtained *E. coli* CM89 harboring pSYC1174 from the American Type Culture Collection (Rockville, MD). Plasmid pSYC1174 carries a fragment of the *E. coli* genome coding for the MAP gene (*map*) and the *map* promoter and high level expression of MAP was reported for this strain (Ben-Bassat et al., 1987) grown in supplemented E medium (Appendix A). We attempted to purify MAP following the published method (Ben-Bassat et al., 1987) but realized that in our hands, this strain did not produce the same high level expression of MAP as was reported. We were fortunate to have received the generous gift of purified MAP from Drs. Brian Matthews and Todd Lowther from the University of Oregon (Eugene, OR). Apparently, Dr. Matthews group has also had difficulties in the overexpression of MAP from the above ATCC strain.

i) MAP Activity Assay

The activity of MAP was determined essentially as described previously (Carter & Miller, 1984; Ben-Bassat et al., 1987). The protein solution to be assayed (10 μL) was added to a solution of peptide Met-Gly-Met-Met (4 mM) in 0.1 M KPB (pH 7.5), 0.2 mM $CoCl_2$ (90 μL), the reaction was incubated for 10-20 min at 30 °C and then was stopped by placing the tube in boiling water for 2 min. To the reaction was added the colour development solution which consisted of 0.2 mg *o*-dianisidine-2HCl, 0.2 mg *L*-amino acid oxidase, and 0.03 mg horseradish peroxidase in 0.1 M Tris, pH 7.4 (0.9 mL) and the reaction was incubated for 45 min at 30 °C. The absorbance at 440 nm was measured against appropriate enzyme and substrate blanks.

ii) Determination of MAP Activity on LaL

The preparation of MAP provided by Dr. B. Matthews was sent as a 2.43 mg/mL solution in 50 mM HEPES (pH 7.9), 150 mM KCl, 1 mM $CoCl_2$ (hereafter referred to as MAP solution). A LaL solution (5 mg/mL) was prepared in Buffer A (0.1 M KPB (pH 7.5), 0.2 mM $CoCl_2$). Two incubations were performed containing:

1) 16.6 μL MAP solution, 80 μL LaL solution and 63.4 μL Buffer A. This sample contained 0.25 and 2.5 mg/mL of MAP and LaL respectively.
2) 20 μL MAP solution, 20 μL LaL solution and 60 μL Buffer A. This sample contained 0.5 and 1 mg/mL of MAP and LaL respectively.

The two samples were incubated at 30 °C. Over the course of 72 hr, aliquots (10 μL) were removed and added to the colour development solution and assayed as described above. In addition, aliquots (5 μL for sample (1) or 15 μL for sample (2)) were chromatographed over the Delta Pak™ C_{18} column (as described in section 4.2.4) and were analyzed by ESMS to detect deletion of the N-terminal methionine. At the end of the incubations, MAP activity on Met-Gly-Met-Met was performed as described above which indicated that MAP activity was sustained over the course of the incubation.

4.2.8. Molecular Modelling and NMR Calculations

Gas phase AM1 (Dewar et al., 1985) and PM3 (Stewart, 1989) semi-empirical energy calculations and geometry optimizations were performed on ethyl methyl sulfide and ethyl trifluoromethyl sulfide using Polak-Ribiere conjugate gradient methods in Hyperchem Version 4.5 (Hypercube Inc., Waterloo, ON). Gas phase *ab initio* energy calculations and geometry optimizations were performed at the RHF/6-31G*//RHF/6-31G* and MP2/6-31G**//RHF/6-31G* levels using either Hyperchem Version 4.5 or

Gaussian 94W (RevC.3; Frisch et al., 1995). Frequency calculations at the RHF/6-31G* level were utilized to confirm the nature of the stationary points and to obtain zero-point vibrational energies (scaled by a factor of 0.89; Hehre et al., 1986). Calculated van der Waals and solvent-accessible surface areas and volumes were calculated under the QSAR option in ChemPlus (Hypercube Inc., Waterloo, ON) based on the Grid method (Grid = 50) described by Bodor and co-workers (Bodor et al., 1989) using the atomic radii suggested by Gavezotti (Gavezotti, 1983). Log P values were obtained using atomic parameters in ChemPlus based on the work of Viswanadhan and co-workers (Viswanadhan et al., 1989). Predicted gas phase fluorine ^{19}F isotropic magnetic shieldings in ppm were calculated for $CFCl_3$, $CF_3SCH_2CH_3$ (trans configuration), and the sulfoxide $CF_3S(O)CH_2CH_3$ at the RHF/6-311+G(d,p)//RHF/6-31G* level utilizing the GIAO method (Wolinski et al., 1990) as implemented in G94(RevC.3). The predicted ^{19}F chemical shifts were then determined for the various compounds by setting the calculated ^{19}F magnetic shielding for $CFCl_3$ equal to zero.

4.2.9. Cyanogen Bromide Reactions

4.2.9.1. Cyanogen Bromide Cleavage of LaL and TFM-LaL Proteins

a) Cleavage of LaL and High Level TFM-LaL

Solutions (1 mg/mL) of *wt* LaL and high level TFM-LaL were prepared utilizing their A_{280} in MQW and 300 µL (300 µg) of each was lyophilized in a small glass vial. The vial was fitted with a cap into which a small hole was created to allow entry of the tip of a glass pipette (for delivery of Ar) and a syringe tip for removal of aliquots. The dry samples were resuspended in 289 µL of 70% (v/v) formic acid and Ar was bubbled through the solution for 2 min. To each was added 11 µL of a freshly prepared solution of cyanogen bromide (50 mg/mL) in 70% formic acid. The vials were further flushed with Ar (1 min), capped and sealed tightly, wrapped in aluminum foil and incubated at room temperature in darkness. Therefore, the final composition of each solution consisted of 1 mg/mL (56 µM) protein and 1.8 mg/mL (17 mM) cyanogen bromide (which represents a 100-molar excess for each of the three methionines in LaL) in 70% formic acid.

At various times (0, 2, 6, 24 and 48 hr) 60 µL (60 µg protein) of each incubation was removed and added to 700 µL of MQW, frozen in liquid N_2 and lyophilized (each of these manipulations were performed under darkness). The time 0 aliquots were removed immediately (approx. 5-10 s) following the addition of the cyanogen bromide. Dried samples were stored at -20 °C for no longer than 24 hr prior to HPLC analysis.

Reverse phase chromatography was performed using the 2.0 mm × 15 cm Delta-Pak™ C_{18} column. The column was equilibrated with 85% buffer A and 15% buffer B and the linear gradient given in Table 4.1 was used for the separation. The lyophilized samples were resuspended in 100 μL MQW, centrifuged and injected onto the column. The eluent was monitored at 215 nm and fractions were manually collected and the times of each fraction recorded. Fractions were subsequently stored at 4 °C prior to analysis by ESMS. The entire lyophilized sample (60 μg) for each time point was injected except for the 48 hr samples in which only half (30 μg) of the sample was loaded (the remainder of each was used in an attempted SDS-PAGE analysis of the digests).

Table 4.1. Gradient conditions for separation of CNBr generated fragments.

Time (min)	Flow Rate (mL/min)	% Buffer A	% Buffer B
0	0.2	85	15
10	0.2	75	25
20	0.2	70	30
60	0.2	60	40
65	0.2	55	45
70	0.2	5	95
71	0.4	5	95
80	0.4	5	95
82	0.4	85	15
90	0.4	85	15

Buffer A: 0.1% TFA in MQW
Buffer B: 0.1% TFA in CH_3CN

b) Cleavage of Low Level TFM-LaL

Preparation of low level TFM-LaL for CNBr cleavage and HPLC analysis was performed as described above except that 0.9 mg of protein was dissolved into 893 μL 70% formic acid (v/v) and then 7.5 μL of freshly prepared CNBr in 70% formic acid (212 mg/mL) was added. The final protein concentration was again 56 μM with a 100-molar excess of CNBr for each methionine. Aliquots of the incubation were removed after 24 h (90 μL) and 48 h (250 μL), diluted 1:10 with MQW, lyophilized and subjected to RP-HPLC as described above. Fractions were collected only for the 48 hr time point.

4.2.9.2. Synthesis of N-Ac-Met and N-Ac-TFM

The acetylation of Met and TFM was performed essentially as described previously for the general acetylation of the α-amino group of peptides (Dell, 1986). The amino acid (0.5 mmol; 75 mg *L*-Met or 102 mg *L*-TFM) was added to MQW (2 mL) containing $NaHCO_3$ (0.55 mmol, 43 mg). The suspension was stirred for 1 min after which a solution (8 mL) of acetic anhydride:MeOH (1:3 v/v) was added, the mixture was stirred for 5 min and the solvent was removed *in vacuo*.

The *N*-acetyl derivatives were purified by chromatography on a μBondapak™ C18 Prep-Pak® cartridge (with the Guard-Pak™ cartridge) using the linear gradients given below for each:

(i) *N*-Ac-*L*-Methionine

Time (min)	Flow Rate (mL/min)	% Buffer A	% Buffer B
0	5.0	100	0
10	5.0	100	0
35	5.0	75	25
40	5.0	5	95
40.01	6.0	5	95
55	6.0	5	95
60	6.0	100	0

(ii) *N*-Ac-*L*-Trifluoromethionine

Time (min)	Flow Rate (mL/min)	% Buffer A	% Buffer B
0	5.0	95	5
10	5.0	95	5
50	5.0	55	45
55	5.0	5	95
55.01	6.0	5	95
65	6.0	5	95
70	6.0	95	5

Buffer A: 0.1% TFA in MQW
Buffer B: 0.1% TFA in CH_3CN

The crude product from the reaction of *L*-Met was dissolved in 2.0 mL of Buffer A and centrifuged. Two individual runs were performed (injecting 1.0 mL for each run) employing gradient (i) and the eluent (monitored at 215 nm) was collected from 23-29 min for each run and lyophilized to dryness yielding 41 mg (43%) of *N*-acetyl-*L*-Met as a clear oil. ^{1}H NMR (250 MHz, CD_3CN with 1 drop $CHCl_3$ as reference (7.23 ppm)) δ 6.59 (1H, d, $J_{NH,\alpha}$ = 7.1 Hz, NHAc), 4.16 (1H, m, Hα), 2.18-2.25 (2H, m, Hγ × 2), 1.60-1.84 (2H, m, Hβ × 2), 1.76 (3H, s, SCH_3) 1.64 (s, 3H, $NHCOCH_3$); ESMS: exact mass calcd for $C_7H_{13}NSO_3$ 191.0616, found 190.91.

The crude product from the reaction of *L*-TFM was suspended in 4.0 mL of MeOH:Buffer A (1:4 v/v) and centrifuged. Two individual runs were performed (injecting 2.0 mL for each run) employing gradient (ii) and the eluent (monitored at 215 nm) was collected from 26-34 min for each run and lyophilized to dryness yielding 48 mg (39%) of *N*-acetyl-*L*-TFM as a clear oil. ^{1}H NMR (250 MHz, CD_3CN with 1 drop $CHCl_3$ as reference

(7.23 ppm)) δ 6.70 (1H, d, $J_{NH,\alpha}$ = 7.4 Hz, NHAc), 4.20 (1H, m, Hα), 2.61-2.74 (2H, m, Hγ × 2), 1.62-1.99 (2H, m, Hβ × 2), 1.67 (s, 3H, $NHCOCH_3$); ESMS: exact mass calcd for $C_7H_{10}NSO_3F_3$ 245.0333, found 244.98.

The other major product (approx. 40-50%) generated under the reaction conditions used to *N*-acetylate *L*-Met and *L*-TFM were the methyl esters of the respective *N*-acetylated amino acids. These could also be isolated and eluted under the chromatographic conditions using the respective gradients between 32-35 min (*N*-Ac-*L*-Met methyl ester) and 37-41.5 min (*N*-Ac-*L*-TFM methyl ester) and were identified by ESMS and ^1H NMR characterization.

4.2.9.3. Treatment of N-Ac-Met and N-Ac-TFM with CNBr

Solutions of *N*-Ac-*L*-Met and *N*-Ac-*L*-TFM (105 mM) were prepared in CH_3CN and 0.38 mL of each solution was removed and the solvent removed *in vacuo*. The amino acids were then dissolved in 70% formic acid (9.05 mL), degassed with Ar (2 min) and then split into 2 × 4.52 mL aliquots. To one aliquot was added an additional 0.48 mL of 70% formic acid and to the other was added 0.48 mL of a 944 mM (100 mg/mL) solution of CNBr in 70% formic acid. Therefore each solution contained 4 mM of the amino acid in the absence or presence of 90 mM CNBr. A reagent control sample containing only 90 mM CNBr in 70% formic acid was also prepared under identical conditions.

The solutions were stirred under Ar in darkness and at times of 1, 24 and 48 h, 1.0 mL aliqouts of each were removed, diluted 1:10 with MQW and lyophilized to dryness in the dark. The dry samples were resuspended in 250 μL of Buffer A, centrifuged and 20 μL of each chromatographed over the 2.0 mm × 15 cm Delta-Pak™ C_{18} column. The column was equilibrated with 100% Buffer A and eluted with the linear gradient given in Table 4.2.

Table 4.2. Gradient conditions for analysis of the reaction products of CNBr with Ac-Met and Ac-TFM.

Time (min)	Flow Rate (mL/min)	% Buffer A	% Buffer B
0	0.2	100	0
5	0.2	100	0
35	0.2	40	60
37	0.4	5	95
43	0.4	5	95
45	0.4	100	0
55	0.4	100	0

Buffer A: 0.1% TFA in MQW
Buffer B: 0.1% TFA in CH_3CN

The eluent was monitored at 210 nm and fractions were collected manually and analyzed by ESMS and TLC (conditions are given in section 4.3.7.1).

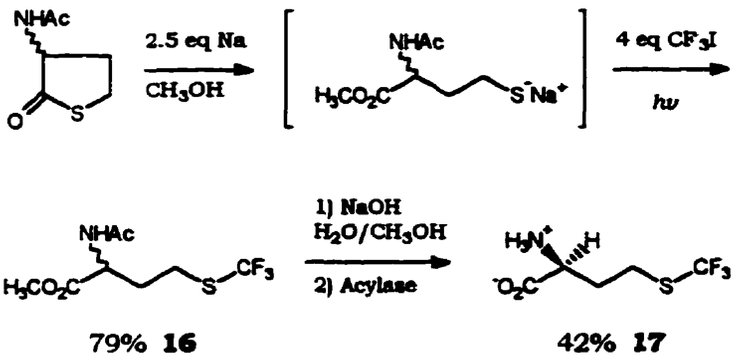
4.2.9.4. Treatment of Serine, Threonine and Tyrosine with CNBr

The free amino acids of Ser, Thr and Tyr (100 μmol) were dissolved into 500 μL of 70% formic acid (v/v). One half of each solution (250 μL) was added to 250 μL of 70% formic acid and to 250 μL of 70% formic acid containing CNBr (500 μmol). Therefore, the solutions contained 0.1 M of the amino acid with or without 1.0 M CNBr. The solutions were degassed with Ar and stored in the dark at 25 °C. Following a 24 h incubation, samples were diluted with MQW (6 mL) and lyophilized in darkness. A small portion (approx. 100 mg) of each dry sample was dissolved into 1 mL of MQW:CH_3CN (1:1, containing 0.1% TFA), centrifuged, and subjected to high resolution ESMS analysis.

4.3. RESULTS AND DISCUSSION

4.3.1. Synthesis of *L*-TFM

The synthesis of *D,L*-TFM was first reported by Dannley and Taborsky (1957) but suffered from low overall yield (11% in 5 steps) with the generation of a racemic product. The current method is based on our previous protocols (Houston & Honek, 1989; Houston, 1992). In the presence of methoxide, N-acetyl-*D,L*-homocysteine thiolactone is converted to the thiolate *in situ* which is reacted with excess trifluoromethyl iodide giving the fully protected racemic amino acid **16**. Alkylation of the thiolate proceeds via a radical mechanism and requires substantial long wave UV irradiation. Saponification of the methyl ester followed by resolution and deprotection with acylase generated the desired *L* isomer of S-trifluoromethylhomocysteine **17** in an overall yield of 33% (Scheme 4.1).



Scheme 4.1.

4.3.2. Optimal Conditions to Produce TFM-Labelled LaL

The analogue *L*-TFM will not support growth of *E. coli* B834(DE3)/pLR102 (discussed in section 4.3.3) and therefore, it was necessary to initially grow the auxotroph in medium supplemented with methionine. As is to be expected, the growth of the auxotroph shows dependency on the concentration of *L*-Met present in the medium (Fig. 4.3). The generation time of the auxotroph B834 (≈ 55 min) was again, less than that for the parent non-auxotrophic strain BL21(DE3) (≈80 min), the results being very similar to those described previously for the growth of the methionine auxotroph *E. coli* DH93 (refer to Fig. 3.20).

Therefore, *E. coli* B834(DE3)/pLR102 was initially grown in minimal medium containing 0.1 mM *L*-methionine to mid-exponential growth (i.e. $OD_{600} \approx 0.6$-0.7) at which stage, the cells were collected and washed to remove residual methionine from the

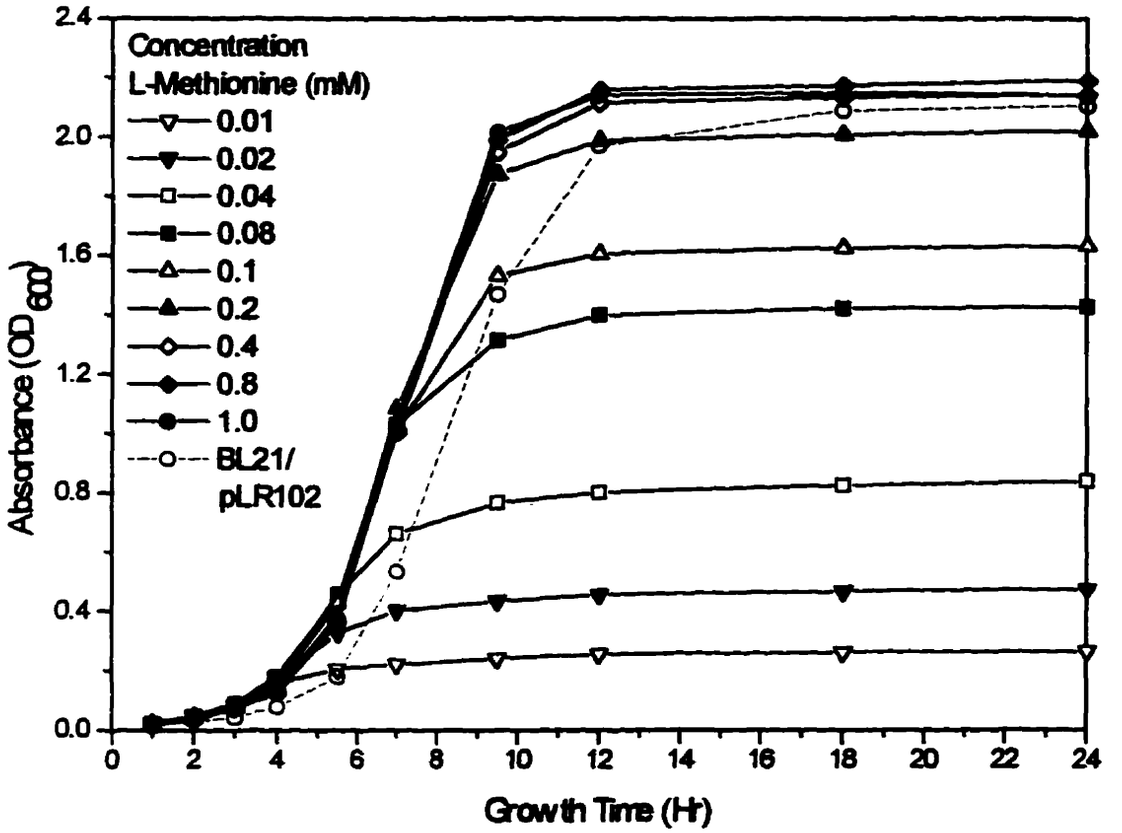


Figure 4.3. Dependence on methionine supplementation for the growth of *E. coli* B834 harboring pLR102. Cultures were grown in M9 medium at 37 °C supplemented with glucose, ampicillin and various concentrations of *L*-Met as indicated by the symbols (in mM). Also shown is the growth curve of *E. coli* BL21/pLR102 (broken line) grown in M9 medium supplemented with glucose.

medium. We wished to limit the amount of *L*-Met carried over from the initial growth. Since the dependency during exponential growth (at least to reach a cell density corresponding to an $OD_{600} \approx 1$) is not adversely affected at the supplement concentrations of 0.1 mM or higher (Fig. 4.3), we chose the lowest of these concentrations, or 0.1 mM, as the supplement concentration. Following the wash procedure, the cells were then reintroduced into one-half the original culture volume of fresh medium supplemented solely with *L*-TFM or in combination with *L*-Met and induced with IPTG. By reintroducing the cells into only one-half the original volume of medium for protein induction and labelling, this not only raised levels of protein production, to be expected, per volume of culture (essentially a two-fold increase as judged from SDS-PAGE analysis) but also reduced the quantity of *L*-TFM needed in half when compared to that amount needed if induction and labelling was brought about in the original volume of culture.

4.3.2.1. High Level TFM-LaL

To obtain high incorporation levels of TFM into LaL (i.e. high level TFM-LaL), cultures were induced in the presence of 1.0 mM of the fluorinated analogue for 9.5 h. Care was taken during the wash procedure to ensure thorough removal of methionine from the media. As will be discussed below, induction in the presence of methionine at concentrations as low as 10 μM will considerably reduce TFM incorporation. LaL expression levels were analyzed for media supplemented with various concentrations of *L*-TFM ranging from 0.01 to 5 mM and duration of induction period. Maximum protein expression was obtained at 1.0 and 2.0 mM *L*-TFM, and since the analogue appears to be toxic at higher concentrations, 1.0 mM was chosen as the supplement concentration. Interestingly, *L*-TFM at 5.0 mM resulted in lower levels of protein expression most likely due to the bactericidal effect of the analogue at this concentration (discussed in 4.3.3). Protein synthesis, in general, was found to increase with time (18 hr > 9.5 hr > 4 hr) in the presence of the inducer, IPTG. However, this phenomenon was found to occur even when no methionine supplement was included during protein induction (compare lane 1 Figure 4.4 A (no methionine and 4 hr induction) to lane 2 Figure 4.4 B (no methionine and 9.5 hr induction). These results suggest that during prolonged growth, the auxotroph is able to recycle endogenous sources of methionine presumably from protein degradation. This new supply of methionine would compete with *L*-TFM to reduce the relative incorporation of the analogue into each of the methionine positions (positions 1, 14 and 107) in LaL. Therefore, even though elevated levels of protein synthesis resulted if induction proceeded for 18 hr, an induction period of 9.5 hrs was chosen for further study.

We define *mono*-, *bis*-, and *tris*-labelled LaL species as enzyme in which 1, 2 or 3 of the 3 methionine positions have been replaced with TFM respectively. Shown in Fig 4.5 (A) is the ESMS spectrum of TFM-LaL prepared under high incorporation conditions. The spectrum clearly demonstrates that TFM has been incorporated at each of the locations normally occupied by methionine in LaL. In addition, partial sequencing of the N-terminus of high level TFM-LaL (results not shown) has indeed established that Met1 and Met14 have been substituted by TFM.

With the assumption that the ionization of each species is the same, the mass spectrum can be interpreted quantitatively. Based on the areas under the peaks, the relative amounts of *tris*:*bis*:*mono*:*wt* are approximately 1:1:0.5:0.1. Statistically, there are three different *mono*- and three different *bis*-labelled species and only one tris-labelled species possible. Considering that the individual species present represent 33% (*mono*), 66% (*bis*) and 100% (*tris*) incorporation respectively and given the relative quantities of

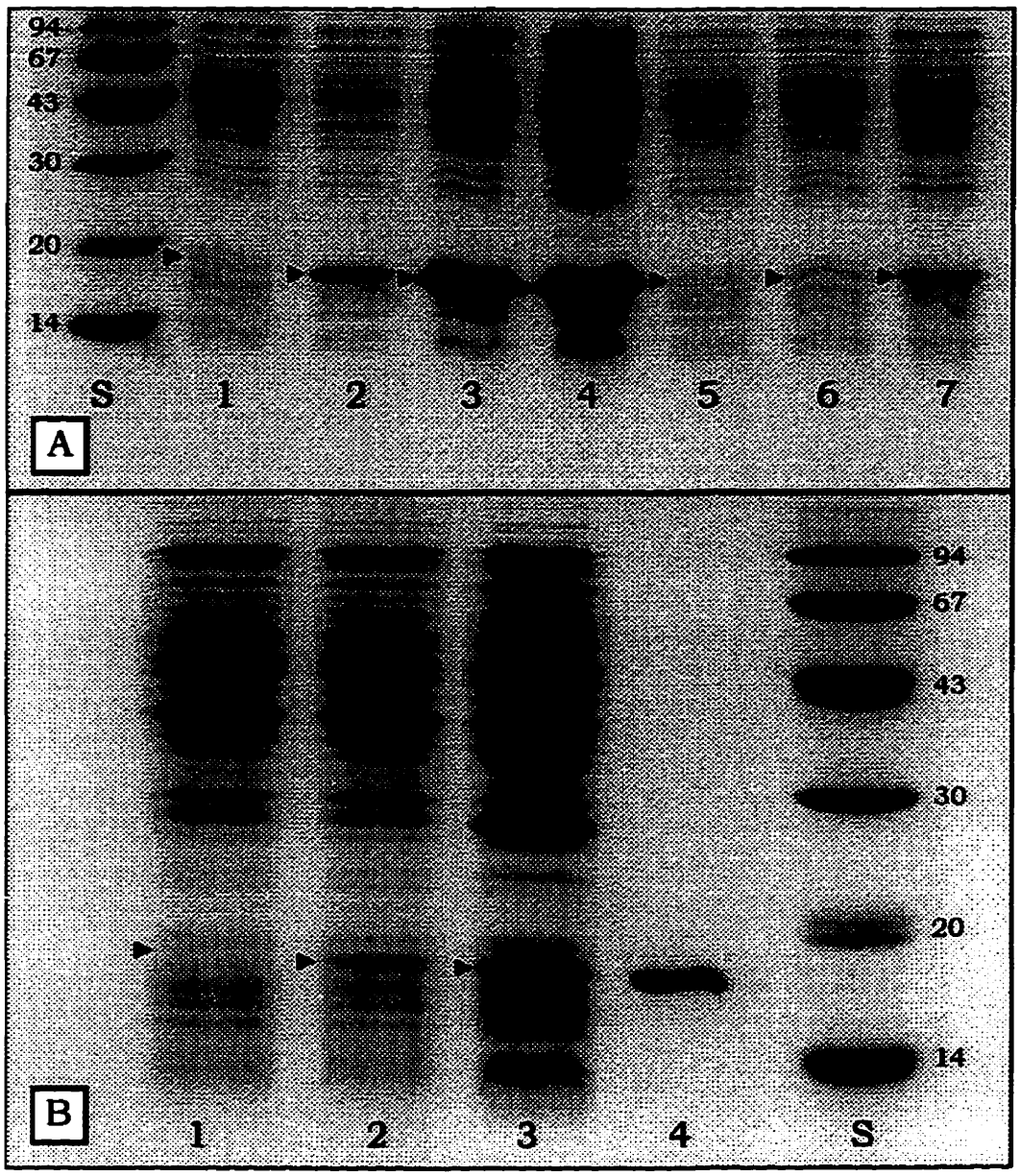


Figure 4.4. Comparison of the expression efficiency of LaL and TFM-LaL under various conditions. *E. coli* B834 harboring pLR102 was grown in $M9_{Glu,amp}$ (containing 0.1 mM *L*-Met) to an absorbance of 0.65 at 600 nm, cells were collected, washed and reintroduced into one-half the initial growth volume of $M9_{Glu,amp}$. Cultures were then treated as described in the following text and aliquots were subjected to 15% SDS-polyacrylamide gel electrophoresis analysis and Coomassie staining.

(**A**) Cells subjected to a 4 hr induction period in the presence of 0.01 mM, 0.1 mM or 1.0 mM of *L*-Met (Lanes 2, 3, 4) or of *L*-TFM (Lanes 5, 6, 7). Lane 1 represents cells induced with no methionine supplement.

(**B**) Cells subjected to 9.5 hr of growth without induction or methionine supplement (Lane 1), with induction and no methionine supplement (Lane 2) or in the presence of 1.0 mM TFM (Lane 3). Lane 4 is purified high level TFM-LaL.

Arrow heads indicate the position of LaL. Lanes S are molecular weight standards (in kilodaltons). The figure is best viewed at arm's length .

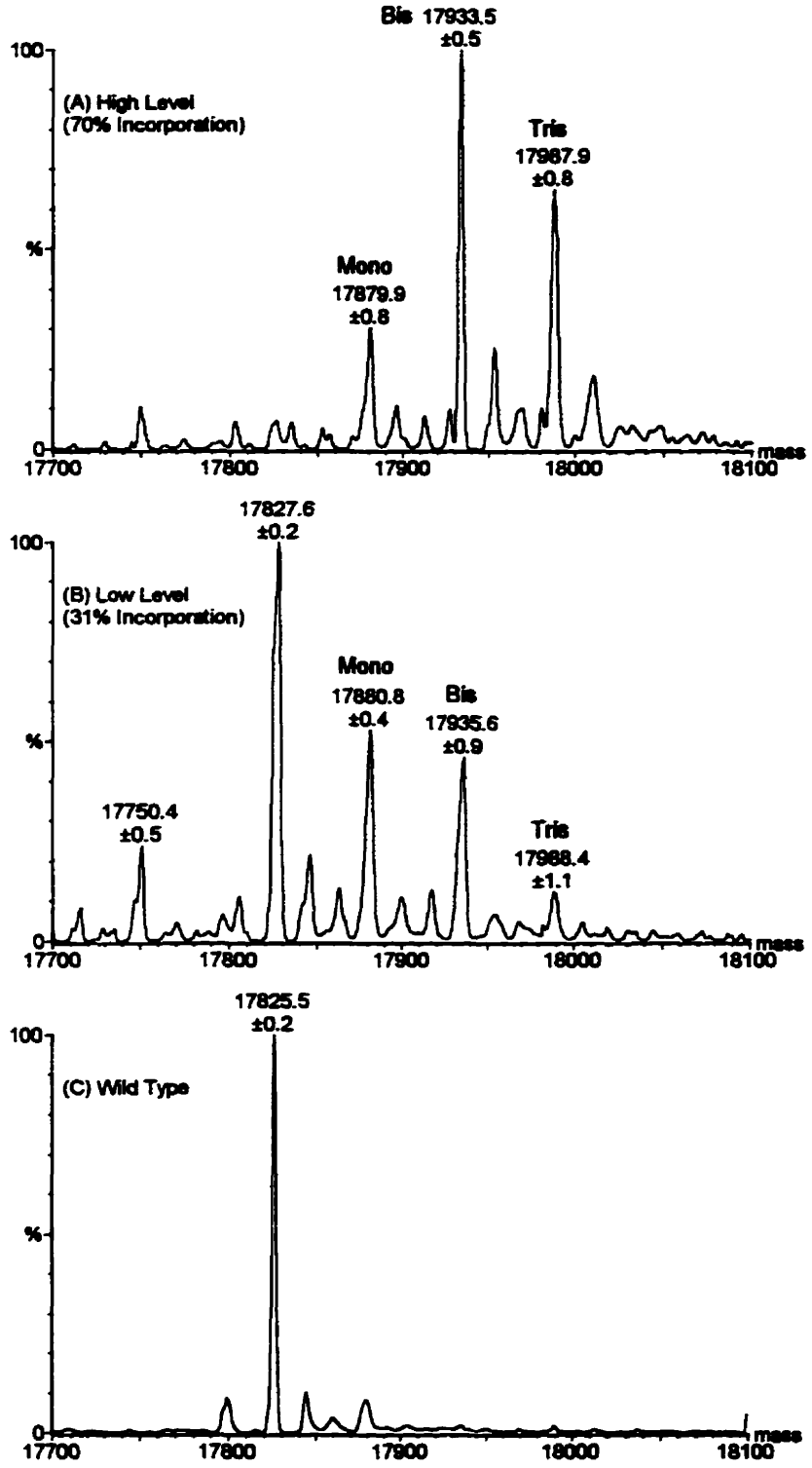


Figure 4.5. Reconstructed electrospray mass spectra of high level TFM-LaL (A), low level TFM-LaL (B) and wild type LaL (C) indicating observed molecular weights (Da). The labels *mono*, *bis* and *tris* refer respectively to lysozyme labelled with one, two or three TFM residues. The calculated molecular weights (Da) are: wild type, 17825.2; *mono*-labelled, 17879.2; *bis*-labelled, 17933.2; *tris*-labelled, 17987.1.

each, the overall extent of incorporation can be calculated to be approximately 70% (a more complete account of these calculations is given in Appendix D). In another prepartion of high level TFM-LaL, incorporation was found to be 74% by the same analysis. The respective protein yields for the 70 and 74% incorporated TFM-LaL were 1.7 and 1.4 mg per litre of labelling/induction medium respectively and high level TFM-LaL could be purified to apparent homogeneity (see lane 4, Fig. 4.4 (B)) by the standard protocol used to purifiy *wt* LaL. Therefore, reproducible levels of incorporation and protein yields can be achieved under the high labelling conditions.

4.3.2.2. Low Level TFM-LaL

In order to establish conditions that would generate lower levels of incorporation, samples for ESMS analysis were prepared by varying the concentration of *L*-methionine during induction from 0.1 nM to 1 mM (each containing 1.0 mM *L*-TFM). Below 10 uM *L*-methionine, the ESMS spectra of TFM-LaL essentially resembled that shown in Fig 4.5 (A). Increasing methionine concentrations led to a proportional decrease of *tris*-labelled LaL with concomitant increase in both *wt* and *mono*-labelled enzyme. A marked difference in the spectra was observed between 10 and 100 uM *L*-methionine and therefore 20, 40, 60, and 80 uM concentrations were also investigated. *L*-Methionine supplements above and including 60 uM resulted in no appreciable levels of species containing TFM as the spectra were very similar to that of the wild type enzyme (Fig. 4.5 (C)). *L*-Methionine supplements of 10, 20 and 40 uM resulted in comparable detectable labelled species and 20 uM was chosen for further investigation. As seen in Fig. 4.5 (B), TFM-LaL prepared under low level incorporation conditions generated each of *tris*-, *bis*- and *mono*-labelled species with *wt* predominating in relative amounts of approximately 0.15:0.5:0.6:1 respectively. As described above, the extent of incorporation is calculated to be approximately 31% and 15 mg/L of purified protein is obtained. The precision of calculating incorporation levels by ESMS was assessed by a triplicate determination of low level TFM-LaL, which gave 30.5 ± 0.7% (see Appendix D for greater details).

4.3.3. TFM Toxicity and Metabolism

In our efforts to optimize conditions for the incorporation of *L*-TFM into proteins, we were also interested in the effects of the analogue on cell growth in the absence of IPTG induction. As is expected, *L*-methionine is required to permit growth of the *E. coli* auxotroph B834(DE3) harboring pLR102, with the total growth increasing with *L*-Met concentration as a function of time (Table 4.3 and refer to Fig. 4.3). Clearly, *L*-TFM will

Table 4.3. Effect of various concentrations of *L*-met or *L*-TFM alone or in concert on the growth of *E. coli* B834 (DE3) harboring pLR102.

	Absorbance (600 nm)[a]													
***L*-Met (mM)**	0	0.01	0.05	0.5	0				0.01	0.05	0.5	0.01	0.05	0.5
***L*-TFM (mM)**	0	0			0.01	0.1	1.0	5.0	1.0			5.0		
Growth Time (hr) 3	0.041 ± 0.002	0.208 ± 0.003	0.365 ± 0.004	0.365 ± 0.003	0.046 ± 0.001	0.046 ± 0.001	0.057 ± 0.001	0.049 ± 0.000	0.217 ± 0.003	0.360 ± 0.004	0.362 ± 0.006	0.052 ± 0.004	0.067 ± 0.002	0.135 ± 0.003
6	0.047 ± 0.002	0.225 ± 0.002	0.869 ± 0.001	1.234 ± 0.018	0.051 ± 0.001	0.051 ± 0.001	0.067 ± 0.001	0.038 ± 0.001	0.228 ± 0.003	0.899 ± 0.014	1.333 ± 0.032	0.032 ± 0.002	0.042 ± 0.002	0.484 ± 0.013
9	0.052 ± 0.001	0.238 ± 0.001	0.993 ± 0.002	1.771 ± 0.010	0.053 ± 0.001	0.055 ± 0.001	0.069 ± 0.002	0.031 ± 0.001	0.212 ± 0.004	1.023 ± 0.029	1.788 ± 0.035	0.022 ± 0.001	0.022 ± 0.002	1.431 ± 0.044
12	0.055 ± 0.001	0.249 ± 0.001	1.070 ± 0.040	1.926 ± 0.017	0.056 ± 0.001	0.059 ± 0.001	0.067 ± 0.001	0.027 ± 0.002	0.188 ± 0.004	1.116 ± 0.024	1.933 ± 0.020	0.020 ± 0.001	0.016 ± 0.003	1.930 ± 0.014
Growth Time (hr)	**Cell Viability: Relative Cell Growth[b]**													
16	*10*	*10*	*10*	*10*	*10*	*10*	*0*	*0*	*10*	*10*	*10*	*0*	*2*	*10*
16	10	10	10	10	10	10	1	0	10	10	10	0	2	10
30							*10*	*0[c]*				*10*	*10*	
30							10	0[c]				10	10	

Top pannel. **An overnight culture of pLR102/B834(DE3) was grown in $M9_{Glu,amp}$ supplemented with 0.5 mM *L*-met. The cells were collected, washed and resuspended in $M9_{Glu,amp}$ and used to innoculate aliquots of $M9_{Glu,amp}$ supplemented with *L*-met, *L*-TFM or both as indicated (performed in triplicate) and samples incubated at 37 °C. The absorbance at 600 nm was recorded at the indicated times. [a] Values represent the average ± S.D.**

Bottom pannel. **Following the initial 12 hr growth, cells from each sample were washed and diluted 1:100 into $M9_{Glu}$ or $M9_{Glu,amp}$ (italicised growth time) each supplemented with 0.5 mM *L*-met and grown at 37 °C. At given times, cultures were observed and cell growth assessed according to the turbidity of the culture. [b] Based on a scale of 0 to 10, 0 representing a clear culture (no observable growth) and 10 representing a heavily turbid culture. [c] No growth observed following 48 hr incubation.**

not support cell growth (Table 4.3). At concentrations examined of 1.0 mM and below, culture growth is minimal and similar to the control where no methionine supplement was included. *L*-TFM at 5.0 mM did appear to have an adverse affect on growth when compared to the control (Table 4.3). Cell growth was also influenced in those cultures supplemented with *L*-TFM together with *L*-Met. At high ratios of *L*-TFM/*L*-Met (5.0 mM/0.01 and 0.05 mM) cell growth is effectively inhibited. An adverse effect on growth is also observed in cultures containing concentrations of *L*-TFM/*L*-Met of 1.0 mM/0.01 mM and 5.0 mM/0.5 mM when compared to those grown with *L*-Met alone at these levels. Cell growth in cultures containing 1.0 mM *L*-TFM and 0.05 or 0.5 mM *L*-Met was not affected. It is apparent that *L*-TFM does impart a dose dependent toxicity to cell growth when the concentration of the analogue is substantially greater than that of *L*-Met. This inhibition may be a consequence of the analogue interfering with the normal metabolic utilization of methionine or with other cellular processes in some manner.

Not only are high concentrations of *L*-TFM detrimental to cell growth, but *L*-TFM also appears to affect the viability of cells. As indicated in Table 4.3, cells grown for 12 hr in the presence of 1.0 mM *L*-TFM alone and 5.0 mM *L*-TFM with 0.01 and 0.05 mM *L*-Met prior to subculturing into fresh media containing 0.5 mM *L*-Met showed a substantial lag period (>16 hr) before reversal of growth inhibition was observed. This lag period extended to greater than 48 hr for cells initially grown in the presence of only 5.0 mM *L*-TFM. No difference in cell viability was noted when cells were subcultured into media with or without ampicillin, suggesting that lack of growth in the former media can not be attributed to plasmid instability. In the case of *Saccharomyces cerevisiae*, 10 μM *L*-TFM produced an indefinite lag period which could be reversed by methionine or S-adenosylmethionine (Colombani et al., 1975).

The growth inhibitory effect of *L*-TFM was initially reported by Zygmunt and Tavormina (1966) for a variety of microorganisms. These workers observed that 100 μg/ml (0.49 mM) of *D,L*-TFM failed to inhibit the growth of *E. coli* in an enriched medium; however, substantial inhibition resulted using a minimal medium. In the latter instance, it appears that *de novo* biosynthesis of methionine or its intracellular utilization may be influenced by TFM and if methionine is present in the medium, this effect is diminished. These results support our observations of growth inhibition with the *E. coli* methionine auxotroph B834(DE3)/pLR102. Interestingly, 1 mM *D,L*-TFM in the presence of 17 μM *D,L*-Met did not inhibit the growth of the methionine-dependent bacterium *Leuconostoc mesenteroides* (Zygmunt & Tavormina, 1966). In *S. cerevisiae*, TFM has been shown to cause a dramatic decrease in the intracellular free pool of methionine which correlated

with inhibition of growth presumably due to its function as a repressor of the biosynthesis of this amino acid (Colombani et al., 1975). This may also be the case in *E. coli*, however we have not fully explored the mechanism of growth inhibition as we were primarily interested in the incorporation of *L*-TFM into proteins.

The incorporation of TFM into LaL also provides some insight into the enzymes involved in protein biosynthesis. The biosynthesis of all proteins begins with methionine or in the case of eubacteria and in eukaryotic organelles such as mitochondria and chloroplasts, with formylmethionine (fMet). Methionyl-tRNA synthetase is responsible for the acylation of both the elongator $tRNA^{Met}$, for insertion of Met into internal peptide linkages, and of the initiator $tRNA_i^{fMet}$. The crystal structure of *E. coli* methionyl-tRNA synthetase has been determined (Brunie et al., 1990) and methionine has been modeled into the active site (Kim et al., 1993). In the proposed model, the methyl group of Met is positioned between the side chains of Phe197 and Trp305. The interactions between methionyl-tRNA synthetase and TFM must therefore be of sufficient nature to accommodate the trifluoromethyl group and permit the analogue to proceed through the synthetic pathway to form TFM-$tRNA^{Met}$ and TFM-$tRNA_i^{fMet}$.

Eubacterial translation initiation also involves formylation of Met-$tRNA_i^{fMet}$ to fMet-$tRNA_i^{fMet}$ by a N^{10}-formyl-tetrahydrofolate-dependent transformylase (RajBhandary, 1994). Whether TFM-$tRNA_i^{fMet}$ becomes formylated cannot be stated with certainty since eubacteria compromised in the biosynthesis of the formyl donor or of the transformylase can sustain diminished growth using unformylated Met-$tRNA_i^{fMet}$ as the initiator (Harvey, 1973; Baumstark et al., 1977; Guillon et al., 1992). It is likely however that TFM-$tRNA^{fMet}$ is formylated. The $tRNA^{fMet}$ can be charged with the methionine analogues norleucine and ethionine and subsequently formylated (Trupin et al., 1966) suggesting that the formylase enzyme is somewhat flexible towards its substrates. In addition, it has been determined that the amino acid acceptor stem of the *E. coli* Met-$tRNA_i^{fMet}$ is quintessential for recognition by the transformylase (Lee et al., 1991). Following translation of fMet-$tRNA^{fMet}$, the formyl group is removed by a peptide deformylase, whose substrate requirements were recently investigated (Meinnel & Blanquet, 1995). If indeed formylation of TFM-$tRNA_i^{fMet}$ does occur, the ensuing deformylation must also proceed since ESMS analysis of TFM-LaL did not give any evidence of formylation of the mature proteins. Finally, results from integration of the resonances obtained from the ^{19}F NMR studies of TFM-LaL suggest that all three methionine locations incorporate *L*-TFM equally well (presented in section 4.3.5)

The characterization of TFM-LaL clearly illustrates the ability of the biosynthetic system to recognize and incorporate *L*-TFM into methionine locations. However, it is also evident that protein expression is substantially reduced when *L*-TFM serves as the methionine source. It is clearly seen (Fig. 4.4 (A)) that the amount of LaL produced is considerably less when induction occurs in the presence of *L*-TFM (Lanes 5, 6, 7) instead of *L*-Met (Lanes 2, 3, 4). This is also reflected in the amount of purified *wt* LaL obtained under *L*-Met enriched conditions (25-30 mg/L) as compared to that obtained for high level TFM-LaL (1-2 mg/L) and low level TFM-LaL (15 mg/L). The decreased expression may be the result of inefficient utilization of *L*-TFM in the translation process. Also, there may be a reduced intracellular availability of *L*-TFM resulting from deficiencies in the transport of the analogue into the cell. In *E. coli*, there exists both a high and a low affinity methionine transport system, the former being ATP-dependent (Kadner, 1977). The high affinity transporter, coded for by the *metD* locus, has broader substrate specificity and is likely the transporter utilized by *L*-TFM. Reduced affinity of the transport system for the fluorinated analogue may result in reduced production of TFM-LaL. It has also been previously shown that *E. coli* cellular viability is dependent on S-adenosyl-*L*-methionine availability, the product of *L*-methionine and ATP condensation catalyzed by S-adenosyl-*L*-methionine synthetase (Satishchandran et al., 1990). *L*-TFM has been previously shown to competitively inhibit the *S. cerevisiae* S-adenosyl-*L*-methionine synthetase (Colombani et al., 1975; Chou & Talalay, 1972). The lack of formation of S-adenosyl-*L*-methionine in the presence of *L*-TFM in *E. coli* may also lower cellular activity and hence TFM-LaL production. This may also explain the impaired growth of the methionine auxotroph when grown in the presence of *L*-TFM (Table 4.3).

4.3.4. Effect of *L*-TFM Incorporation on Activity and Propensity to Inhibition by Chitooligosaccharides

The activities and inhibition of the TFM-labelled lysozymes were assessed turbidimetrically using *E. coli* substrate cells as detailed previously in Chapter 3. Incorporation of *L*-TFM into LaL generated catalytically active proteins that were inhibited by chitooligosaccharides (Fig. 4.6). Preparations of high and low level TFM-LaL exhibited relative activities comparable to *wt* lysozyme indicating that *L*-TFM incorporation did not result in any substantial structural alterations deleterious to enzyme activity.

Compared to *wt* LaL, both high and low level TFM-LaL were inhibited equally well by $(GlcNAc)_3$ but appeared to show less inhibition by $(GlcNAc)_{4-6}$ (Fig. 4.6). In Chapter 3, it was suggested that the different chitosaccharides may assume different modes on binding to LaL. If such is the case, then it is reasonable to assume that length of the $(GlcNAc)_n$

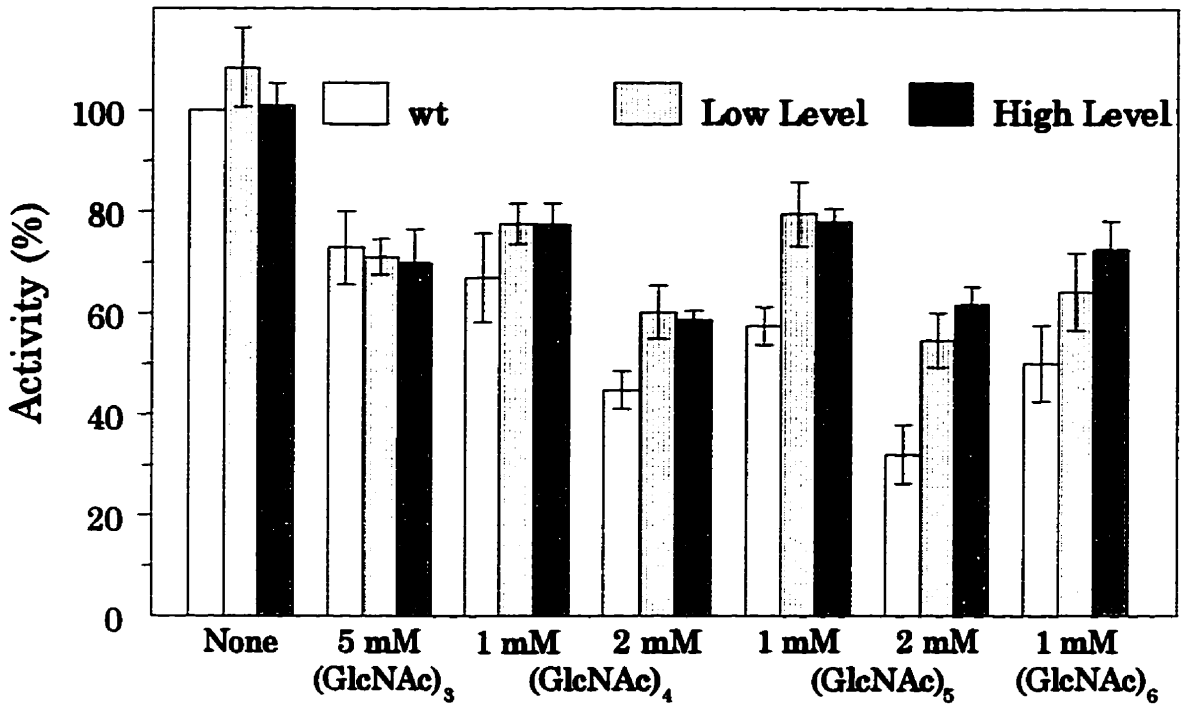


Figure 4.6. Comparison of the activity and inhibition of *wt* and TFM-LaL. Illustrated are the activities for *wt* LaL, low level and high level TFM-LaL in the absence (None) or presence of chitooligosaccharides ($(GlcNAc)_n$) at the indicated concentrations.

inhibitor dictates which particular sequential subsites of the protein are predominantly recruited for binding. $(GlcNAc)_3$ may bind in such a manner that none of the three methionine residues in LaL are in the vicinity of its preferred binding site and hence it interacts equivalently with *wt* and TFM-LaL. Conversely, a methionine residue(s) may be present or involved in the association with $(GlcNAc)_{4-6}$. If so, the presence of a TFM residue may interfere with proper binding and a reduced affinity reflected in decreased inhibition would be observed. In addition, the labelled enzymes (i.e. high and low level TFM-LaL) generally exhibited comparable inhibition to each other although a slight difference in inhibition between the two levels may be suggested for $(GlcNAc)_5$ at 2 mM and $(GlcNAc)_6$ (Fig. 4.6).

There is literature support for the involvement of methionine in carbohydrate binding. It was previously discussed (see 4.1.1) that Met108 of the arabinose-binding protein adopts specific conformations to optimize nonpolar interactions in association with its different ligands (Vermersch et al., 1991). Furthermore, the crystal structures of *E. coli* exomuramidase Slt70 (Thunnissen et al., 1995a), HEWL (Lumb et al., 1994) and T4L (Kuroki et al., 1993), enzymes which like LaL must bind and cleave peptidoglycan, indicate

a methionine residue in or near the vicinity of their active sites (refer to Figs. 4.2 and 3.4). A methionine, Met 502, has also been found to be near the active site in the crystal structure of *E. coli* β-galactosidase (Jacobson et al., 1994). A similar interaction might be expected for LaL although there is no structural evidence for or against this hypothesis. Nevertheless, the ^{19}F NMR evidence presented in the following section suggests that the methionine residues themselves may well experience ligand binding indirectly even though they may not be in direct contact with the ligand.

4.3.5. ^{19}F NMR Studies on TFM-LaL

4.3.5.1. ^{19}F NMR Spectra of High and Low level TFM-LaL and the Effect of $(GlcNAc)_5$ Binding on the Spectra

The ^{19}F NMR spectra of *L*-trifluormethionine and high level TFM-LaL are compared in Figure 4.7. When externally referenced to $CFCl_3$ (0.00 ppm), the resonance of the trifluormethyl group of *L*-TFM appears at a chemical shift of -40.05 ppm (Fig. 4.7 A). The resonances corresponding to the trifluormethyl groups of the TFM residues incorporated into LaL also appear in the same chemical shift region (Fig. 4.7 B). No additional resonances are observed over the chemical shift range of -66 to 8 ppm in both high level TFM-LaL (Fig. 4.7 B) and low level TFM-LaL (results not shown).

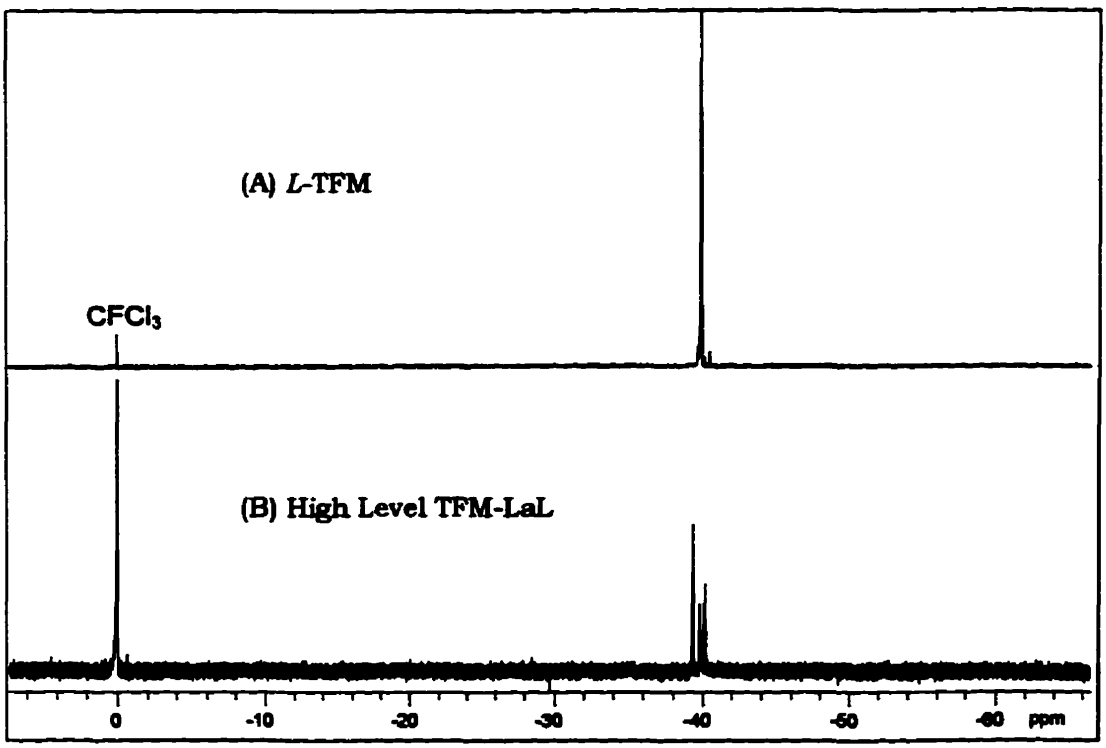


Figure 4.7. Comparison of the ^{19}F NMR spectra of *L*-TFM and high level TFM-LaL.

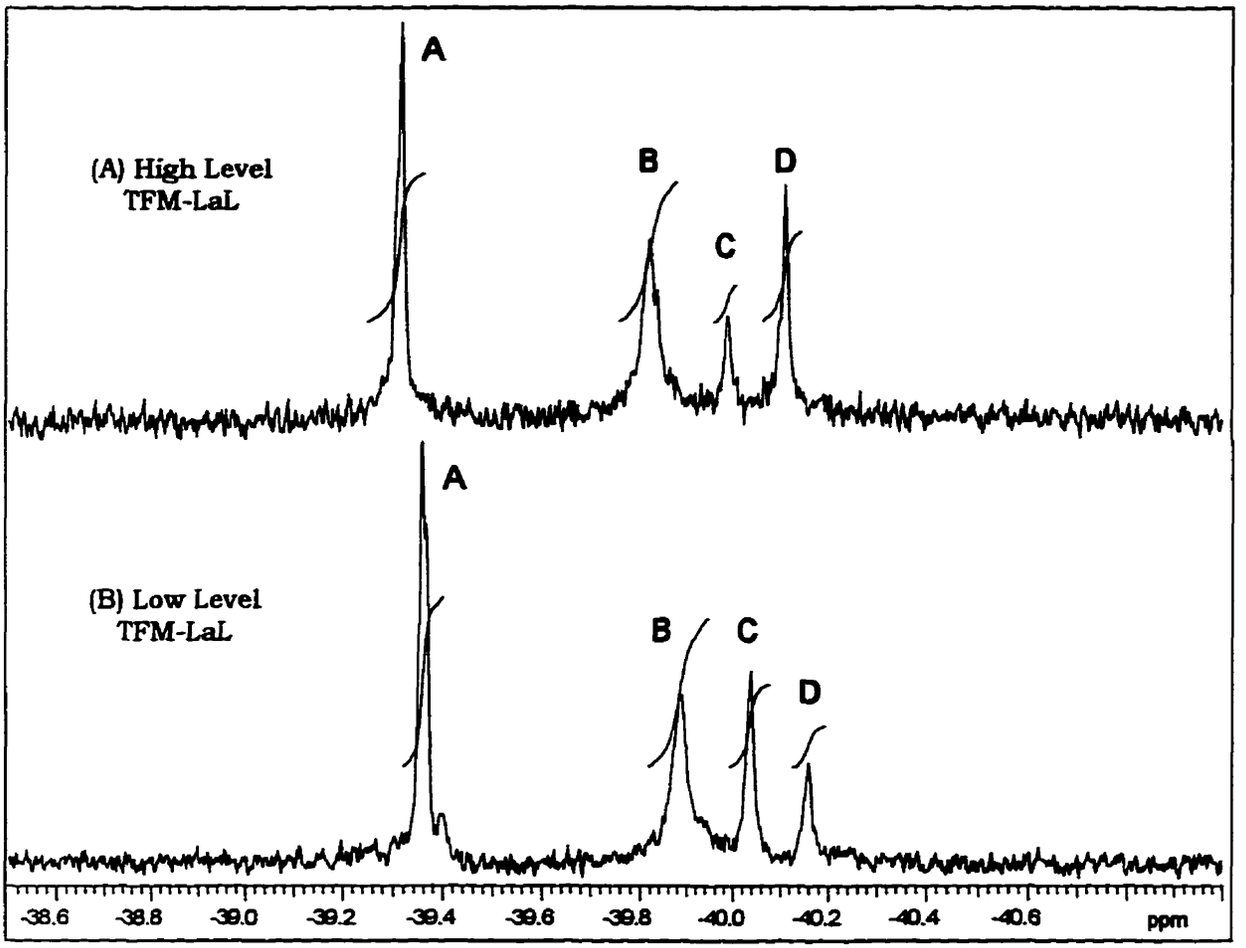


Figure 4.8. ^{19}F NMR spectra of high and low level incorporated TFM-labelled LaL.
(A) Lysozyme labelled with 70% TFM. Protein concentration was 270 μM in 55 mM KPB, pH 7.0. A total of 2000 scans were acquired.
(B) Lysozyme labelled with 31% TFM. Protein concentration was 1.53 mM in 50 mM KPB, pH 7.0. A total of 1500 scans were acquired.

The expanded ^{19}F NMR spectra of TFM-labelled lysozyme prepared under high (70%) and low (31%) level incorporation conditions are presented in Figure 4.8. In both cases four well resolved resonances spanning a chemical shift range of approximately 0.9 ppm is observed. The resonances are rather sharp with the line widths at half height measured to be approximately 7-9 Hz for peaks A, C and D and 15 Hz for resonance B in both spectra. Since the line shape of the resonances is not affected at the two different protein concentrations studied (0.27 and 1.53 mM for high and low level TFM-LaL respectively), no appreciable self-aggregation of TFM-LaL is suggested at the concentrations used for these NMR studies. Other NMR data are summarized in Table 4.4. The small upfield shift (≤ 0.06 ppm) noted for resonances in low level TFM-LaL when

Table 4.4. ^{19}F NMR data for high and low level TFM-LaL.

Resonance[a]	T_1 (s)[b]		Percentage of Total Integrated Area[c]		Chemical Shift[d]	
	High[e]	Low[f]	High	Low	High	Low
A	0.80 ± 0.10	0.84 ± 0.04	33.9 ± 1.2	37.2 ± 1.3	-39.32	-39.36
B	0.38 ± 0.03	0.40 ± 0.01	33.5 ± 1.3	33.3 ± 1.8	-39.82	-39.88
C	0.51 ± 0.17	0.54 ± 0.04	11.8 ± 0.6	19.1 ± 0.5	-39.99	-40.04
D	0.50 ± 0.07	0.59 ± 0.06	20.8 ± 1.1	10.4 ± 1.5	-40.11	-40.16

[a] Refer to Fig. 4.8 for assignments.
[b] ^{19}F spin-lattice relaxation times determined by the inversion recovery method.
[c] Determined by comparing the integral of each resonance to the sum of all resonance integrals in the same spectrum ($n \geq 3$).
[d] Referenced to $CFCl_3$ (0.00 ppm) ± 0.02 ppm.
[e] High level TFM-LaL labelled with 70% TFM. Protein concentration was 0.27 mM in 55 mM KPB, pH 7.0.
[f] Low level TFM-LaL labelled with 31% TFM. Protein concentration was 1.53 mM in 50 mM KPB, pH 7.0.

compared to the analogous ones for high level TFM-LaL may arise from the differences in protein and buffer concentrations of the two samples.

The most striking feature of the ^{19}F NMR spectra of high and low level TFM-LaL is the appearance of 4 resonances and the differences in the intensity of peak C and D between the respective labelled proteins (Fig. 4.8 and Table 4.4). We believe that peaks C and D arise from the same positional methionine and that the relative intensity of each of these two peaks is dependent on the extent of TFM incorporation (discussed below). Lysozyme contains three methionines at positions 1, 14 and 107 and integration of the individual resonances indicates that TFM is incorporated equally well at the three Met positions. Resonances A and B each represent approximately one third of the total integrated area while the integral sum for C and D account for the remaining one third (Table 4.4).

It is possible to suggest from the NMR data that resonance A might correspond to the N-terminal TFM residue. Resonance A exhibits the furthest downfield shift in accord with deshielding effects from a protonated α-amino group. In addition, A is the narrowest resonance (i.e. peak width) and shows the largest relaxation delay (T_1, Table 4.4; the T_1 for *L*-TFM was found to be 2.6 ± 0.02 s) suggestive of the higher mobility that a less restricted N-terminal amino acid may encounter than one found in the interior of the protein.

Although the resonances have not yet been assigned, they are still useful as probes. The observed inhibition of TFM-LaL by chitooligosaccharides prompted us to investigate the effect of $(GlcNAc)_5$ binding on the ^{19}F NMR spectra. Changes in chemical shift on binding $(GlcNAc)_5$ may be a reflection of changes in the environment of the trifluoromethyl group either due to proximity of saccharide, conformational alteration or both. In both high and low level TFM-LaL, resonances B and C/D show very modest upfield shifts on addition of $(GlcNAc)_5$ that increase with increasing sugar concentration while resonance A remains relatively unchanged (Table 4.5). These results are consistent with our previous observations of a slight alteration in two of the three ^{13}C resonances upon binding of $(GlcNAc)_5$ to [*methyl*-^{13}C]Met-labelled LaL (refer to section 3.3.2.4). The spectra of high level TFM-LaL in the presence of $(GlcNAc)_5$ are shown in Figure 4.9 (A). Upon sugar binding resonance B not only exhibits the largest upfield shift but also displays substantial line broadening. The line width at half height increases from 15 Hz (in the absence of $(GlcNAc)_5$) to approximately 26-30 Hz in the presence of $(GlcNAc)_5$ at each concentration studied. The general appearance (peak line widths) of the spectra of low level TFM-LaL upon sugar binding is identical to that shown for high level TFM-LaL except that the intensities of resonances C and D are reversed (Fig. 4.9 (B)).

Table 4.5. Effect of $(GlcNAc)_5$ binding on the ^{19}F NMR resonances of TFM-LaL.

Resonance[a]	Δ δ (ppm) Resulting from $(GlcNAc)_5$ Binding[b]					
	2 mM		5 mM		10 mM	
	High[c]	Low[c]	High	Low	High	Low
A	-0.01	nd	-0.02	-0.02	-0.03	-0.02
B	-0.06	nd	-0.11	-0.09	-0.14	-0.11
C	-0.04	nd	-0.06	-0.05	-0.07	-0.07
D	-0.03	nd	-0.05	-0.05	-0.06	-0.06

[a] Refer to Fig. 4.9 for assignments.
[b] Differences relative to chemical shifts given in Table 4.4 for unliganded TFM-LaL. The negative sign (-) indicates a shift to higher field.
[c] High (70%) and low (31%) level incorporated TFM-LaL respectively.
nd; not determined.

Figure 4.9.

Effect of $(GlcNAc)_5$ binding on the ^{19}F NMR resonances of TFM-labelled LaL.

(A) Spectra of LaL labelled with 70% TFM (high level TFM-LaL). All samples contained 270 uM high level TFM-LaL in 55 mM KPB, pH 7.0. A total of 2000-3000 scans were acquired. The samples also contained no, 2, 5, and 10 mM $(GlcNAc)_5$ as indicated on the spectra.

(B) Spectra of LaL labelled with 31% TFM (low level TFM-LaL). All samples contained 1.53 mM low level TFM-LaL in 55 mM KPB, pH 7.0. A total of 1500 scans were acquired. The samples also contained no, 5, and 10 mM $(GlcNAc)_5$ as indicated on the spectra.

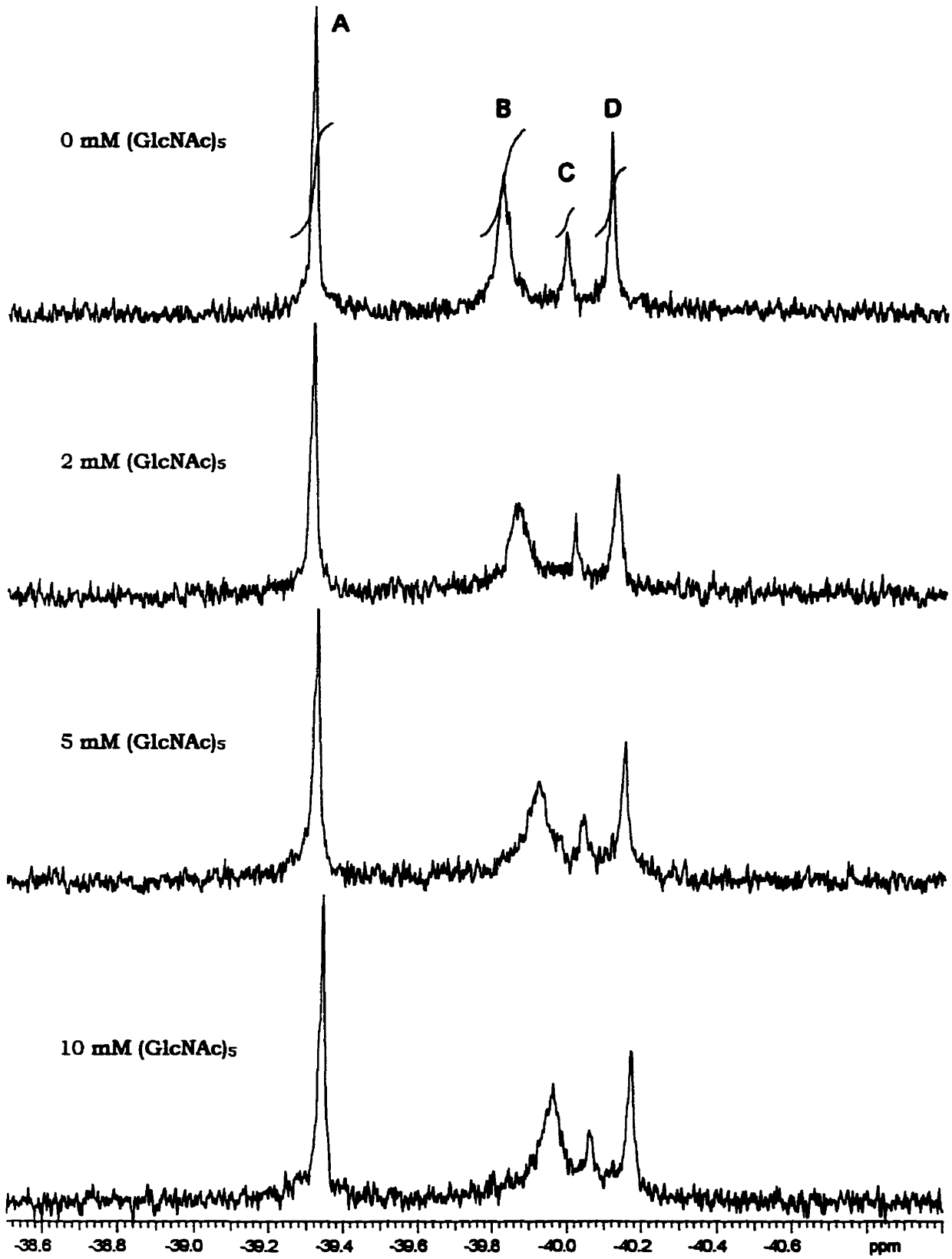


Figure 4.9 (A). High Level TFM-LaL.

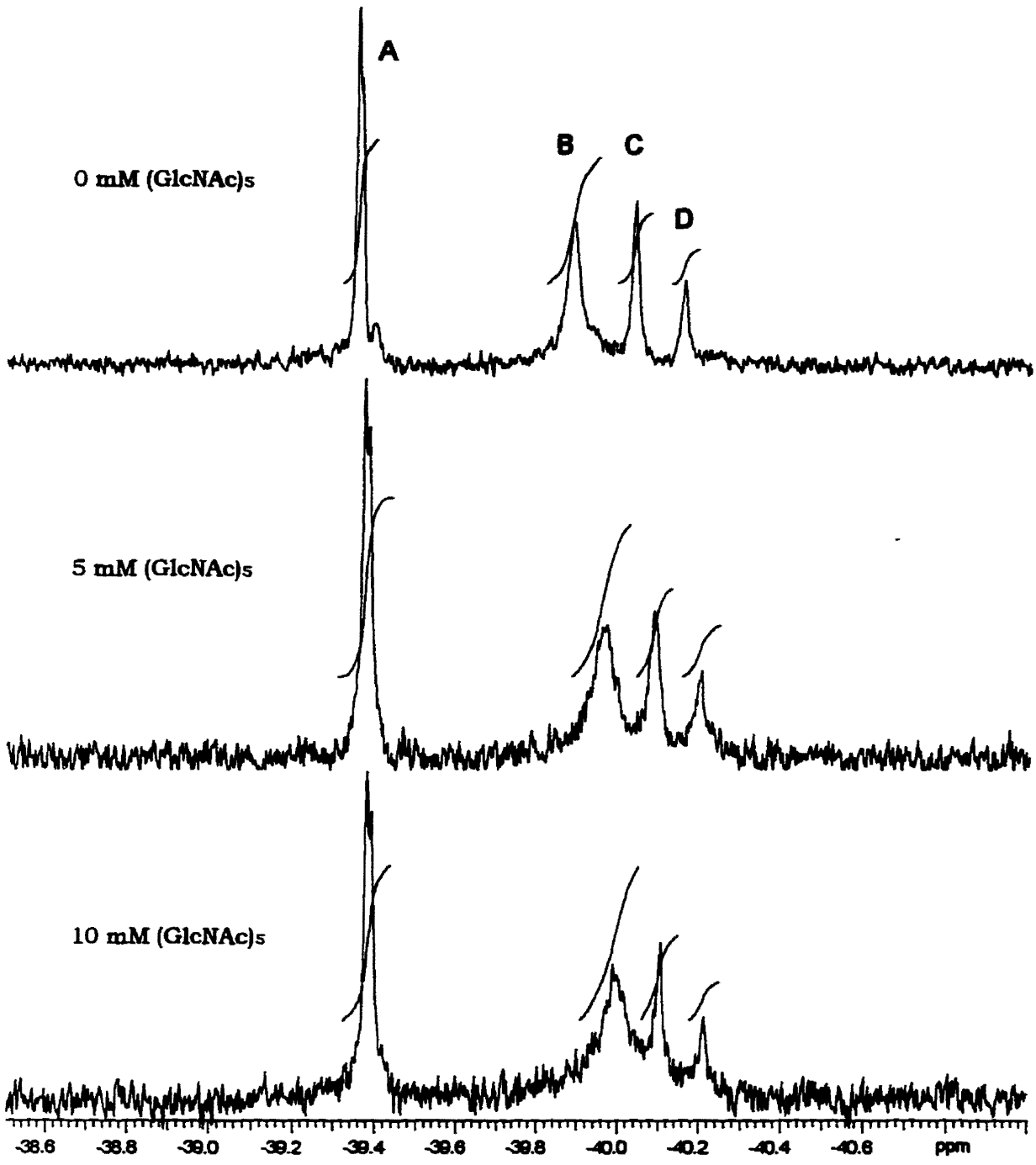


Figure 4.9 (B). Low Level TFM-LaL.

4.3.5.2. Possible Nature of the C/D "Double Resonance"

The NMR results obtained for peaks C and D suggest that they arise from the same positional methionine. If one examines the data for these peaks independently for both high and low level TFM-LaL then within experimental errors (i) the integral sum of C/D approximately equals that of A or B, (ii) they have equivalent spin-lattice relaxation times which differ from A and B and (iii) undergo corresponding chemical shift changes upon $(GlcNAc)_5$ binding (Tables 4.4 and 4.5). These similarities may indicate that the trifluoromethyl groups contributing to C and D occupy comparable microenvironments in the protein structure that differ sufficiently to produce chemically non-equivalent environments resulting in distinct resonances.

Several possibilities could explain the existence of a double resonance corresponding to a single residue and each is addressed below: (1) chemical exchange of the trifluoromethyl group between different microenvironments; (2) chemically different forms of the enzyme (protein heterogeneity) producing microenvironments around the trifluoromethyl group(s) that differ amongst the particular protein species; and (3) distinct, non-interconverting populations of enzyme whose structures are dependent or influenced by the presence of either methionine or TFM at a given position. The extent of incorporation would thereby dictate the relative proportions of each population.

Chemical exchange can be attributable to rotation of the TFM side chain or conformationally different forms of the enzyme. In either case, these exchange processes must be slow enough on the NMR time scale in order to observe separate resonances for a single residue (Wagner & Wüthrich, 1986). These processes could very well reflect attempts by the protein to accommodate the trifluoromethyl group. The volume occupied by the side chain of TFM is expected to be larger than that of *L*-methionine (discussed in section 4.3.6) and local regions in the protein in proximity to the methionine(s) may not be sterically permissive to this substitution. Readjustment of the TFM side chain through rotation may be required and will depend on the flexibility of this side chain. It may also be possible that either the TFM may move within a flexible region of the backbone or there could be movement of an adjacent protein segment to accommodate the TFM.

We believe the possibility of exchange to be unlikely from the following observations. Firstly, spectra obtained for high incorporation TFM-LaL at 22, 35 and 45 °C (results not shown) were indistinguishable regarding peaks C and D. A minor sharpening of resonance B as well as a resonance corresponding to denatured protein was noted at the elevated temperatures and a spectra taken after maintaining the sample at 45 °for 50 min revealed only a single sharp resonance at -39.96 ppm indicative of complete

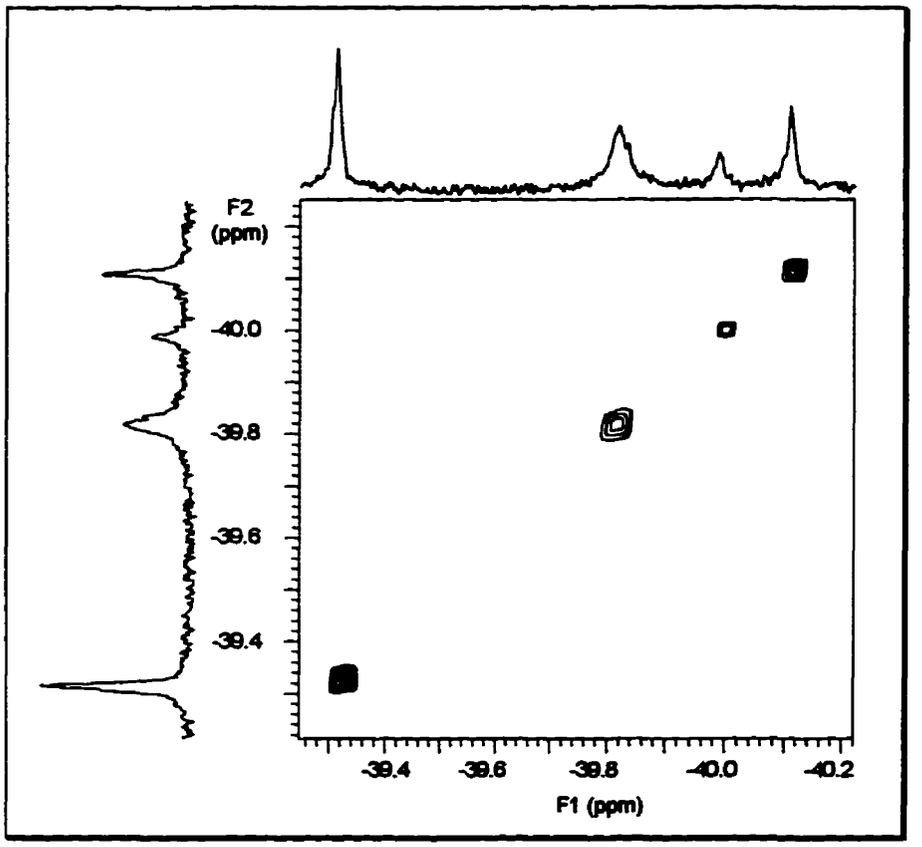


Figure 4.10. Two-dimensional EXSY spectra of high level TFM-LaL.

denaturation. Secondly, we have also recorded phase-sensitive 2D exchange spectra (EXSY) of high level TFM-LaL at mixing times from 0.1 to 0.9 s and found no evidence of chemical exchange for C and D from the absence of any correlations between their resonances (Fig. 4.10).

The second possibility as to the nature of the double resonance is protein heterogeneity. A perceptive observation made by Li et al. (1989) was that doubling of resonances in the ^{19}F NMR spectra of 6F-Trp-labelled rat cellular retinol-binding protein II was due to the presence or absence of the N-terminal methionine residue (the relative intensities of the two components of the two resonances corresponded to the relative amount of initiator methionine present). We also have evidence of N-terminal heterogeneity. ESMS of several preparations of *wt* (refer previously to section 2.3.2.2 and Fig. 2.28) and TFM-LaL (Fig. 4.5) demonstrated the presence of species corresponding to the loss of a methionine residue, probably arising from co- or post-translational processing by methionine aminopeptidase. The results can not indicate whether an N-terminal Met or TFM is removed. For example, in Fig. 4.5 (B) the peak at 17750 Da could

be representative of a *mono*-labelled TFM-LaL which has lost a N-terminal Met or of a *bis*-labelled TFM-LaL which has lost a N-terminal TFM. However, the typically low abundance of these species (< 5%) could not account for the intensity ratio observed for C and D.

With the intention of possibly identifying the N-terminal TFM resonance, attempts were made to utilize recombinant methionine aminopeptidase (Roderick & Matthews, 1993) to remove the N-terminal trifluoromethionine of TFM-LaL *in vitro*. However this enzyme failed to remove the N-terminal methionine residue in *wt* LaL presumably due to either the presence of a non-optimal amino acid residue, valine, in the adjacent position or inaccessibility of the N-terminus. As indicated in Fig. 4.11, incubation of *wt* LaL with MAP for 68 hr resulted in only a marginal increase in the percentage of LaL lacking a methionine residue (presumably the N-terminal methionine). Hence, this approach was not extended to achieve the deletion of the N-terminal residue from LaL preparations containing TFM residues.

Although it is well known that methionine residues are susceptible to oxidation (Brot & Weissbach, 1991), we have no ESMS evidence for the occurrence of such modifications in TFM-LaL which might contribute to protein heterogeneity and the origin of the C/D resonances. In addition, we have found that *L*-TFM itself is inert to hydrogen peroxide oxidation under conditions which are known to rapidly convert *L*-methionine into its sulfoxide (Brot et al., 1984). As well, gauge-independent atomic orbital (GIAO) calculations (Wolinski et al., 1990) involving the *ab initio* calculation of nuclear magnetic resonance (NMR) chemical shifts were used to determine the change in the average ^{19}F chemical shift of the trifluoromethyl group on sulfur oxidation. GIAO calculations on ethyl trifluoromethyl sulfide and its sulfoxide at the RHF/6-311+G(d,p)//RHF/6-31G* level indicated a chemical shift change from -34.8 ppm for the sulfide (relative to $CFCl_3$ as reference) to -62.8 ppm for its sulfoxide. Hence, the C/D ^{19}F NMR resonances do not originate from the presence of trifluoromethionine sulfoxide residues in LaL.

Proline cis-trans isomerization can also be considered to produce chemical heterogeneity in a protein and was used as a possible explanation to account for two resonances observed for 5F-Trp-labelled D-galactose/D-glucose chemosensory receptor (Luck & Falke, 1991a). Although LaL contains 5 prolines, it is unlikely that the nature of the double resonance arises from isomerization since (i) lysozyme labelled with the non-perturbing probe [*methyl*-^{13}C]Met did not produce doubling of any of the three resonances observed in its [^{1}H-^{13}C]HMQC spectra (refer to Fig. 3.21) and (ii) the relative occurrence

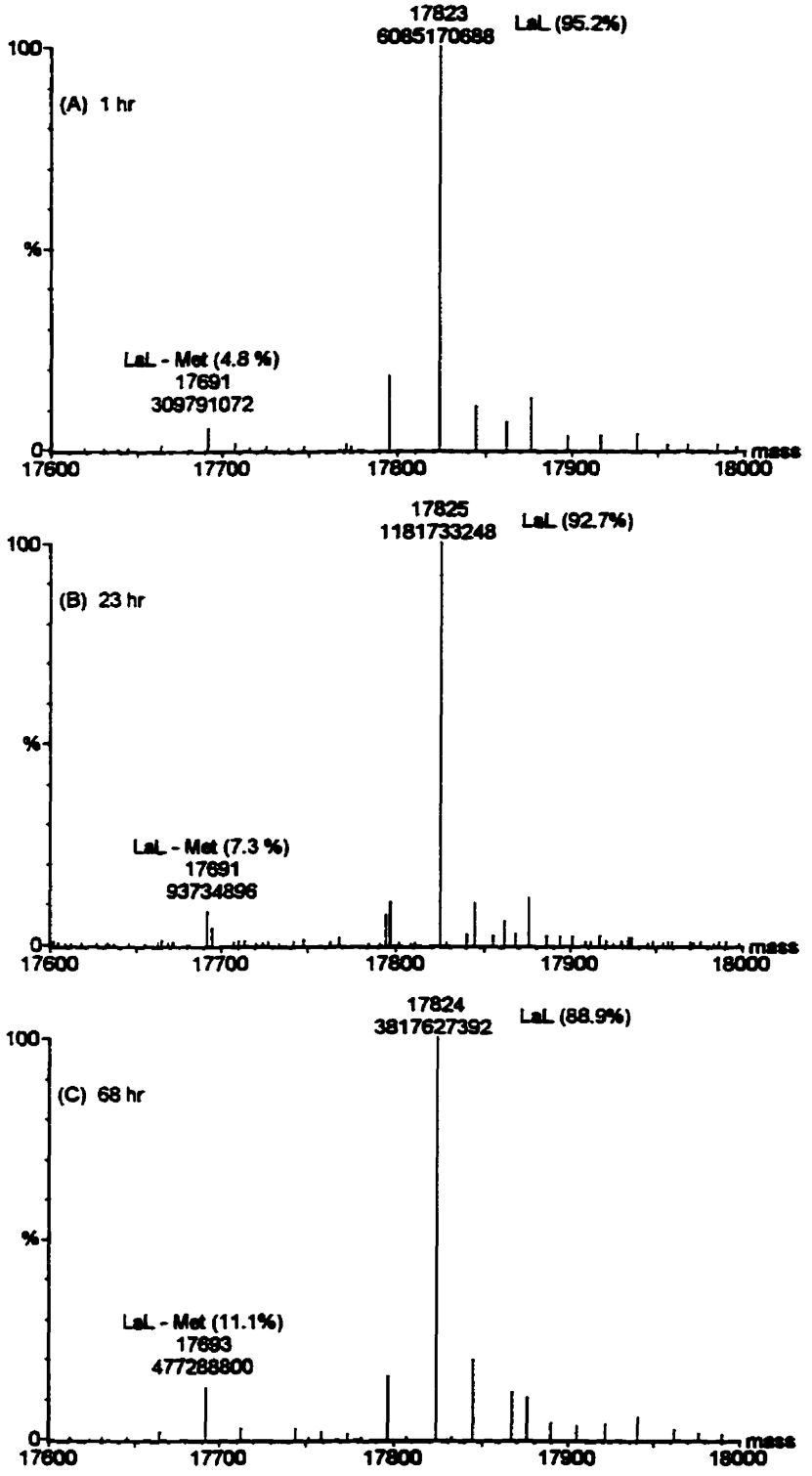
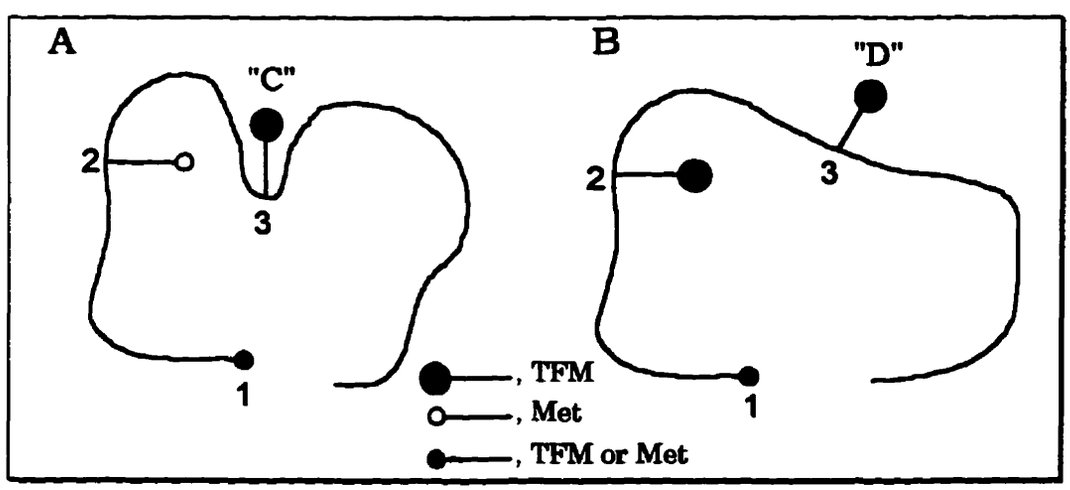


Figure 4.11. Methionine aminopeptidase activity on *wt* LaL. LaL (2.5 mg/mL) and MAP (0.25 mg/mL) were incubated at 30 °C. At the indicated times, aliquots were removed and chromatographed over the Delta Pak™ C_{18} column and the fraction containing LaL was subjected to ESMS. Shown are the bar spectra indicating the mass and intensity of the peaks corresponding to LaL and LaL less the mass of a methionine. The relative proportions of the two species in each spectrum is indicated as a percentage.

of separate protein species (and therefore, relative intensities of C and D) would depend on the degree of isomerization and not, as is observed, on the extent of incorporation. One could however argue that isomerization may be influenced or even induced by the extent of TFM incorporation and therefore, the intensities of C and D could represent varying degrees of one or several isomerizations.

The third argument possibly explaining the double resonances is the existence of distinct, non-interconverting conformers of TFM-LaL whose relative populations are dictated by TFM incorporation. An exaggerated model depicting this concept is shown in Scheme 4.2. The methionine positions are arbitrarily designated as 1, 2 and 3. An assumption is made that the TFM (large solid circle, Scheme 4.2) at position 3 gives rise to the double resonance C/D. In situation A, position 2 is occupied by methionine (small open circle) and a particular local environment exists around the TFM at position 3, giving rise to resonance C. Incorporation of the more sterically demanding TFM at position 2 results in an altered protein conformation (situation B) which places the TFM at position 3 in a new environment, hence giving rise to resonance D. Obviously when methionine occupies position 3, no resonance is detected but when position 3 is TFM, it serves as an indicator as to the occupancy of position 2. With lower incorporation levels, there is less likelihood of TFM at position 2 and situation A and resonance C dominate. With higher levels of incorporation, the degree of incorporation at position 2 increases and situation B and resonance D dominate. Our integration results appear to support such a scenario. With low level TFM, position 2 should statistically be occupied with TFM 31% of the time when position 3 is TFM, and there would be a 2:1 ratio of resonance C to D as is observed.



Scheme 4.2.

Conversely, high level TFM would result in a 70% likelihood of TFM occupying position 2 when position 3 is TFM and there would exist a 2:1 ratio of D to C, and this is also observed. Since resonance C prevails under lower incorporation levels, it may be more indicative of the unperturbed structure. The model assumes that the two situations are not dependent on the occupancy of position 1, which could be Met or TFM (small solid circle). The chemical shift difference between C and D is very small (0.12 ppm) and could reflect very minor differences in the overall protein structures in the two situations. However, the chemical shift of fluorine is extremely sensitive to the local environment and even slight differences between the protein structures can generate sufficiently distinct microenvironments.

4.3.6. Molecular Modelling Calculations

As the replacement of methionine residues in protein may be of general use to the study of protein-ligand and protein-protein interactions by ^{19}F NMR spectroscopy, it is interesting to explore some of the fundamental properties of the side chain of TFM compared to that of *L*-methionine. The thiobutane side chain of *L*-methionine is important for a number of biochemical properties. The side chain itself is more flexible than those of related amino acids such as leucine and isoleucine and this has been ascribed to the longer CH_3S-CC bond which reduces the steric repulsions usually seen for all carbon alkyl chains (discussed previoulsy in section 4.1.1).

Experimental evidence over the years, based on electron diffraction (Oyanagi & Kuchitsu, 1978) and microwave vibrational spectroscopy (Nogami et al., 1975), has indicated that the gauche and trans (or anti) conformations of ethyl methyl sulfide are close in energy and in fact both are found in the gas and liquid phases. Recently Durig and coworkers (Durig et al., 1991) have reported that in fact the trans conformation may be slightly more stable in the gas phase than the gauche by approximately 0.38 kcal/mol. This flexibility of the thiobutane moiety permits numerous conformations to be taken up by the methionine side chain. In addition, the sulfur atom in methionine is a polarizable atom and this has been suggested to play a role in sequence independent recognition of non-polar protein surfaces.

In the case of TFM, one needs to consider the possible conformational and electronic effects that the incorporation of fluorine has on the methionine structure. In terms of conformation, we have utilized both semi-empirical (AM1, PM3) as well as *ab initio* calculations to study ethyl methyl sulfide and ethyl trifluoromethyl sulfide as models for Met and TFM. Since methionine is usually encountered in the hydrophobic interior of

most globular proteins, gas phase calculations may in fact be appropriate indicators for this system.

The semi-empirical methods, AM1 and PM3, predict a slightly more stable gauche conformation than that of the trans conformation for ethyl methyl sulfide of 0.17 kcal/mol and 1.00 kcal/mol respectively. For ethyl trifluoromethyl sulfide, AM1 and PM3 also predicted a slightly more stable gauche conformation of 0.33 kcal/mol and 1.12 kcal/mol indicating that there appears to be very little energy difference between the conformations of the fluorinated and the non-fluorinated methionine side chain. The gauche angles predicted for the AM1 and PM3 geometry optimized structures for ethyl methyl sulfide are 72.4° and 70.6° and for ethyl trifluoromethyl sulfide are 78.6° and 75.5° respectively. A slightly larger angle is therefore predicted for the gauche conformation of TFM although as stated the energies of the trans and gauche conformations are very close.

High level *ab initio* calculations with geometry optimizations at the //RHF/6-31G* level have been shown to rather accurately predict the experimental stabilities and structures of sulfides such as ethyl methyl sulfide (Durig et al., 1991; Markham & Bock, 1995). At the RHF/6-31G*//RHF/6-31G* level, both the trans conformers of ethyl methyl sulfide and ethyl trifluoromethyl sulfide are more stable than their gauche conformers by 0.41 kcal/mol and 0.39 kcal/mol respectively. However, if electron correlation is taken into account utilizing the MP2/6-31G**//RHF/6-31G* level, the gauche conformer of ethyl methyl sulfide and ethyl trifluoromethyl sulfide are predicted to be only slightly more stable than the trans by 0.002 kcal/mol and 0.27 kcal/mol respectively. It appears that even in the fluorinated case, the trans and gauche conformers are very close in energy and are probably both sampled and are energetically attainable in TFM. The dihedral angles for the gauche conformations of ethyl methyl sulfide and ethyl trifluoromethyl sulfide at this level of calculation are 70.5° and 80.8° respectively indicating an "opening up" of the gauche structure in the case of TFM which may contribute to the overall properties of a TFM-containing protein such as LaL.

The associated bond lengths and bond angles of the side chains of norleucine, methionine, selenomethionine and TFM from geometry calculations at the //RHF/6-31G* level are presented in Table 4.6. The bond lengths and angles for CF_3-S and CF_3-S-CH_2 (Table 4.6) are seen to be intermediate among the values found for other analogues which have been successful replacements for methionine (Budisa et al., 1995 and references therein). The Mulliken charge on the sulfur atom in TFM and the highest-occupied molecular orbital (HOMO) energy for this analogue compared to the methionine and

Table 4.6. Structural and Electronic Characteristics of *n*-Butane and Analogues.

	$CH_3CH_2CH_2CH_3$ X = CH_2	$CH_3SCH_2CH_3$ X = S	$CH_3SeCH_2CH_3$ X = Se	$CF_3SCH_2CH_3$ X = S
Bond Length CH_3-X or CF_3-X (Å)	1.53	1.81	1.94	1.79
Bond Length X-CH_2 (Å)	1.53	1.82	1.96	1.82
Bond Angle H_3C-X-CH_2- (degrees)	113.1	100.2	98.2	99.3
Mulliken Charge on X (electrons)	-	0.113	-0.023	0.153
Dipole Moment (Debye)	0.00	1.80	1.74	2.66
HOMO Energy (eV)	-	-9.027	-0.310	-10.380
Van der Waals Surface Area (Å^2)	108.9	111.4	115.9	121.0
Solvent-Accessible Surface Area[a] (Å^2)	240.8	243.8	251.2	257.4
Van der Waals Volume (Å^3)	78.4	80.6	85.0	87.5
Solvent-Accessible Volume[a] (Å^3)	321.3	326.9	339.3	349.7
log P	2.09	0.80	-	2.73

[a] Probe Radius = 1.4 Å

selenomethionine side chains may be used to explain the decreased nucleophilicity of the sulfur atom in TFM and its reduced reactivity with hydrogen peroxide and cyanogen bromide (a detailed account of the reaction of cyanogen bromide with TFM-LaL is presented in section 4.3.7). However, it should be noted that the CF_3S- moiety itself does not entirely preclude the possibility of its ligation to metal centers. For example, the crystal structure of a CF_3-sulfide ligand to platinum in *cis*-dichloro[1,2-bis(trifluoromethylthio)propane]platinum (II) indicates ligation of the sulfur to platinum and hence TFM may indeed exhibit ligation in proteins where normally methionine is encountered (Manojlovic-Muir et al., 1977). In the case of fluorinated aromatic amino acids such as those utilized for Trp, Phe and Tyr, the electron density of the aromatic π systems will also be affected although this has not detracted from their highly successful application as ^{19}F nmr probes.

Van der Waals and solvent accessible surface areas and volumes (Table 4.6) based on //RHF/6-31G* level geometries for the TFM side chain are somewhat larger than the values calculated for the norleucine, methionine and selenomethionine side chains. Side chain van der Waals volumes calculated for norleucine, methionine, selenomethionine, telluromethionine and TFM utilizing PM3-determined geometries resulted in volumes of 78.6 Å^3, 81.0 Å^3, 85.0 Å^3, 93.7 Å^3 and 89.2 Å^3 respectively. Telluromethionine has recently been successfully incorporated into several proteins, although telluromethionine has been found to be extremely sensitive to oxidation (Boles et al., 1994; Budisa et al., 1995;

Kambrook et al., 1996). With *ab initio* //RHF/3-21G* level geometries, a level reasonably accurate to accomodate the tellurium atom, the van der Waals and solvent accessible volumes for ethyl trifluoromethyl sulfide (87.8 Å^3 and 349.9 Å^3 respectively) and 2-tellurobutane (93.7 Å^3 and 363.7 Å^3) indicate that the telluromethionine side chain is somewhat larger than the TFM side chain. Steric effects with TFM nevertheless must play a role as indicated by the observation of two NMR signals for a single TFM. However, these effects must be subtle as the enzyme is still catalytically active and the other TFM residues exhibit only single ^{19}F signals. Particular methionine positions may be more or less affected by the introduction of TFM at their respective positions. It is interesting to note that selenomethionine incorporation, widely utilized as a heavy atom derivative in protein crystallography, may produce a thermally less stable enzyme although the X-ray structure of the selenomethionine-containing enzyme may not be visibly different from the wild type enzyme (Boles et al., 1991). Recently, Bernard and co-workers (Bernard et al., 1995) have reported that selenomethionine incorporation into phosphomannose isomerase changes the kinetic parameters of the enzyme and that, for this particular protein, even selenomethionine has introduced too large a steric/electronic effect into the protein which alters its fundamental kinetic parameters.

The CF_3-group is hydrophobic and the log P value for ethyl trifluoromethyl sulfide was calculated to be 2.73 compared to 0.80 and 2.09 for ethyl methyl sulfide and *n*-butane respectively. This predicted hydrophobicity of a TFM residue is exemplified experimentally by the increase in retention time observed on reverse-phase HPLC by LaL molecules containing higher levels of TFM. It was observed that samples of TFM-LaL desalted over the Delta Pak™ C_{18} in preparation for ESMS analysis tended to produce somewhat broader peaks than did *wt* LaL employing a sharp elution gradient (refer to section 4.2.4 for details on this gradient). When TFM-LaL was chromatographed using a much shallower gradient it was possible to obtain partial separation of the component species and analyze these individual fractions by ESMS (Fig. 4.12). Clearly, retention times on the C18 hydrophobic support increased with proportionate content of TFM, with *tris*-labelled TFM-LaL having the strongest interactions. The results demonstrate the increased hydrophobic character to the protein imparted by the trifluoromethyl group. Although it is to be expected that the protein adopts an extended or denatured conformation when chromatographed, it is interesting to note that the *bis*-labelled species elutes in two major fractions (peaks 3 and 4) suggesting a possible positional influence on retention time. If the methionine residues are important in the hydrophobic effect for folding, the incorporation of TFM into the protein appears not to greatly affect the folding process since the resulting TFM-LaL is enzymatically active and does bind inhibitors in a similar fashion to wild type LaL.

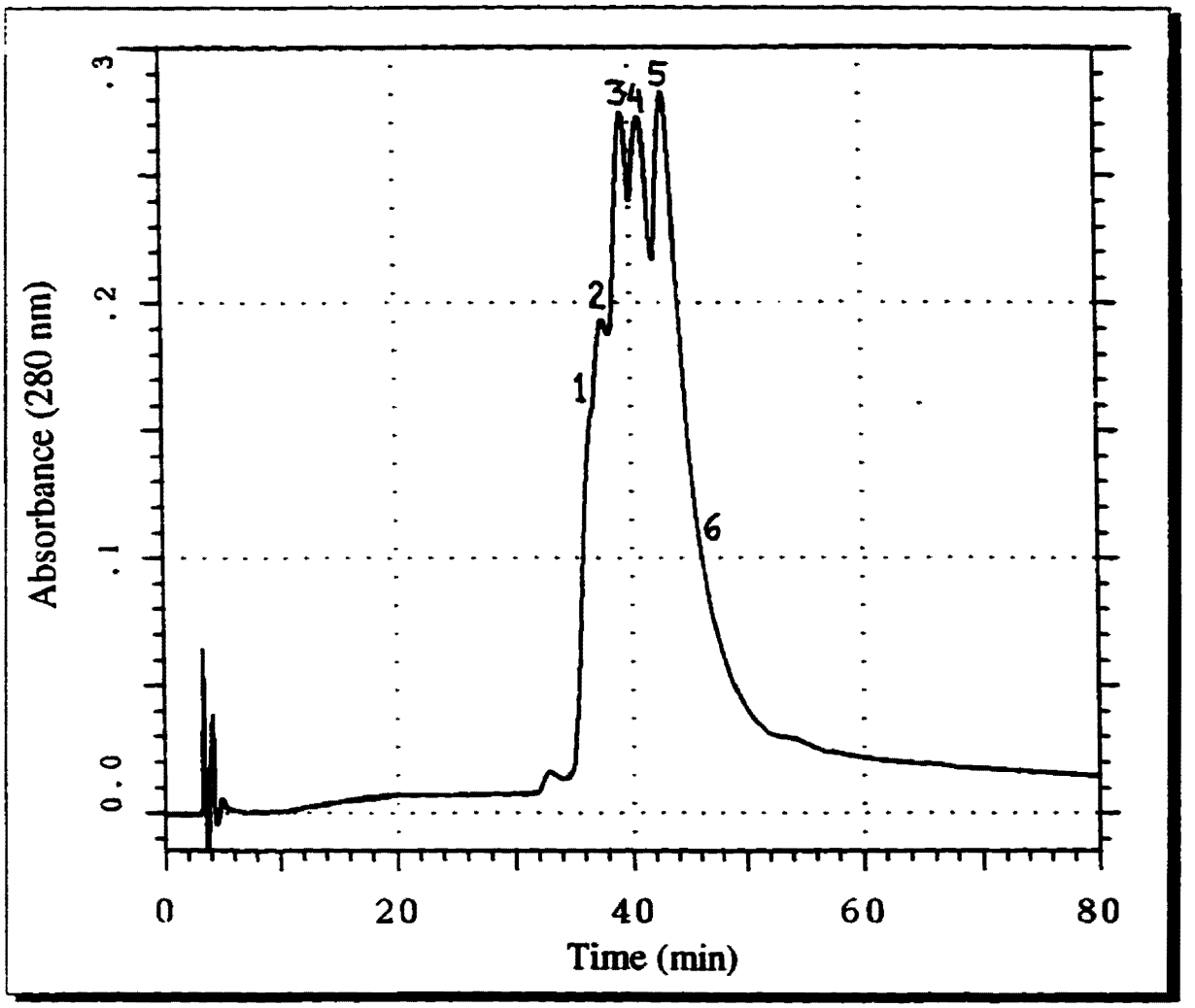


Figure 4.12. Reverse phase chromatography of high level TFM-LaL.

The sample was chromatographed over the Delta Pak™ C_{18} column employing linear gradients from buffer A (0.1% TFA in MQW) to buffer B (0.1% TFA in CH_3CN) and a flow rate of 0.2 mL/min. The column was equilibrated with 85% Buffer A and the gradient was 0-10-20-80 min, 25-35-37-43% buffer B. Fractions (1-6) were collected and subjected to ESMS to identify the respective components present.

The major (and minor in parentheses) species detected in each fraction were:

(1) *wt* LaL and *mono*-TFM-LaL (Met⁻ *wt* LaL)
(2) *mono*-TFM-LaL (*bis*-TFM-LaL, Met⁻ *mono*-TFM-LaL)
(3) *bis*-TFM-LaL (Met⁻ *mono*-TFM-LaL)
(4) *bis*-TFM-LaL (Met⁻ *bis*-TFM-LaL)
(5) *tris*-TFM-LaL (*bis*-TFM-LaL)
(6) *tris*-TFM-LaL.

The prefix Met⁻ denotes a species of mass corresponding to loss of a single Met residue.

4.3.7. Cyanogen Bromide Treatment of LaL and TFM-LaL Proteins

Cyanogen bromide treatment of proteins is a method widely used for protein fragmentation. The low natural abundance of methionine in proteins makes for a limited number of peptide fragments advantageous for separation and isolation for primary sequence studies. Since the discovery of the reaction in 1961 (Gross & Witkop, 1961) and its first application to the cleavage of methionyl peptide bonds in ribonuclease (Gross & Witkop, 1962), the technique has undoubtedly become the most widely used chemical reaction for protein fragmentation.

The proposed mechanism of the cleavage reaction (Gross, 1967) is shown in Fig. 4.13. The sulfur of methionine nucleophilically attacks the cyanide group of CNBr generating an intermediary cyanosulfonium salt. Nucleophilic attack by the methionyl carbonyl oxygen on the sulfonium results in release of methyl thiocyanate and formation of the iminolactone. Hydrolysis of the iminolactone generates the cleavage products, the

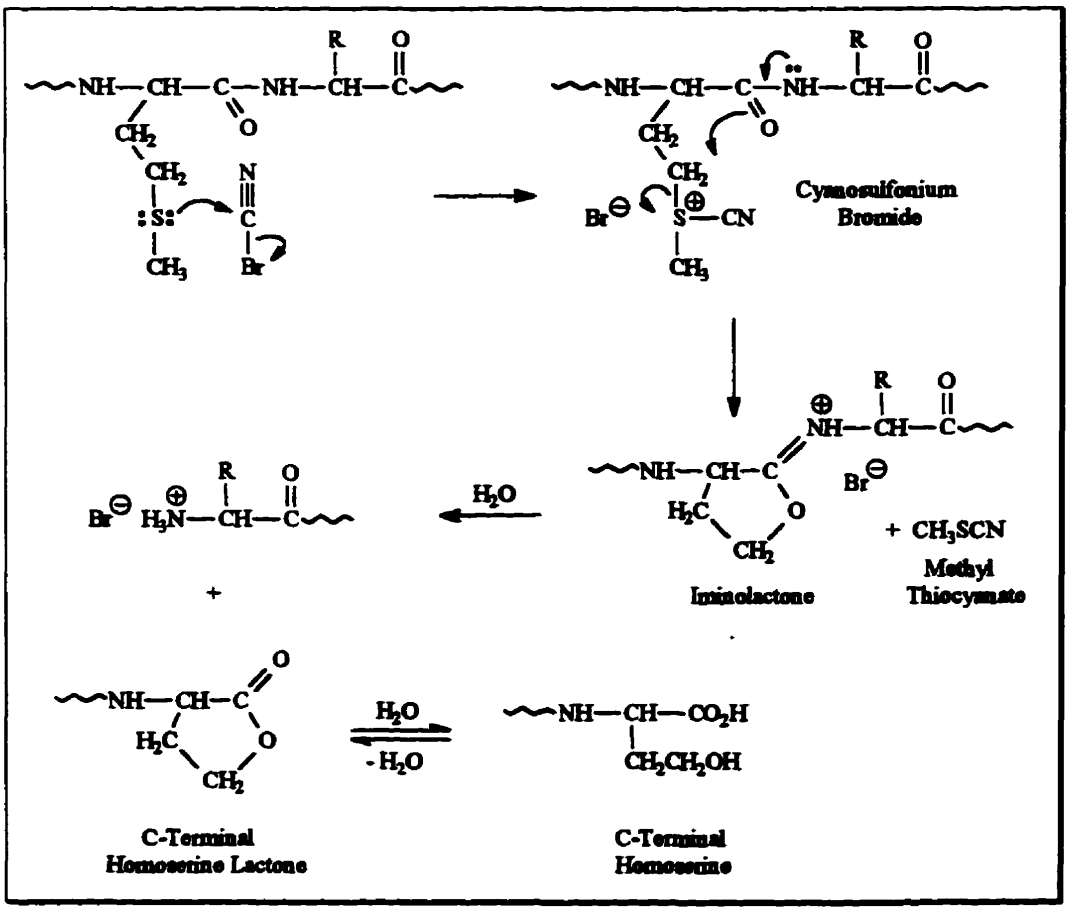


Figure 4.13. Mechanism of CNBr cleavage of methionyl peptides (Gross, 1967).

amino peptide and the C-terminal peptide. All fragments produced (except for the C-terminal peptide of the protein) will contain a mixture of either a C-terminal homoserine lactone or C-terminal homoserine which are in equilibrium.

We were interested in establishing whether proteins containing TFM would be susceptible to fragmentation by reaction with CNBr. To this end it was first desirable to establish conditions that would result in optimal fragmentation of LaL since there were limited quantities of high level TFM-LaL. A survey of the literature provided the following typical cleavage conditions. The protein (10-20 mg/mL) is dissolved in 70% formic acid and solid CNBr (1-2 mg of reagent/mg of protein) is added directly to the protein solution. CNBr is usually present at a 20- to 100-fold molar excess with respect to methionine. Aqueous trifluoroacetic acid and 0.1 M HCl have also been used as the reaction medium although 70% formic acid is the most generally used (Fontana & Gross, 1986). The acidic solvent protonates ε-amino groups and other basic groups suppressing their cyanylation. At neutral or alkaline pH mono- and di-amino acids are involved in several reactions with CNBr (Schreiber & Witkop, 1964). In addition, 70% formic acid is a strong protein unfolding reagent resulting in exposure of Met residues to CNBr (Fontana & Gross, 1986). Since methionine sulfoxide does not react with CNBr (Gross, 1967) the reaction is performed under an inert atmosphere to suppress oxidation.

Initial attempts were done to ascertain if protein concentration would affect the efficiency of cleavage. Reactions with *wt* LaL were performed at 0.1, 1.0 and 10 mg/mL employing a 100-molar excess of CNBr for 24 h period (under standard conditions). Analysis of the reactions by RP-HPLC revealed that extensive and essentially equal cleavage was obtained at the protein concentrations of 1.0 and 10 mg/mL however a marginal yet noticeable decrease in fragmentation occurred at 0.1 mg/mL over the time of incubation. It is evident that a reduction in the concentration of the reactants, namely the protein methionines and CNBr, is accompanied with a slower reaction rate which is observed for the lowest concentration of reactants. Since the 1.0 and 10 mg/mL concentrations afforded analogous results and as it was deemed advantageous to be able to manipulate larger reaction volumes and therefore, less concentrated protein samples, further studies were performed using protein at 1.0 mg/mL. Therefore, the cleavage conditions utilized for LaL and TFM-LaL consisted of 1 mg/mL (56 μM) protein and 1.78 mg/mL (16.8 mM) cyanogen bromide (100-molar excess/methionine) in 70% formic acid.

Fragmentation of a protein with CNBr is expected to produce a mixture of peptides, with yields exceeding 80% being typical (Spande et al., 1970). Some doubts have been raised concerning the quantitative nature of the overall reaction (Link & Stark, 1968). In order to aid in the interpretation of subsequent results, all possible CNBr generated peptides from LaL and TFM-LaL and their respective masses are illustrated in Fig. 4.14 and the reader is urged to make frequent reference to Fig. 4.13 for the following discussion.

Treatment of LaL and high level TFM-LaL (for these experiments, the high level TFM-LaL used contained 74% TFM) with CNBr were performed under identical conditions. Evaluation of cleavage was accomplished by reverse phase HPLC of the reactions and mass spectral analysis of the fragments produced. The reactions were observed over time with the intention of possibly identifying a difference in the cleavage rate between Met and TFM residues. The HPLC chromatograms of the time points taken (0, 2, 6, 24 and 48 h) are shown in Fig. 4.15.

An examination of the chromatograms for the digests of LaL demonstrate that over time, LaL was nearly quantitatively cleaved into smaller peptides. Analysis of the fragments produced by CNBr by ESMS and MALDI-MS (Table 4.7) confirmed that the peptides generated were consistent for cleavage of *wt* LaL at the methionine positions (refer to Fig. 4.14). Comparison of the relative intensities of the peaks on the chromatograms and their mass identification offer the following. The early time points consisted exclusively (0 h, pk 1-3) or predominantly (2 h, pk 6) of intact LaL and only a small generation of peptides in the 2 h reaction that would result from only a single cleavage at either M14 (2 h, pk 1 "AM*l*", pk 6 "E") or M107 (2 h, pk 4 "DM*l*", pk 5 "C") or cleavage at both M14 and M107 (2 h, pk 2, "B*l*"). Chromatograms of aliquots from the longer reaction times (6, 24 and 48) showed a higher degree of cleavage with a concomitant decrease of the larger molecular weight fragments. Very little intact LaL remained (or could be detected) in both the 24 and 48 h samples although there still existed a significant amount of fragment E (24 h, pk 7; 48 h, pk 11). Elution of the peptides from the column essentially occurred in order of increasing molecular weight except for peptide B and D/DM which preceded and co-eluted with peptide C respectively.

The reaction of LaL with CNBr went efficiently but was not quantitative. A quantitative cleavage of LaL would result in a equal molar production of peptides A, B and C. However, it appears that fragmentation at M107 and to a higher extent at M1, was much less than that at M14. Inefficient reaction at M107 would result in the persistence of peptide E, as was mentioned above, although substantial cleavage did occur at M107 as

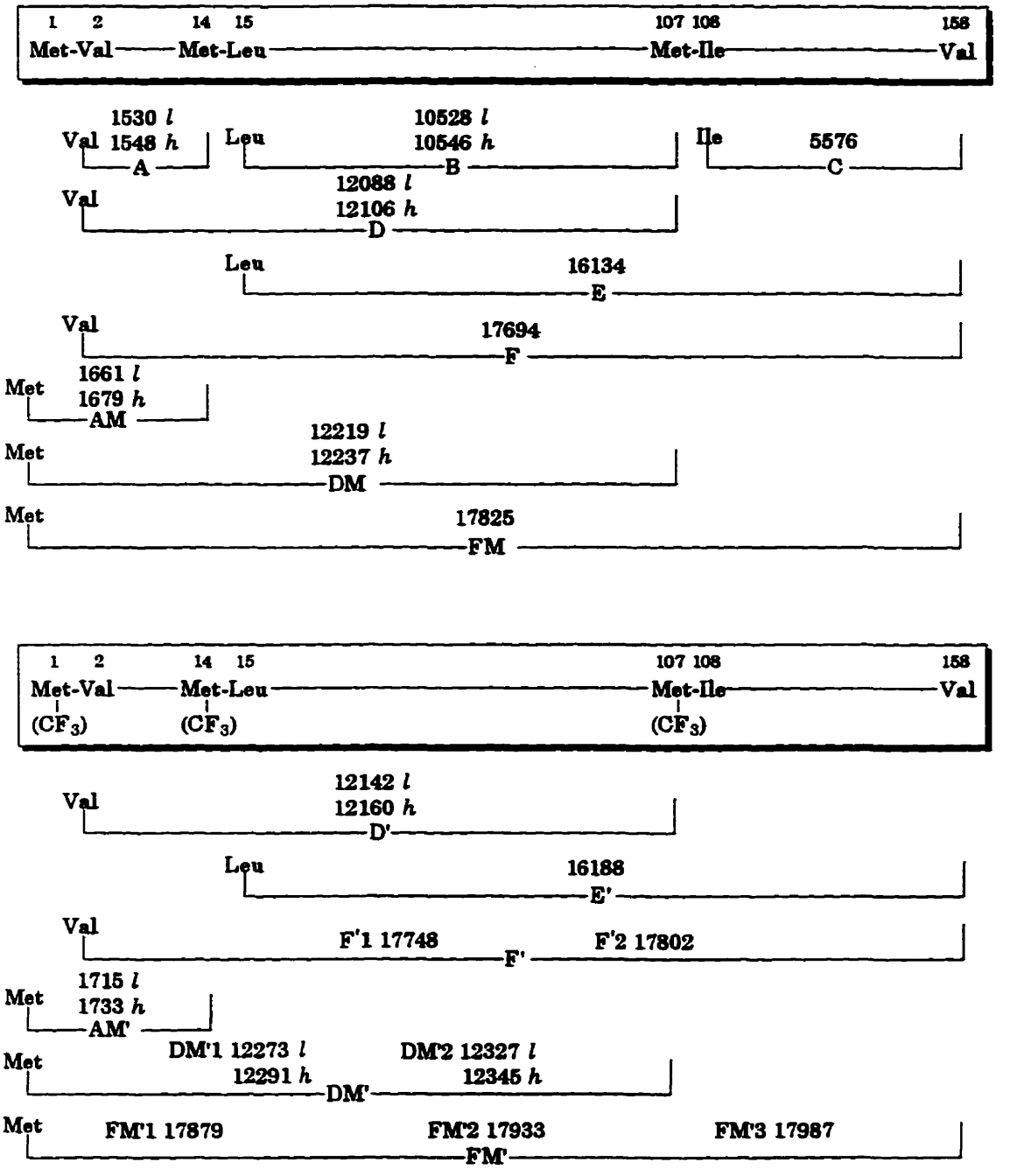


Figure 4.14. Possible products from the CNBr cleavage of *wt* LaL (top) and TFM-LaL (bottom). The products for TFM-LaL also include those for LaL and those given are in addition to and include those shown for *wt* LaL. Given at the top of each is the sequence and residue numbering of LaL. Each fragment is designated with a capitol letter (A-F) and for simplicity, those retaining the N-terminal Met (or TFM) are further designated with an M. Where appropriate, peptides are also designated with an "*l*" (homoserine *l*actone) or an "*h*" (*h*omoserine) indicative of the C-terminus. Fragments containing TFM are also designated with a letter followed with a "'" (prime). If a fragment could contain 2 or 3 TFM, the prime symbol is followed by a number indicating the total number. The calculated molecular weight of each fragment is given.

Figure 4.15.

RP-HPLC chromatograms monitoring the reaction of *wt* LaL and high level TFM-LaL (74% TFM incorporation) with CNBr. Reactions contained 56 uM of the respective proteins and 16.8 mM CNBr in 70% formic acid. At the indicated times, aliquots (60 μL, 60 μg protein) were removed and dried by lyophilization and chromatographed over a Delta Pak™ C_{18} column (2.0 mm x 15 cm) as outlined in section 4.2.9.1 using the gradient listed in Table 4.1 (only 30 μg was chromatographed for the 48 h samples). The abscissa axis is time (min) and the ordinate axis is absorbance at 215 nm and the ordinate scales are unique for each chromatogram.

Left Panels: Chromatograms for *wt* LaL.
Right Panels: Chromatograms for high level TFM-LaL.

The indicated peaks were collected and analyzed by ESMS and/or MALDI-TOF MS and the results are given in Tables 4.7 and 4.8. Chromatograms indicated as "untreated" were obtained from chromatography of the respective proteins (60 μg) without exposure to CNBr or to formic acid. Peaks appearing after 74 min result from the gradient employed.

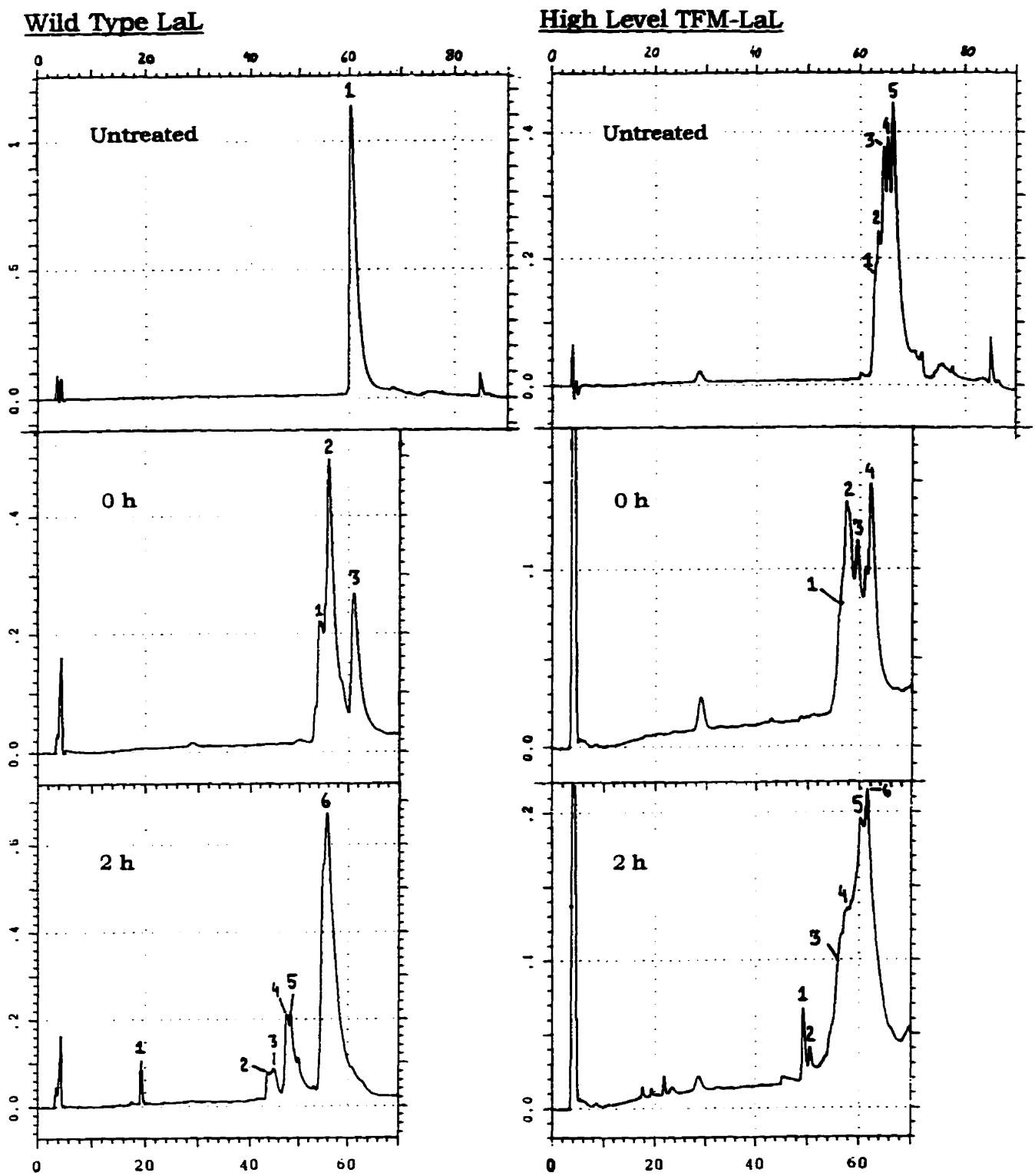

Wild Type LaL
High Level TFM-LaL
Untreated
Untreated
0 h
0 h
2 h
2 h
0
20
40
60
80

Figure 4.15.

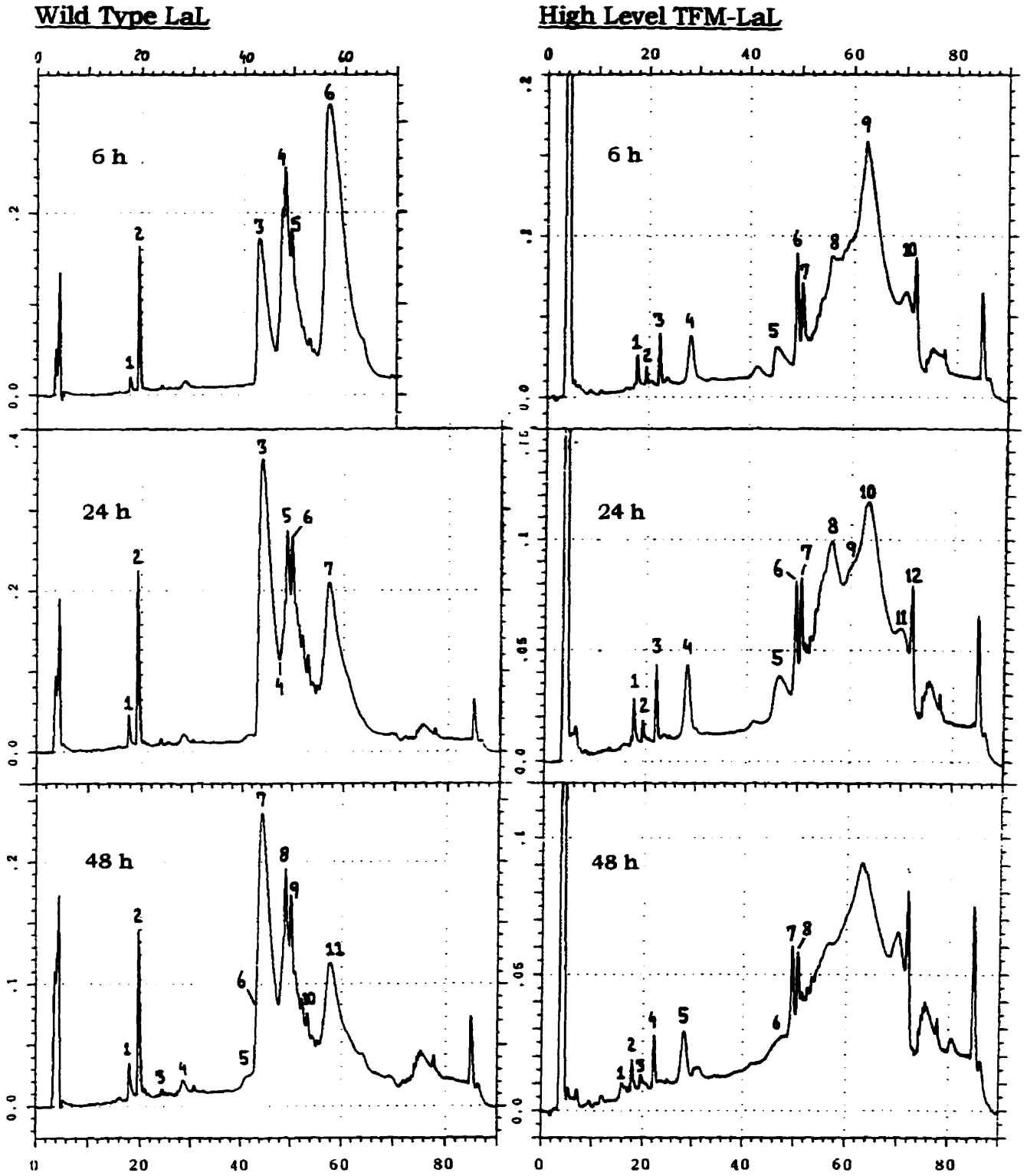


Figure 4.15. (cont'd)

Table 4.7. **Summary of HPLC results of CNBr generated fragments of *wt* LaL. Refer to Figures 4.15 and 4.14 for peak and fragment designations.**

Time (hr)	Peak No.	RT (min)	Fragment Designation	Molecular Weights (Da) Observed from ESMS (*ES*) or MALDI TOF MS (*MAL*)	Degree of Formylation
0	1	54.2	FM	*ES*: 17870[5]	1-2
	2	56.4	FM	*ES*: 17820[5]	
	3	61.3	FM	*ES*: 17825[5]	
2	1	19.3	AM*l*	*ES*: 1660[1]	
	2	43.8	B*l*	*ES*: 10526[1] > 10555 > 10586	0 > 1 > 2
			B*h*	*ES*: 10573[1] > 10602 ≈ 10540	1 > 2 ≈ 0
	3	45.1	inc.	inc.	
	4	47.6	DM*l*	*ES*: 12215[5] >12245	0 > 1
	5	48.6	C	*ES*: 5575[5]	
			DM*l*	*ES*: 12245[5] > 12270 ≈ 12215	1 > 2 ≈ 0
	6	55.2-57.0	FM	*ES*: 17820[5] > 17850	0 > 1
			E	*ES*: 16185[5] ≈ 16160 > 16130 ≈ 16215	2 ≈ 1 > 0 ≈ 3
6	1	18.0	inc.		
	2	19.5	AM*l*	*ES*: 1662[2]	
	3	43.0-44.2	B*l*	*ES*: 10556[1] ≈ 10528 > 10584	1 ≈ 0 > 2
			B*h*	*ES*: 10545[1] > 10573	0 > 1
	4	48.0-49.3	C	ES: 5575[1]	
			DM	MAL: 12275	2
	5	49.8	inc.		
	6	57.2	E	*ES*: 16186[1] ≈ 16158 > 16216 ≈ 16135	2 ≈ 1 > 3 ≈ 0
24	1	17.7	A*l*	*ES*: 1529[1] *MAL*: 1532	
			AM*h*	*MAL*: 1680	
	2	19.4	AM*l*	*ES*: 1660[1] *MAL*: 1660	
			AM*h*	*MAL*: 1707	1
	3	44.1	B*l*	*ES*: 10583[1] ≈ 10555 > 10612 > 10528	2 ≈ 1 > 3 > 0
			B*h*	*ES*: 10600[1] ≈ 10572 > 10629 > 10545	2 ≈ 1 > 3 > 0
				MAL: 10570	1-2
	4	47.2	B	*MAL*: 10625	3-4
	5	49.1	C	*ES*: 5575[1] *MAL*: 5583	
			DM*l*	*ES*: 12272[1] > 12304 ≈ 12247	2 > 3 ≈ 1
	6	50.0	C	*ES*: 5576[1] ≈ 5603 *MAL*: 5589 ≈ 5618	0 ≈ 1
	7	57.3	E	*ES*: 16186[1] ≈16214 >16158 ≈16240 >16270 >16129	2 ≈3 >1 ≈4 >5 >0
				MAL: 16210	
48	1	17.7-18.6	A*l*	*MAL*: 1533	
			AM*h*	*MAL*: 1681	
	2	19.6-20.5	AM*l*	*MAL*: 1658	
			AM*h*	*MAL*: 1705	1
	3	24.5	inc.	inc.	
	4	28.6	inc.	inc.	
	5	40.7-42.5	inc.	inc.	
	6	43.1	B	*MAL*: 10565	1
	7	44.3	B	*MAL*: 10638	4
	8	48.5-49.5	C	*MAL*: 5570	
	9	50.0-50.8	C	*MAL*: 5630	2
	10	53.2-54.0	C	*MAL*: 5664	3
	11	57.1-59.2	E	*MAL*: 16288	5-6
			C	*MAL*: 5700	4-5
			FM	*MAL*: 17955	4-5

Refer to the following page for description of column titles.

The following description of column titles is applicable to Tables 4.7, 4.8, and 4.9.

Time: time at which an aliquot of the LaL or TFM-LaL CNBr incubation was removed, diluted with Milli-Q water and lyophilized.

Peak No.: peak numbering for fractions collected for each respective chromatogram (refer to either Fig. 4.15 and 4.16).

RT: retention time of the peaks collected. A single number represents the medium time of a fraction in which the eluent was collected approximately 0.1 min to either side of the number given. Larger fractions obtained are denoted with the time range over which the eluent was collected.

Fragment designation: those fragments in each peak collected which could be observed and identified from MS analysis. Given are the most probable sources of the fragments. In some cases of TFM-LaL digestions, the data could be attributable to more than one possible fragment and these are designated with an asterisk (*). The designations are those given as outlined in Fig. 4.14. When appropriate, designations are further described with an "*l*" (homoserine lactone) or an "*h*" (homoserine) indicative of the C-terminus. When a particular peak comprises more than one peptide, the fragments are listed in order from major to minor components (as quantitated by MS).

Molecular Weights: Samples were subjected to either ESMS, MALDI TOF MS or both. The superscript value given with the first listed mass determined by ESMS is the resolution (in Da) to which the raw multiply charged data were solved during the MaxEnt reconstruction to produce true molecular weights. In those fragments where formylation was observed, the masses are listed in order of their respective intensities (amounts) observed as either approximately equal to ($\approx$) or greater than (>) the next listed mass.

MALDI-MS TOF spectra were externally calibrated for masses lying in the following ranges: i) using the M+1/M+2 (17826-8914 Da) ions of LaL; ii) using the M+2/M+4 (8914-4457 Da) ions of LaL; or iii) using the M+1 ions of Gramicidin S† and of a synthetic peptide‡ of known sequence and mass (1142.7†-3330‡ Da). All samples were prepared using α-cyano-4-hydroxycinnamic acid as the ionization matrix.

Degree of formylation: the number of formyl groups that correspond with the weights observed in ESMS spectra of fragments exhibiting formylation. The numbers are given in the same order that represent the relative intensity (amount) of the mass they coincide with. For fragment masses determined only by MALDI, the degree is estimated to the most representative value.

inc.: samples in which MS data were inconclusive or of poor quality to afford a fragment designation or molecular weight measurement.

reflected in the generation of peptide C (eg. 24h pk 5-6; 48 h, pk 8-10). Failure to cleave at M14 would result in persistence of peptides D or DM. Only a minor presence of DM*l* was detected in the 24 h sample (pk 5) which was not detected in the 48 h aliquot, suggesting that reaction at M14 went efficiently and that prolonged reaction times result in an enhanced yield. It has been reported that low yields of cleavage have been attributed to the presence of particularly resistant sequences such as Met-Ser or Met-Thr (Titani et al., 1972) or of Met-Cys (Doyen & Lapresle, 1979) however these sequences do not occur in LaL. The nature of the incomplete reaction at M107 as compared to M14 is not clear. It is possible that oxidation to the sulfoxide may have occurred but none of the ESMS analysis of peptides generated from LaL (and of TFM-LaL) gave any indication of oxidation.

More notable was the deficiency of reaction at M1 in *wt* LaL. The majority of the peptides generated that include the N-terminal region retained the N-terminal methionine. As shown in Tables 4.7-4.9, the predominant form of peptides identified that could contain the N-terminal portion of either LaL or TFM-LaL were peptides AM, AM', DM, DM', FM or FM' and not A, D, D', F or F'. This is clearly seen in the 48 h chromatogram of *wt* LaL (Fig. 4.15). There is substantially more of peptide AM (pk 2) than of peptide A (pk 1; there is also a minor contribution of AM*h* in pk 1). The relative proportion of the two peptides remains relatively constant for the early time points as well. The results suggest that the N-terminal methionine reacts slower than do the internal methionine residues. Previous investigations (Inglis & Edman, 1970) have demonstrated that methionine reacts much slower than does N-acetylmethionine with CNBr (as determined by release of methyl thiocyanate) and it was suggested that a free amino group reduces the rate of cyclization of the cyanosulfonium to the iminolactone (Fig. 4.13). It is also possible, as will be discussed later, that the source of fragment A may have only arisen from LaL in which the initiator Met was already absent.

Although 100% fragmentation of *wt* LaL did not occur, the same HPLC and MS analysis of high level TFM-LaL demonstrated that cleavage was substantially reduced. A comparison of the chromatograms obtained for LaL and high level TFM-LaL (Fig. 4.15) illustrate the different profiles observed for each with time. Following 24 and 48 h, mostly all of the digestion products for LaL eluted before 60 min while a substantial portion of the high level TFM-LaL products eluted after 60 min. Although expected products were identified (Table 4.8), the proportion of smaller peptides to larger molecular weight

Table 4.8. Summary of HPLC results of CNBr generated fragments of High level TFM-LaL. Refer to Figures 4.15 and 4.14 for peak and fragment designations.

Time (hr)	Peak No.	RT (min)	Fragment Designation	Molecular Weights (Da) Observed from ESMS (*ES*) or MALDI TOF MS (*MAL*)		Degree of Formylation
0	1	56.0-57.0	inc.			
	2	57.2-58.7	FM,FM'2,3	*ES*: $17820^5 > 17845 > 17935 \approx 17980$		0 > 1
	3	59.7	FM'3	*ES*: 18035^5		2
	4	62.4	FM	*ES*: 17825^5		
2	1	49.3	C	*ES*: 5572^2		
	2	50.6	inc.			
	3	55.7-56.7	inc.			
	4	57.7-58.7	FM, FM'1,2	*ES*: $17825^5 > 17870^5 > 17935^5$		
	5	60.5-61.4	*FM'2	*ES*: $17960^5 > 17985 > 18010 \approx 17930$	* or FM'3	1> 2 > 3 ≈ 0
			or *FM'3	17985 > 18010		0 > 1
			F'2	*ES*: $17830^5 > 17800 \approx 17855$		1 > 0 ≈ 2
	6	61.6	FM'3	*ES*: $17985^5 > 18015$		0 > 1
6	1	17.8		nd		
	2	19.6		nd		
	3	22.2		nd		
	4	28.6		nd		
	5	45.5		nd		
	6	49.1		nd		
	7	50.3		nd		
	8	56.3		nd		
	9	62.5		nd		
	10	71.2		nd		
24	1	17.7	A*l*	*MAL*: 1533		
			AM*h*	*MAL*: 1679		
	2	19.6	AM*l*	*MAL*: 1666		
	3	22.1	AM'*l*	*ES*: 1714^1		
	4	28.5	inc.			
	5	46.6	B	*MAL*: 10626		3-4
	6	49.5	C	*ES*: 5573^1 *MAL*: 5596		
	7	50.6	C	*ES*: $5601^1 \approx 5573 >> 5630$		1 ≈ 0 >> 2
	8	55.2-57.2	inc.			
	9	61.1-62.1	inc.			
	10	63.1-64.1	*FM'3	*ES*: $18014^2 \approx 18042$		1 ≈ 2
			or *FM'2	18014 ≈ 18042		3 ≈ 4
	11	70.5	inc.			
	12	72.0	inc.			
48	1	15.8-17.0	inc.			
	2	17.8-18.3	A*l*	*MAL*: 1531		
	3	19.7-20.5	AM*l*	*MAL*: 1660		
			AM'*h*	*MAL*: 1736		
	4	22.1-23.0	AM'*l*	*MAL*: 1711		
			AM'*h*	*MAL*: 1757		1
	5	28.0-29.5	inc.			
	6	46.3-48.2	inc.			
	7	49.1-50.3	C	*MAL*: 5570		
	8	50.6-51.5	C	*MAL*: 5610		1

fragments was considerably less than with LaL. These fragments most likely result from cleavage at positions within high level TFM-LaL occupied by the residual Met (not TFM) residues.

For example, formation of fragment B involves cleavage at both M14 and M107. In the cleavage of LaL, there is strong production of fragment B which elutes at approx. 44 min (24 h, pk 3; 48 h, pk 6-7). Assuming that TFM is non-reactive towards CNBr and cleavage occurs only at the residual Met positions, then there is only a 7% chance of Met occupying both positions and hence, resulting in the formation of B (based on 74% incorporation into high level TFM-LaL, the probability of both positions 14 and 107 being Met is 0.068, the probability of one being Met and the other TFM is 0.385, and the probability of both being TFM is 0.548). As can be seen in the high level TFM-LaL chromatograms (Fig. 4.15), fragment B was identified in the 24 h sample (pk 5) and by analogy in the 48 h sample (pk 6, although a mass determination was not obtained) however its generation occurred in very low amounts, apparently even lower than 7% and even more significantly, dramatically less than observed for LaL.

Treatment of low level TFM-LaL with CNBr resulted in a degree of fragmentation that was observed to be intermediate to LaL and high level TFM-LaL. The HPLC profiles for the 24 and 48 h low level TFM-LaL reaction aliquots (Fig. 4.16), when compared to the anlogous ones for wt and high level TFM-LaL (Fig. 4.15), make evident that greater fragmentation resulted than with high level TFM-LaL yet was less than that observed for LaL. This is consistent with greater reactivity with CNBr due to the higher proportion of Met to TFM residues in low level TFM-LaL than in high level TFM-LaL. Using once again the example of the formation of peptide B, the 31% incorporation level of low level TFM-LaL would establish a 0.096 probability of both positions 14 and 107 being TFM, a 0.428 probability of one being Met and the other TFM, and a 0.476 probabilitiy of both positions occupied with Met. As is observed in Fig. 4.16, the relative-area (compared to the entire chromatogram area) of the peak for peptide B (48 h, pk 4) does indeed appear to validate the probability of its production. It should be stated that the actual amount of peptide B produced would be less than that predicted statistically as this would assume quantitative cleavage at the methionines. Characterization of the CNBr fragments generated with low level TFM-LaL are presented in Table 4.9.

Figure 4.16.

RP-HPLC chromatograms monitoring the reaction of low level TFM-LaL with CNBr. The reaction contained 56 uM of low level TFM-LaL and 16.8 mM CNBr in 70% formic acid. At the indicated times, aliquots (90 μL, 90 μg for 24 h; 250 μL, 250 μg for 48 h) were removed and dried by lyophilization and chromatographed over a Delta-Pak™ C_{18} column (2.0 mm × 15 cm) as outlined in section 4.2.9.1 using the gradient listed in Table 4.1. The abscissa axis is time (min) and the ordinate axis is absorbance at 215 nm and the ordinate scales are unique for each chromatogram.

The indicated peaks were collected and analyzed by ESMS and/or MALDI-TOF MS and the results are given in Table 4.9. The chromatogram indicated as "untreated" was obtained from chromatography of low level TFM-LaL (100 μg) without exposure to CNBr or to formic acid. Peaks appearing after 74 min result from the gradient employed.

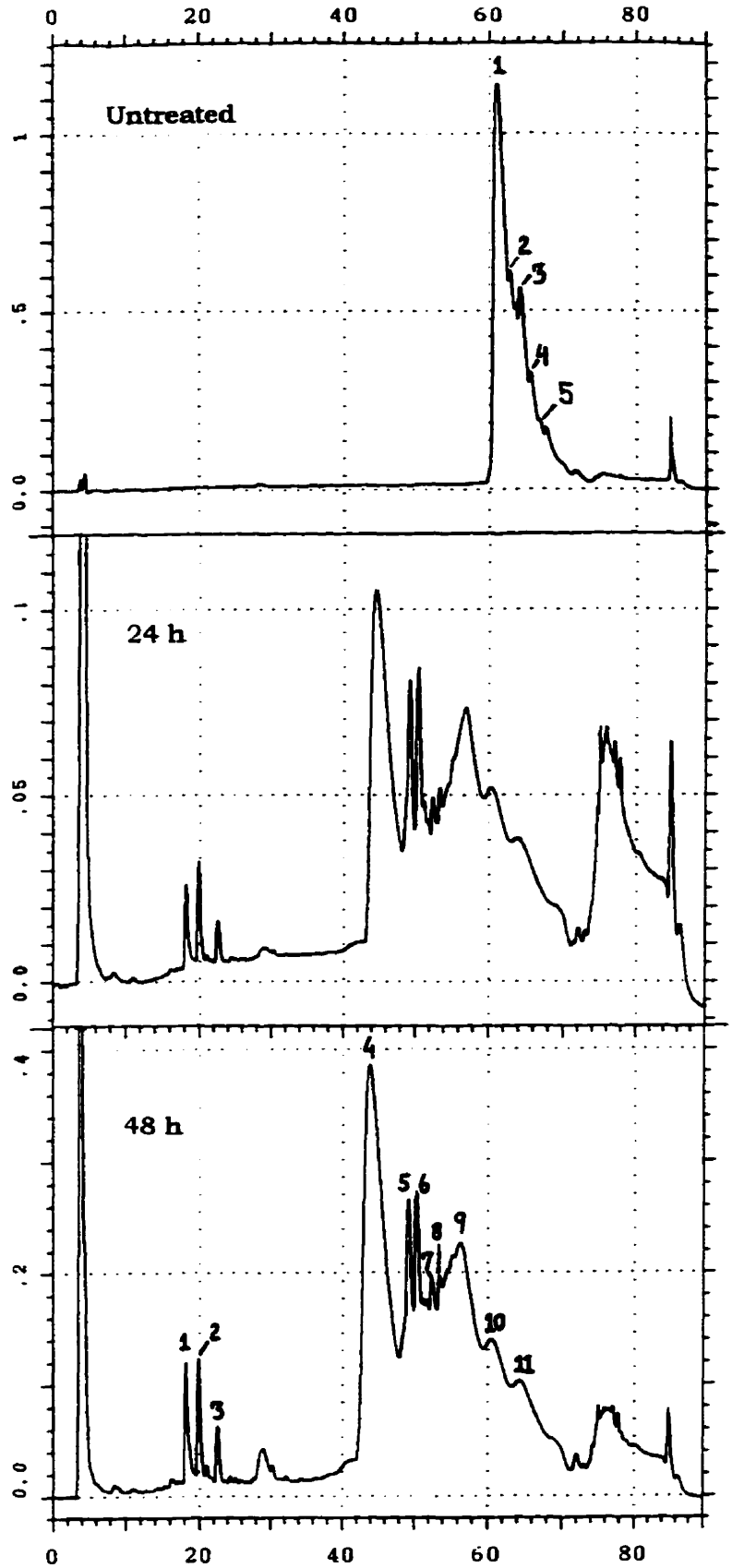
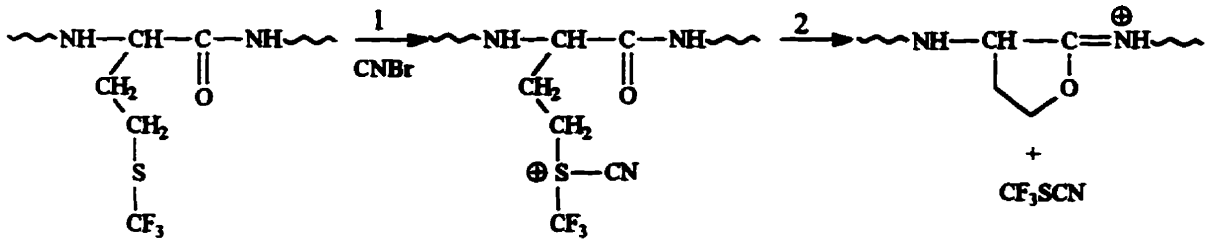


Figure 4.16.

Table 4.9. Summary of HPLC results of CNBr generated fragments of Low level TFM-LaL. Refer to Figures 4.16 and 4.14 for peak and fragment designations.

Time (hr)	Peak No.	RT (min)	Fragment Designation	Molecular Weights (Da) Observed from ESMS (*ES*) or MALDI TOF MS (*MAL*)	Degree of Formylation
48	1	17.8-19.2	A*l*	*MAL*: 1529	
			AM*h*	*MAL*: 1675	
			A*h*	*MAL*: 1575	1
	2	19.8-20.7	AM*l*	*MAL*: 1660	
			AM*h*	*MAL*: 1706	1
	3	22.3-23.2	AM'*l*	*MAL*: 1711	
			AM'*h*	*MAL*: 1758	1
	4	43.0-46.3	B*l*	*ES*: $10611^1 \approx 10583 > 10639 > 10669 \approx 10556$	3 ≈ 2 > 4 > 5 ≈ 1
			B*h*	*ES*: $10598^1 \approx 10627 \approx 10656 \approx 10685 > 10570$	2 ≈ 3 ≈ 4 ≈ 5 > 1
				MAL: 10600	
	5	49.0-50.0	C	*ES*: 5574^1 *MAL*: 5582	
	6	50.3-51.0	C	*ES*: $5602^1 > 5574$ *MAL*: 5613	1 > 0
	7	52.4-53.0	C	*ES*: $5630^1 > 5602$ *MAL*: 5646	2 > 1
	8	53.5-54.0	D'*l*	*ES*: $12255^5 > 12280 \approx 12225 > 12195 \approx 12300$	4 > 5 ≈ 3 > 2 ≈ 6
			*DM'1*l*	*ES*: $12355^5 > 12380 \approx 12410 \approx 12330 > 12445 \approx 12300$	3 > 4 ≈ 5≈ 2 >6≈ 1
			or *DM*l*	$12355^5 > 12380 \approx 12410 \approx 12330 > 12445 \approx 12300$	5 > 6 ≈ 7≈ 4 >8≈ 3
			or *DM'2*l*	$12355^5 > 12380 \approx 12410 \approx 12330 > 12445$	1 > 2 ≈ 3 ≈ 0 > 4
			C	*ES*: $5630^5 \approx 5600 > 5660 > 5690 \approx 5575$ *MAL*: 5663	2 ≈ 1 > 3 > 4 ≈ 0
	9	55.8-57.4	DM'2*l*	*ES*: $12440^5 \approx 12465 \approx 12485 \approx 12410 > 12385 > 12355$	4 ≈ 5 ≈ 6 ≈ 3> 2>1
				MAL: 12416	
			*E	*ES*: $16215^5 \approx 16240 > 16270 \approx 16185 > 16160$	3 ≈ 4 > 5 ≈ 2 > 1
			or *E'	$16215^5 \approx 16240 > 16270 \approx 16185$	1 ≈ 2 > 3 ≈ 0
				MAL: 16270	
			C	*MAL*: 5672	3-4
	10	60.0-61.5	E'	*ES*: $16270^5 > 16240 \approx 16295 \approx 16320 > 16350 > 16210 > 16380$	3 > 2 ≈ 4 ≈ 5 > 6 > 1 ≈ 7
	11	64.0-65.2	E'	*ES*: $16325^5 > 16365 > 16300$	5 > 6 > 4
			FM'3	*ES*: $18040^5 > 18010$	2 > 1

Taken together, the results demonstrate that either TFM containing peptides are resistant to CNBr cleavage or that the cleavage reaction is much slower. We have not conclusively established the chemical basis for this lack of reaction but offer some experimental insight towards this end (see section 4.3.7.1). It is highly probable, however, that the trifluoromethyl group reduces the nucleophilicity (electron density) of the sulfur atom and thus prevents its reaction with CNBr (step 1, Scheme 4.3). If formation of the trifluoromethylcyanosulfonium salt did occur, then the ensuing cyclization reaction (step 2) would be expected to occur since CF_3SCN would be a better leaving group than CH_3SCN. The lesser nucleophilic character of the sulfur atom in ethyl trifluoromethyl sulfide (the side chain of TFM) compared to ethyl methyl sulfide (the side chain of Met) has been predicted by the *ab initio* calculations described previously.



Scheme 4.3.

The detailed analysis performed was done with the goal of specifically separating and identifying peptides produced from TFM-LaL and not from LaL. Unfortunately, the formylation of peptides (discussed in section 4.3.7.2) and the presence of either C-terminal homoserine or lactone resulted in peptide heterogeneity causing peak broadening and peak overlap of fractions eluting from 40-70 min (and to a lesser extent of earlier fractions) which could not be resolved on the column. This heterogeneous nature of fractions collected often gave very complicated raw data (multiply charged) spectra by ESMS that could not be processed to give trustworthy results and only the unambiguous ESMS results are given in Tables 4.7-4.9. Complimentary analysis of fractions by MALDI MS proved to be somewhat more sensitive and forgiving to sample composition but does not offer the same accuracy for large molecular weight peptides as does ESMS. However, MALDI MS results were in excellent agreement with ESMS data obtained for the low molecular weight peptides.

Nonetheless, it was still possible to extract additional information. Peptides containing a TFM had longer retention times than the corresponding Met containing peptide. The fragment DM*l* generated from LaL (Table 4.7 , 24 h, pk 5) eluted at 49.1 min while DM'1*l* and DM'2*l* (Table 4.9, 48 h, pk 8-9) produced from low level TFM-LaL were observed to elute over 53.5-57.4 min. Similarly, fragments AM', E' and FM'3 had longer retention times than AM, E and FM. The results once again demonstrate the increased hydrophobic character imparted to peptides by the trifluoromethyl functionality.

This phenomena is more appropriately demonstrated in the chromatograms of untreated *wt* LaL and high level TFM-LaL. When chromatography of the time 0 samples (Fig. 4.15) resulted in the unusual multiple peak profiles observed (single peaks were expected especially for LaL) samples of untreated *wt* LaL and TFM-LaL were chromatographed to observe how they behaved as well as to establish where "intact" protein eluted on the gradient. Untreated *wt* LaL eluted in a single peak at 60.4 min while untreated high level TFM-LaL eluted in several peaks over 62-68 min (Fig. 4.15)

reminiscent of the partial separation of the component species of TFM-LaL presented previously (refer to Fig. 4.12).

As was mentioned previously, cleavage was limited at position M1 as was exemplified by the greater formation of peptide AM relative to peptide A for the reaction with *wt* LaL. The relative amounts of the analogous peaks in high level TFM-LaL were reversed; there is more of peptide A than of AM (Fig. 4.15, eg. 24 h, pk 1 > 2; 48 h, pk 2 > 3). The remainder of the N-terminal peptide is accounted for by the corresponding TFM containing peptide AM' (24 h, pk 3; 48 h, pk 4). Even for high level TFM-LaL, one would expect to see a greater ratio of AM/A as in *wt* LaL since the ratio would be governed by the degree and lack of reaction at M1. However, high level TFM-LaL is a mixture of proteins, not only in terms of TFM content but also in the absence or presence of the initiator Met (i.e. the N-terminal could be Met, TFM or Val2). The higher proportion of peptide A simply reflects the fact that a greater percentage of high level TFM-LaL has had its initiator Met removed during its biosynthesis. The synthesis of LaL under conditions of methionine starvation (recall that under conditions to produce high level TFM-LaL, there is no exogenous supply of methionine in the medium during induction of protein synthesis) may induce cellular processes to increase Met availability such as the removal of initiator methionines from proteins by aminopeptidases. Low level TFM-LaL is synthesized under less demanding conditions. The presence of comparable amounts of peptide A and AM (Fig. 4.16, 48 h, pk 1 and 2 respectively) illustrates the higher biosynthetic retention of the initiator Met than is seen with high level TFM-LaL. Also, there is a greater relative amount of AM' produced from high level TFM-LaL than with low level TFM-LaL, consistent with a higher incorporation level of TFM in the former. The type of characterization of the N-terminal achieved may lend itself as a convenient and accurate method to examine possible N-terminal heterogeneity resulting from the incorporation of other methionine analogues into LaL.

Analysis of the data has also revealed that a greater extent of peptide fragments bearing a C-terminal homoserine lactone were generated than those bearing a C-terminal homoserine. In certain cases, the two were also separable chromatographically with the lactone exhibiting marginally longer retention times (eg. see Tables 4.7-4.9 for AM*l*, AM*h* and AM'*l*, AM'*h*). Peptides obtained from the CNBr reaction and bearing a C-terminal homoserine lactone have found additional applications. The homoserine lactone ring is sufficiently activated allowing selective coupling of peptides through their C-terminus to amino-polymers (Horn & Laursen, 1973) or with the amino terminal of another peptide and the latter has been shown to be useful in the case of semi-synthetic analogues of cytochrome *c* and trypsin inhibitor (Dykes et al., 1974; Sheppard, 1979).

Other approaches are possible to establish if TFM-containing proteins are susceptible to chemical cleavage by CNBr. A qualitative evaluation can be obtained by SDS-PAGE analysis. Samples from the 48 h reaction time of *wt* LaL and high level TFM-LaL were subjected to electrophoresis using the PhastSystem with both High Density and Homogeneous 20 precast gels (results not shown). Although the results illustrated a lower degree of cleavage for high level TFM-LaL than for LaL, the resolution of the higher molecular peptides and visualization (staining) of the lower ones were poor, and no true quantitative data could be obtained. N-Terminal analyses of samples could also be performed (i.e. a single cycle of EDMAN degradation). The results should show relatively equal amounts of the respective N-terminal residues of the product peptides, namely Met1 (or Val2), Leu15 and Ile108 for *wt* LaL. These amino acids would also be identified for high level TFM-LaL but in lower amounts, and TFM would be the predominating N-terminal residue. Nevertheless, the quality and level of results obtainable from these methodologies would not match that extracted from the HPLC/MS mapping performed.

The course of cyanogen bromide reaction has also been followed by analysis of the CH_3SCN released by gas chromatography (Inglis & Edman, 1970). Detection of CF_3SCN, either by gas chromatography or NMR would give evidence that cyanylation of TFM had resulted (step 1, Scheme 4.3); however the liberation of CF_3SCN itself would not offer explicit evidence for peptide bond cleavage (although it would give evidence of alkylation of $R\text{-}CH_2\text{-}S\text{-}CF_3$). Previous investigations by Carpenter & Shiigi (1974) have demonstrated that a side reaction occurs in which reaction with CNBr results in formation of *O*-peptidyl-homoseryl peptides without cleavage of the peptide (Fig. 4.17). In model peptides approximately 10% of the side reaction (pathway a) competes with the normal reaction (pathway b; Fig. 4.17). Therefore, misleading information about the degree of peptide bond cleavage would be obtained solely from detection of methyl thiocyanate. As well, over 50% of *N*-acetylmethionine undergoes this side reaction yielding *O*-acetylhomoserine (Fig. 4.18). In this instance, the carbonyl oxygen of the acetyl moiety is more nucleophilic than the carbonyl oxygen of the free carboxyl group and is thus more prone to displace the methyl thiocyanate of the intermediate cyanosulfonium despite the formation of the less favored six-membered ring. Subsequently, the authors suggest that in proteins, the degree to which the side reaction takes place may vary with the nature of the amino acids located on both the amino and carboxyl side of the methionine residues. Therefore, the observed incompleteness of reactions of proteins with CNBr could be attributed to the occurrence of the side reaction in which the methionines react with CNBr yet the peptide chain remains intact.

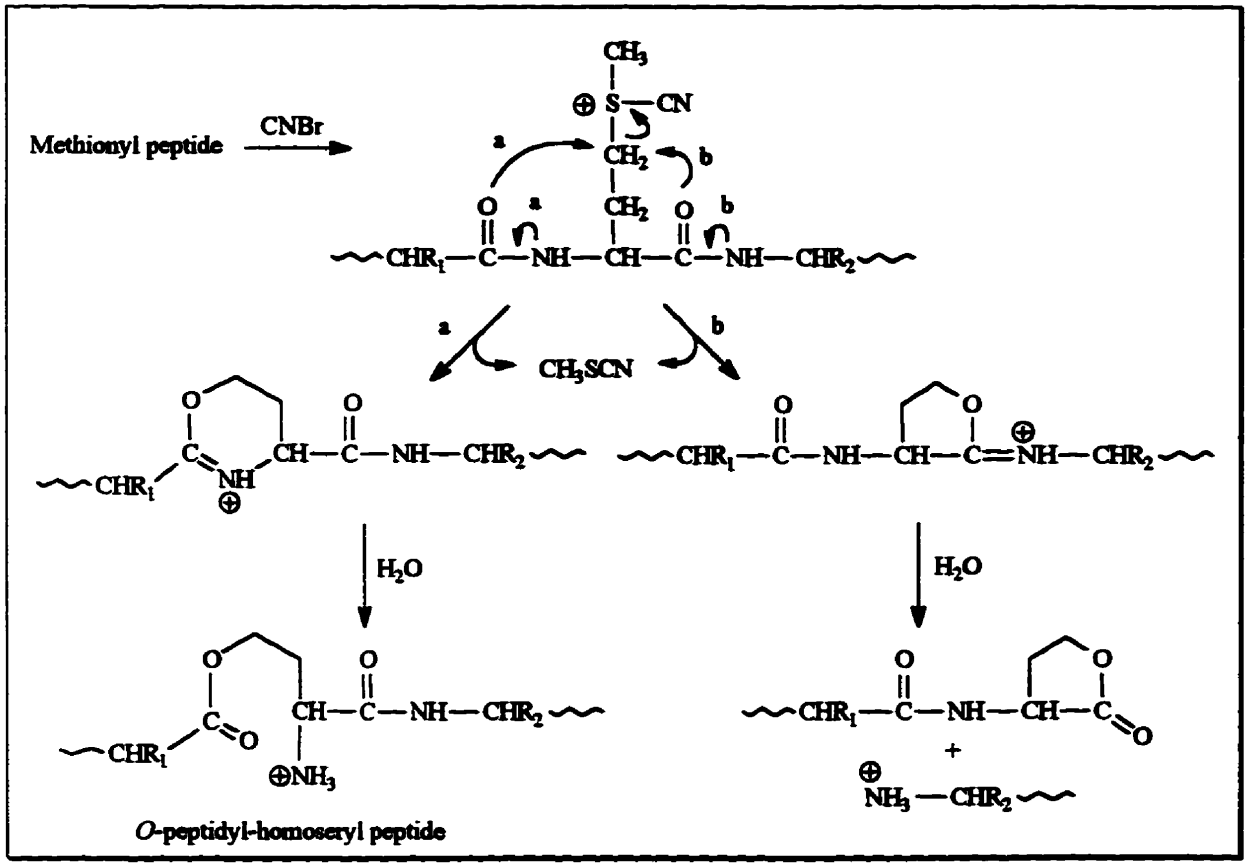


Figure 4.17. Mechanisms of the reaction pathways during treatment of methionyl peptides with CNBr. Elimination of CH_3SCN by normal attack of the methionyl carbonyl oxygen (pathway b) results in peptide cleavage. Attack of the carbonyl oxygen of residue R_1 (pathway a) yields a six-membered dihydrooxazine ring which hydrolyses to give an *O*-peptidyl-homoseryl peptide without cleavage of the peptide (Carpenter & Shiigi, 1974).

4.3.7.1. Effect of CNBr reaction on N-Ac-Met and N-Ac-TFM

To establish whether TFM is susceptible to either the normal or side reaction, both *N*-Ac-*L*-Met and *N*-Ac-*L*-TFM were synthesized and their reaction with CNBr investigated. The acetylated amino acids were prepared in approx. 40% yield from the free amino acids in the presence of aqueous $NaHCO_3$ and Ac_2O/MeOH mixture (Dell, 1986). Typical acetylation conditions (Ac_2O/Et_3N/DMF or AcCl/Py/CH_2Cl_2) were attempted with *L*-Met but resulted in very low yields of *N*-Ac-*L*-Met and generated a mixture of products in which it was deemed difficult to isolate the desired acetylated product. In these cases, ^1H NMR analysis of the crude products suggested that perhaps *S*-methyl-*N*-acetyl-homocysteine thiolactone was being produced. This could result if the acetylating agent

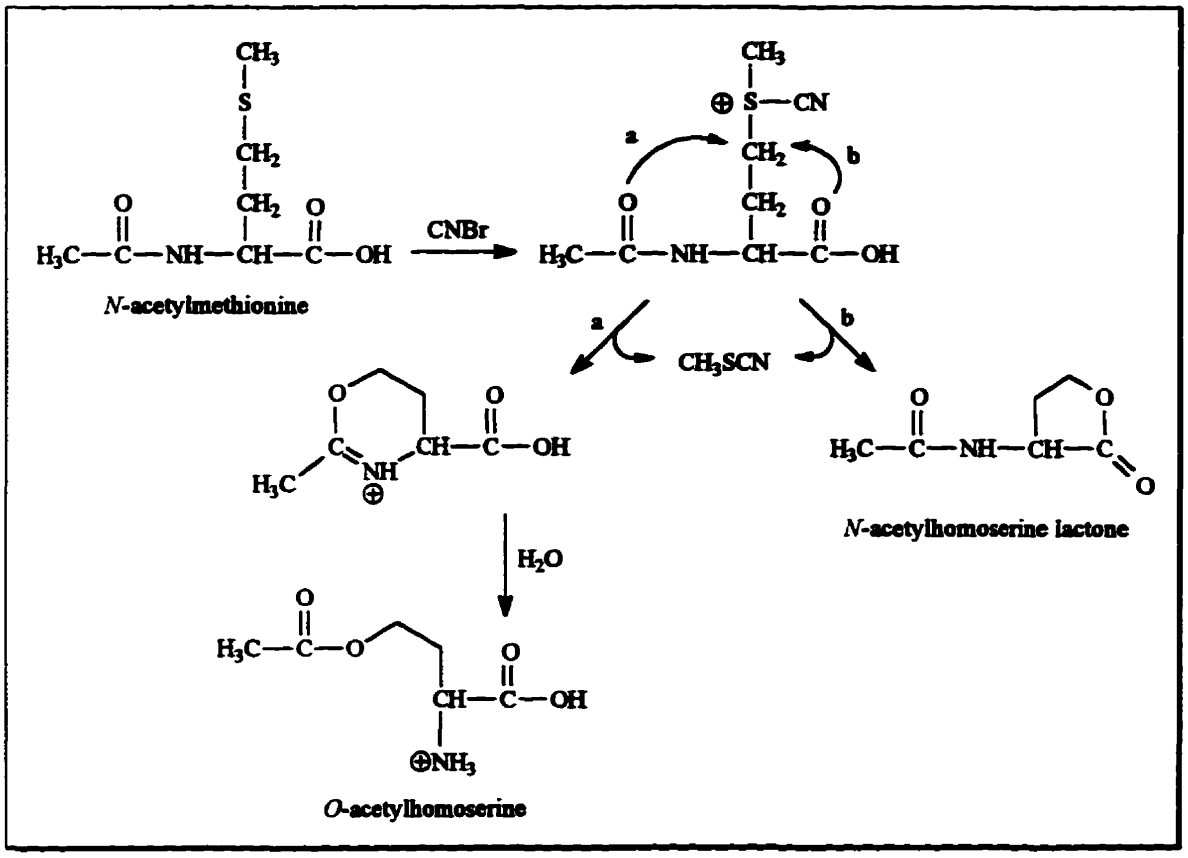


Figure 4.18 Diagram of the reaction pathways during treatment of *N*-acetylmethionine with CNBr. The normal pathway (b) results in formation of *N*-acetylhomoserine lactone while the alternative reaction (a) results in formation of *O*-acetylhomoserine (Carpenter & Shiigi, 1974).

formed a mixed anhydride with the carboxylic acid of Met with subsequent cyclization resulting from attack by the sulfur atom of the side chain onto the methionyl carbonyl.

The reaction of *N*-Ac-Met with CNBr will yield both *N*-acetylhomoserine lactone and *O*-acetylhomoserine as outlined in Fig. 4.18 and the products are both produced in approximately equal amounts (Carpenter & Shiigi, 1974). Both *N*-Ac-Met and *N*-Ac-TFM (4 mM) were subjected to CNBr (90 mM) treatment under standard conditions. Over the course of the reactions, the amount of unreacted or remaining starting material was assessed quantitatively by HPLC analysis and compared to control reactions in which CNBr was omitted.

The analysis of the CNBr reaction with *N*-Ac-Met indicated that there was complete disappearance of the starting material and conversion to the expected products. The chromatograms of the treated sample clearly indicate the loss of the peak corresponding to *N*-Ac-Met (R_T 20.1 min) and the emergence of two new peaks, *a* and *b* (Fig. 4.19 A) whose respective areas increased with time (Table 4.10). Integration of the peak corresponding to

Figure 4.19.

RP-HPLC chromatograms monitoring the reaction of *N*-Ac-Met and *N*-Ac-TFM with CNBr. Samples contained 4 mM of the amino acid in the presence (+ CNBr, 90 mM) or absence (- CNBr) of CNBr in 70% formic acid. Chromatographic conditions are given in section 4.2.9.3 using the gradient listed in Table 4.2. The abscissa axis is time (min) and the ordinate axis is absorbance at 210 nm and the ordinate scales are equal for each chromatogram.

(A) Chromatograms for the reaction of *N*-Ac-Met and the Reagent Control (90 mM CNBr/70% formic acid only) at times of 1 and 24 h.

(B) Chromatograms for the reaction of *N*-Ac-TFM at times of 1 and 48 h.

Peaks indicated with an asterisk (*) are believed to arise from residual reagents as observed in the reagent control. Integration data for selected peaks are listed in Table 4.10.

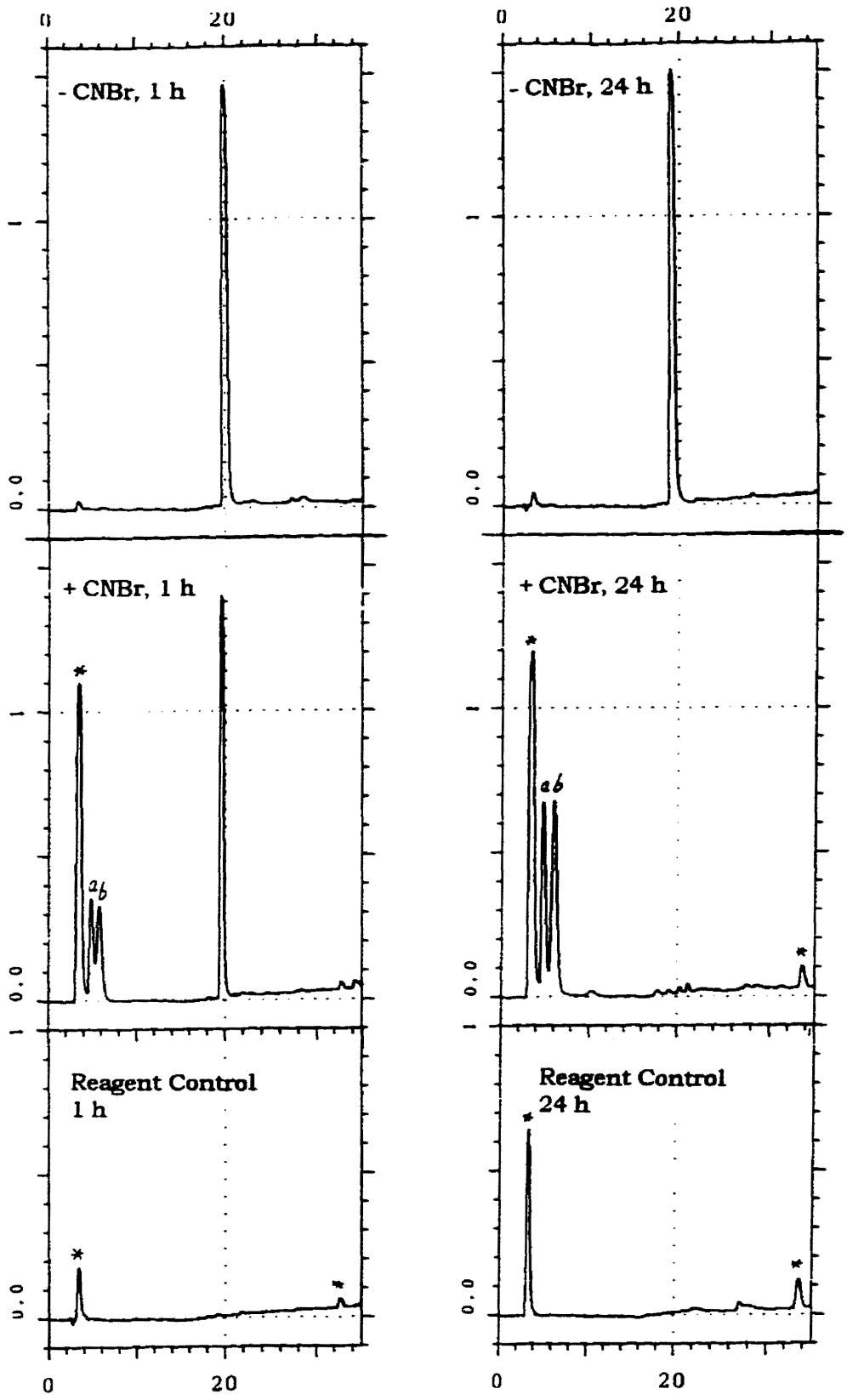

0
20
- CNBr, 1 h
- CNBr, 24 h
+ CNBr, 1 h
+ CNBr, 24 h
a b
Reagent Control
1 h
Reagent Control
24 h
0.0
1

Figure 4.19 (A).

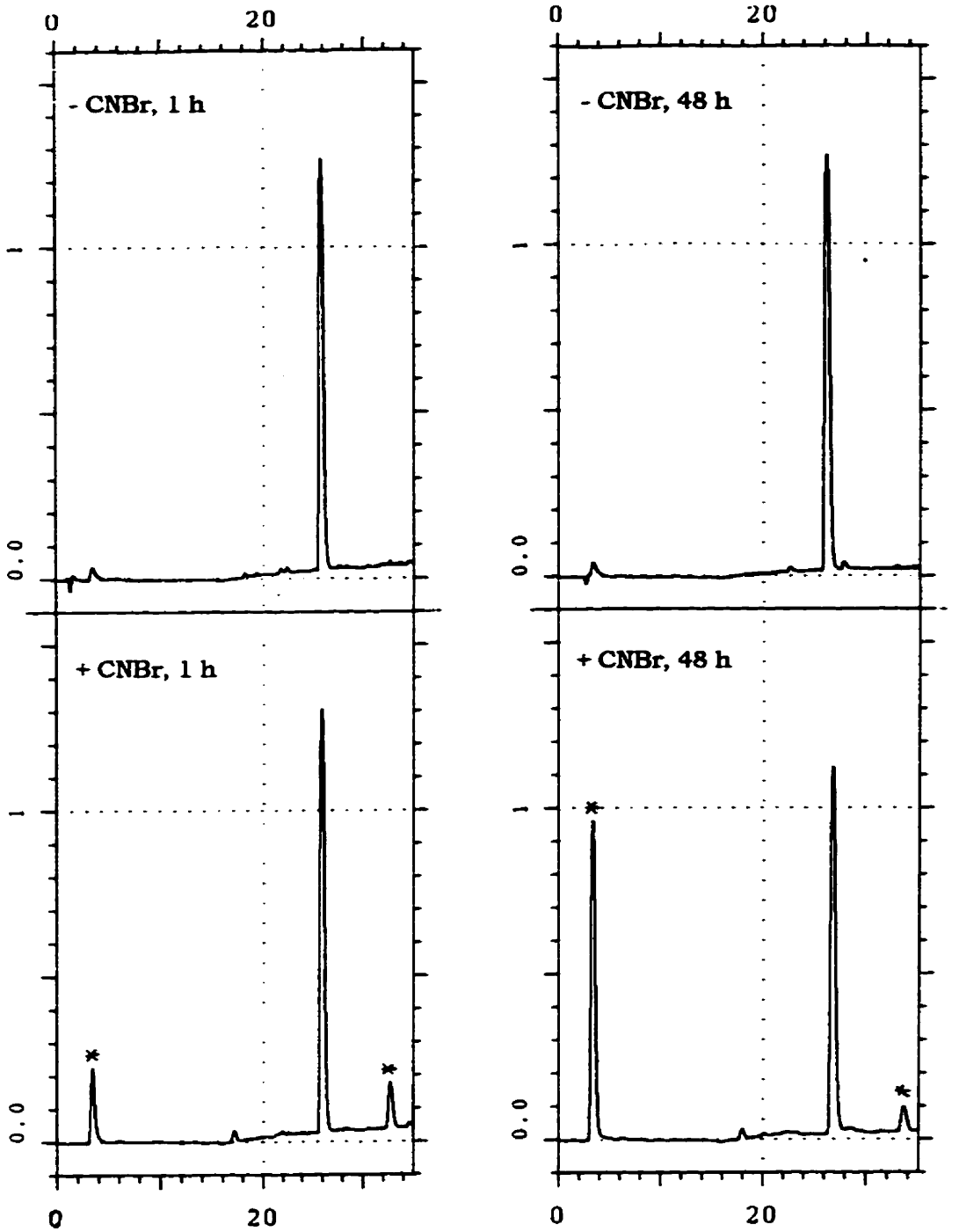

0
20
- CNBr, 1 h
1
0.0
0
20
- CNBr, 48 h
1
0.0
+ CNBr, 1 h
1
0.0
0
20
+ CNBr, 48 h
1
0.0
0
20

Figure 4.19 (B).

Table 4.10. Integration data for the reactions of *N*-Ac-Met and *N*-Ac-TFM with CNBr.

Sample[†]		R_T[‡] (min)	Area (AU × min) 1 h	24 h	48 h
N-Ac-Met	+ CNBr	20.1 ± 0.5	0.71	0.03	0.02
	- CNBr	"	1.08	1.04	1.07
	peak *a*	4.9 ± 0.04	0.20	0.37	0.31
	peak *b*	5.9 ± 0.1	0.23	0.47	0.46
N-Ac-TFM	+ CNBr	25.9 ± 0.2	0.68	0.66	0.65
	- CNBr	"	0.65	0.67	0.70

[†]Refer to Fig. 4.19 (A) and (B) for peak designations. [‡] The average retention time of the appropriate peak.

N-Ac-Met revealed approximately 35% of *N*-Ac-Met had reacted after 1 h and over 97% was consumed following 24 h and 48 h (Table 4.10). Similar studies have shown over 95% loss of *N*-Ac-Met following 4 h of reaction (Carpenter & Shiigi, 1974). No loss of *N*-Ac-Met was observed in the control untreated sample (Table 4.10).

Peaks *a* and *b* were collected and were virtually indistinguishable by ESMS analysis. Both were identified to contain species with masses corresponding to those of *N*-Ac-homoserine lactone and *O*-Ac-homoserine confirming that the expected products described in Fig. 4.18 were produced (no other significant peaks were observed in the spectra over the acquired mass range of 100-200 Da). The ESMS spectra of peak *a* typically indicated a higher intensity for the mass corresponding to *O*-Ac-homoserine (161.07 Da) while the spectra for peak *b* indicated a higher intensity for the mass corresponding to *N*-Ac-homoserine lactone (143.06 Da) although the relative amounts of each varied somewhat between repeated acquisitions.

It is possible that *N*-Ac-homoserine (with identical mass to *O*-Ac-homoserine) was also generated either through hydrolysis of *N*-Ac-homoserine lactone or from transfer of the acetyl group (*O*→*N* acyl shift) from the hydroxyl to the α-amino of *O*-Ac-homoserine (although unlikely under the acidic conditions). Therefore fractions *a* and *b* were subjected to TLC (silica; 1-butanol:H_2O:acetic acid, 4:1:1) and revealed a ninhydrin positive spot (R_f = 0.30). As a result, a free amino group must be present and supports generation of *O*-Ac-homoserine. As well, TLC (silica; hexane:EtOAc, 3:1) of fractions *a* and *b* revealed a 254 nm absorbing species (R_f = 0.41) indicative of *N*-Ac-homoserine lactone and a ninhydrin reactive species (R_f = 0.0) again, an indication of *O*-Ac-homoserine.

It is not clear regarding the weak interaction of the products with the Delta Pak™ C_{18} column or as to why they both elute in two separate fractions when compared to the elution pattern for *N*-Ac-Met. The Delta Pak™ column is composed of support particles with 300 Å pore size, which may be too large for efficient interactions with *O*-Ac-homoserine and *N*-Ac-homoserine lactone especially if they are compact molecules. As well, experience with various columns has often revealed that single species will elute in more than one fraction from either lack of uniformity or deterioration of the packing material or the involvement of two or more physical separation modes operating during chromatography with a given column. Regardless of the aforementioned anomalies or as to the composition of the reaction products, the analysis was suitable to monitor and clearly indicated near quantitative loss of *N*-Ac-Met during its reaction with CNBr.

The peaks indicated on the chromatograms with an asterisk (*) were also present in the reagent control samples which simply contained 90 mM CNBr in 70% formic acid (Fig. 4.19 A, bottom panels). Although CNBr and formic acid are volatile and should be removed during lyophilization of samples, the presence of these peaks in the reagent control samples could indicate their incomplete removal or perhaps an impurity in the CNBr used. It is expected that both CNBr and formic acid would elute in the void volume of the column which appears to take place. The intensity of these peaks varied in the reagent control samples and in the *N*-Ac-Met and *N*-Ac-TFM treated samples more than likely due to differences in their efficiencies of removal during lyophilization or to their "trapping" in samples in which amino acids were present. Analysis of these (*) fractions by ESMS and TLC did not give any indication as to their origin.

The analysis of the reaction of *N*-Ac-TFM with CNBr demonstrated that it was not consumed during its reaction. Even following a 48 h reaction period, the chromatograms of the treated sample (Fig. 4.19 B) and integration of the *N*-Ac-TFM peak (Table 4.10) revealed that the amount of remaining *N*-Ac-TFM was the same as to the control untreated sample and did not vary with time (within the error of integration). The results therefore indicate that *N*-Ac-TFM does not react with CNBr to enter either the normal or side reaction. Subsequently, the lack of fragmentation observed with high level TFM-LaL cannot be attributed to the side reaction in which reaction with CNBr occurs without cleavage of the protein. Taken together with the results of treatment of TFM-LaL, it can be concluded that the presence of the trifluoromethyl moiety renders TFM resistant to reaction with CNBr in comparison to the reaction that occurs with methionyl residues. As mentioned previously, the lack of reaction more than likely stems from lack of initial cyanylation of the sulfur atom (step 1, Scheme 4.3).

4.3.7.2. Formylation of CNBr Reaction Products

As mentioned previously, chromatography of the CNBr treated protein samples tended to give broad peaks, unusual for the Delta-Pak C_{18} micro-bore column used which from experience, generally gives sharp peaks. Initial analysis of peptides gave ESMS spectra showing multiple adducts of 28 Da, consistent with formylation of the peptides (i.e. HCO_2H + H-X-peptide → H_2O + HCO-X-peptide would result in an increase of 28.01 mass units to the peptide). The effect of formylation is clearly demonstrated in the ESMS spectra of fragment B (Fig. 4.20). The masses and the intensities of the peaks observed indicate that the degree of formylation of both the peptide with a C-terminal homoserine lactone (B*l*) and C-terminal homoserine (B*h*) has occurred to similar extents and the amounts in each formylation "series" appear to follow a Gaussian (normal) distribution. As mentioned previously, the spectra also indicates the generation of a greater amount of a homoserine lactone termini than of a free homoserine.

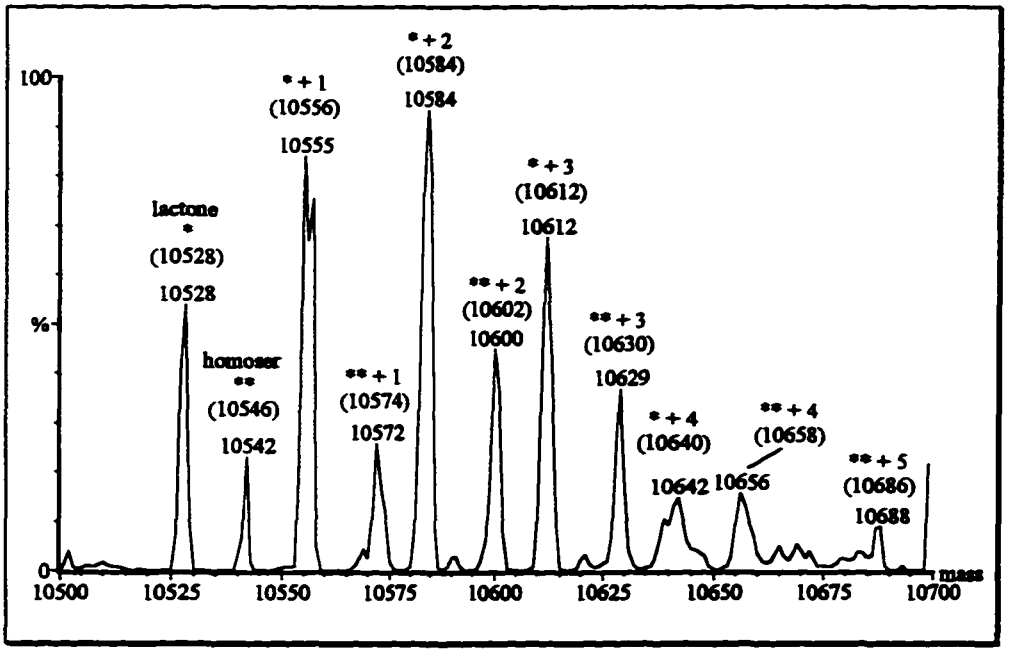


Figure 4.20. ESMS spectrum of CNBr generated fragment B demonstrating the extent of formylation. Fragments derived from the parent homoserine lactone are designated with " * " and from the parent homoserine with " ** " and the numbers given following the + sign represent the total amount of formyl groups associated with each mass peak. Shown are the observed masses and the calculated masses are given in parantheses. The spectrum was resolved to a resolution of 1 Da. The sample was obtained from HPLC purification of an independent reaction of CNBr with *wt* LaL under standard conditions.

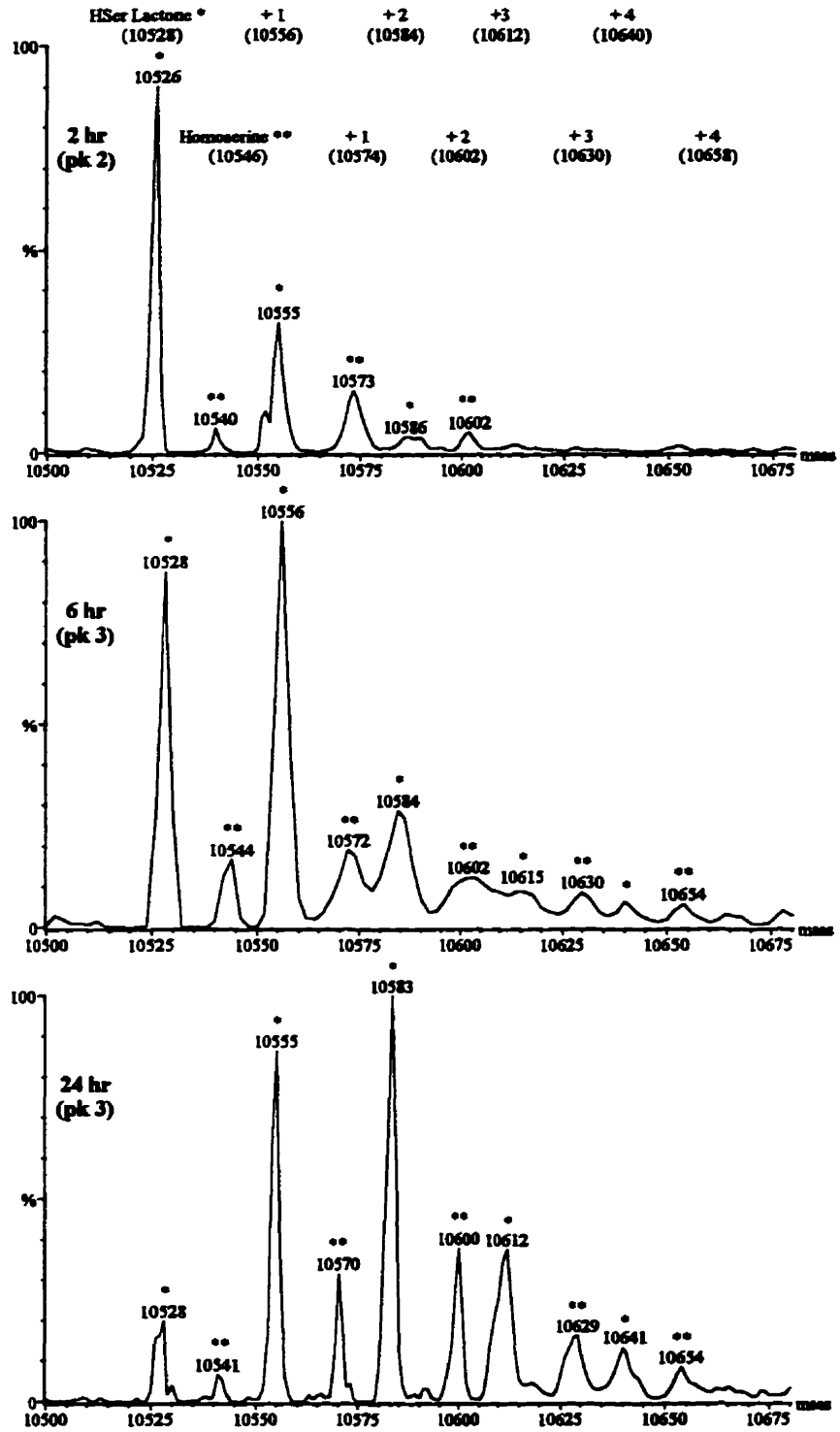


Figure 4.21. ESMS spectra for fragment B obtained from *wt* LaL illustrating the higher degree of formylation with time. The 2 and 24 hr spectra were resolved to a resolution of 1 Da and the 6 hr spectra to 2 Da. Shown at top of the 2 hr pannel are the calculatated masses for the homoserine lactone (*) and homoserine (**) peak series. The numbers given following the + sign represent the total amount of formyl groups associated with each mass peak.

There was also a dependence of the extent of formylation with the duration of the reaction. The ESMS spectra of peptide B reveals that with increasing reaction times the distribution of peaks has shifted to higher molecular weight (Fig. 4.21). This was observed for essentially all peptides exhibiting formylation. An examination of the data presented in Tables 4.7 and 4.8 will indicate that corresponding peptides isolated from the longer reaction samples tended to show stronger modification.

Since the CNBr cleavage reactions were performed in 70% formic acid, the reaction solvent must act as the source of the formyl groups. Anomalous reactions resulting in the blockage of the N-terminal amino acid of peptides and proteins treated with CNBr in formic acid have been reported but the nature of the blocked products have not been fully characterized (Shively et al., 1982; Richard et al., 1984). Formylation of peptides following CNBr/formic acid treatment has been suggested on the basis of FAB mass spectral analysis (Fukuhara et al., 1985; Goodlett et al., 1990) and aberrant chromatographic behaviour (Tarr & Crabb, 1983) of the product peptides. Although a reactive serine or threonine hydroxyl has been suggested to be the site of formylation (Tarr & Crabb, 1983; Goodlett et al., 1990), the supporting evidence is conjectural. Consequently, the nature as to which group(s) on peptides are acting as the acceptor sites for formylation has not been clearly defined.

Due to the low pH of the reaction, it is unlikely that amides of ε-amino, N-terminal amino or imidazole groups were forming from the CNBr/formic acid reaction with LaL. Instead, the evidence suggests that the most likely modification of peptides resulted from the acid catalyzed esterification between formic acid and hydroxyl groups on the peptides. Table 4.11 lists the amino acid composition of the CNBr generated fragments and it became apparent that peptides containing any portion from L15-V158 underwent formylation. All of the Ser, Thr and Tyr residues of LaL are located in this region and an experiment was performed in which each of these free amino acids were subjected to standard CNBr conditions. Analysis of the reactions by high resolution ESMS (Table 4.12) revealed that the spectra for Ser and Thr gave not only the peaks for the parent amino acid, but also a peak at plus 28 Da, evidence of formylation. The spectra for Tyr did not give any indication of formylation. This implies not only that the phenolic hydroxyl group is unreactive under these conditions but also that indeed the α-amino does not become formylated. In light of these results, it is evident that formylation of the side chains of Ser and Thr will occur in formic acid. Treatment in the presence of CNBr did not have any

Table 4.11. Amino acid composition of CNBr generated fragments of LaL (as predicted from the primary sequence).

Amino Acid	Met1-Met14 (AM)†	Leu15-Met107 (B)†	Ile108-Val158 (C)†
Ala	1	6	6
Arg	1	7	4
Asn	2	3	1
Asp	1	7	5
Cys	0	0	1
Gln	1	6	2
Glu	1	3	3
Gly	0	9	6
His	0	2	1
Ile	1	2	6
Leu	1	11	2
Lys	1	8	3
Met	1	0	0
Phe	1	2	2
Pro	0	4	1
Ser	**0**	**6**	**3**
Thr	**0**	**5**	**1**
Trp	0	3	1
Tyr	0	4	1
Val	1	4	2
Met*	1	1	

Met* denotes a C-terminal homoserine or homoserine lactone.
† Fragment designations are given in Fig. 4.14.

Table 4.12. ESMS data for Ser, Thr and Tyr treated under various conditions.

Amino Acid	Condition†	Observed Masses	Relative Intensity (%)‡
Ser	A	104.98	100
	B	104.98; 132.97	83; 17
	C	104.90; 132.93	85; 15
Thr	A	119.01	100
	B	118.98; 146.97	95; 5
	C	118.96; 146.96	95; 5
Tyr	A	181.06	100
	B	181.05	100
	C	181.03	100

† Amino acids were (A) untreated, (B) treated in 70% formic acid only or (C) treated in 70% formic acid in the presence of CNBr (0.1 M amino acid, without (A) or with (B)1.0 M CNBr, 24 h reaction). ‡ Determined from the areas of the respective mass peaks observed in the ESMS spectra.

influence as to the amount of formylation suggesting that CNBr does not catalyze the esterification. As well, a greater extent of formylation ensued with Ser (15-17%) than with Thr (5%), consistent with the higher assumed reactivity of the 1° hydroxyl of Ser to the 2° alcohol of Thr (Table 4.12).

The analysis of the HPLC fractions of the CNBr treated protein samples also lends support to the modification of Ser and Thr residues. Fragment B contains a total of 6 Ser and 5 Thr residues, and species representative of the addition of up to 5 formyl groups were detected in the spectrum (Fig. 4.20). In contrast there are 9 Ser and 6 Thr residues present in fragment E', and in agreement with a higher occurrence of the residues, a total of 7 modifications were observed (Fig. 4.22). Fragment C has only 3 Ser and 1 Thr and the spectrum reveals the presence of all species possible including one with 4 formyl group additions (Fig. 4.22). It also appears that the number of formyl groups detected more closely coincides with the number of Ser residues present in the peptide, again suggesting that formylation of Ser proceeds more readily than with Thr. Figure 4.22 was also included to demonstrate that the existence of 2 series of peaks for peptides ending in a homoserine lactone or homoserine was not artifactual (see Figs. 4.20 and 4.21 which demonstrate the two series of peaks). Both peptides C and E' terminate with Val, and since there is no C-terminal heterogeneity, only a single series is seen for these fragments (Fig. 4.22).

Further support of the formylation of hydroxyl species is evident from analysis of the N-terminal fragments. Peptide A (or AM) contains no Ser or Thr, and formylation of these peptides occurs only when the peptide ends in homoserine (and not homoserine lactone). Since only a single hydroxyl group is present in these peptides, only masses consistent with a single formylation of peptides are observed (see Tables 4.7-4.9 for peptides A*h*, AM*h* and AM'*h*).

Formylated peptides were also observed to have increased retention times when chromatographed over the C-18 column. For example, peptide AM*h* (formylated) eluted approximately 1-2 min later than the corresponding unformylated peptide. As mentioned previously, peak broadening was associated with formylation. Fragment C was observed to elute from 49.0-57.4 min (pk 5-9, Table 4.9) and the peptide eluting in the later fractions had a higher degree of formylation than the earlier ones. The addition of the formyl group(s) therefore increases the hydrophobic character of the peptides.

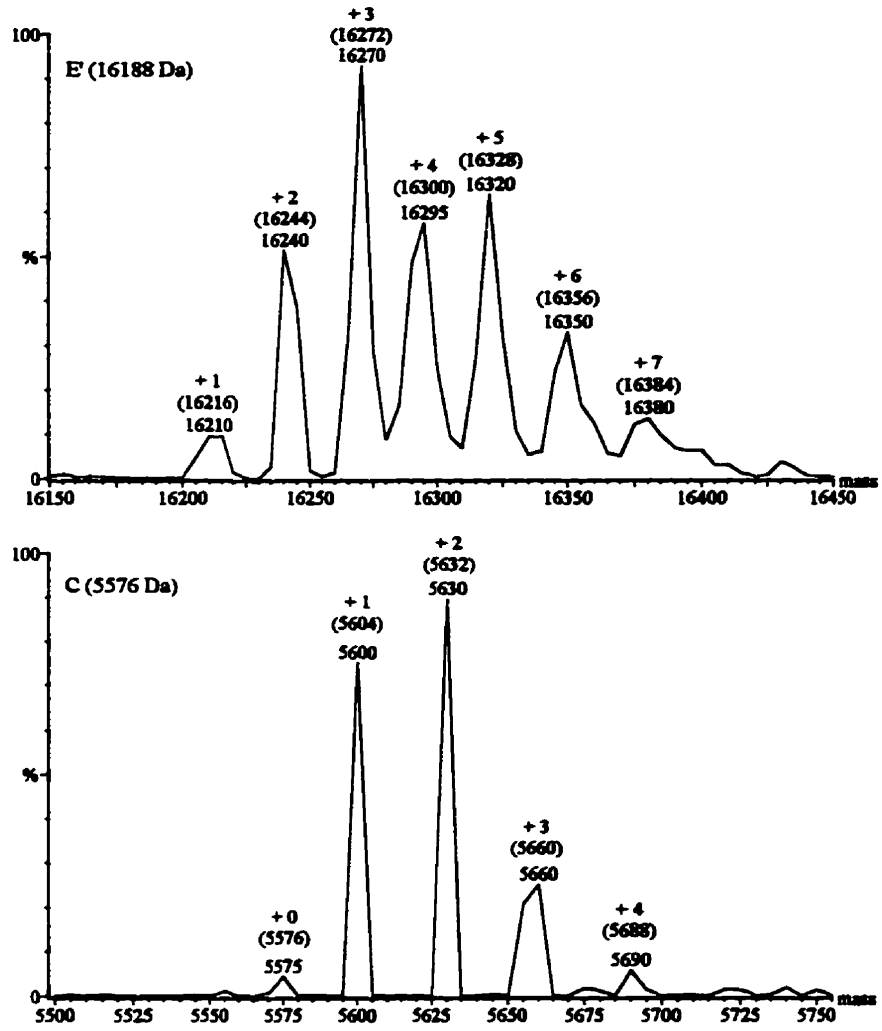


Figure 4.22. ESMS spectra of CNBr generated fragments E' and C. The spectra were obtained from peak 10 and peak 8 (see Fig. 4.16 and Table 4.9) for E' and C respectively from low level TFM-LaL. The numbers given following the + sign represent the total amount of formyl groups associated with each mass peak. Shown are the observed masses and the calculated masses are given in parantheses. The spectra were resolved to a resolution of 5 Da.

What is not clear is the peculiar elution pattern observed for the time 0 samples of *wt* LaL and high level TFM-LaL (Fig. 4.15). One would expect that the short exposure of these samples to the formic acid (prior to dilution and lyophilization) would result in minimal formylation and therefore, the samples should behave chromatographically similar as the untreated samples. This was not observed and no straightforward explanation concerning the elution profile or to the mass analysis results obtained for the

fractions collected can be offered. In a separate experiment, LaL was treated in neat 70% formic acid (no CNBr) for 24 h. Chromatography of this sample gave a single peak with a similar retention time but of a broader nature than for untreated LaL. The ESMS spectra of this sample (Fig. 4.23) also demonstrates that the protein became formylated with evidence of species with up to 6 modifications detected. However, the majority of the protein remained unmodified and a lesser extent of formylation was observed than with the CNBr generated peptides. This may suggest a decreased exposure of the reactive Ser and Thr residues of the protein which become more accessible in peptides resulting from CNBr fragmentation.

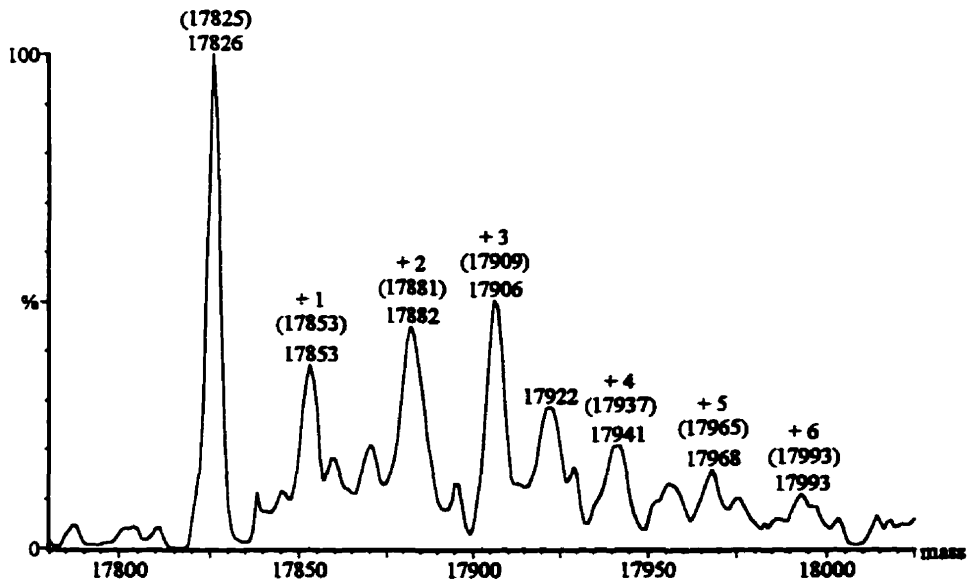


Figure 4.23. ESMS spectrum of LaL treated in 70% formic acid for 24 h illustrating its formylation. The numbers given following the + sign represent the total amount of formyl groups associated with each mass peak. Shown are the observed masses and the calculated masses are given in parantheses. The spectra was resolved to a resolution of 1 Da.

Formic acid is conventionally used as a reaction medium for the CNBr treatment of proteins. The evidence presented strongly suggests the hydroxyl groups of Ser, and to a lesser extent Thr in addition to the hydroxyl group of C-terminal generated homoserine, can become formylated. It is noteworthy to caution others that this phenomena might interfere with the further analyses of peptides generated from the intended fragmentation of a protein with CNBr. These concerns have been expressed by others. Formyl esterification of serine and threonine was the presumed cause of HPLC peak broadening

and decreased resolution of peptides obtained from the cleavage of proteins by CNBr in formic acid and these effects were shown to be reversible by treatment of the peptides with aminoethanol (Tarr & Crabb, 1983). In addition to formic acid, trifluoroacetic acid (Lee & Shively, 1990) and 0.1 N HCl (Spande et al., 1970) are other mediums that have been applied for the CNBr reaction with proteins. However, trifluoroacetylation of peptides has also been observed by FAB mass spectrometry but at a lower abundance than formylation when the reactions were performed in TFA or formic acid respectively (Goodlett et al., 1990). Substitution of formic acid with acetic acid, TFA or HCl has been reported to prevent blocking of the N-terminal of peptides observed when formic acid is used for CNBr cleavage (Shively et al., 1982). Another report suggests that CNBr reactions conducted in the presence of HCl and not formic acid are cleaner in terms of the products produced (Richard et al., 1984). Our results appear to be the most detailed to date on side reactions produced by CNBr cleavage of proteins in formic acid and have unambiguously defined specific sites of chemical modification.

4.3.8. Summary and Future Work

In our ongoing involvement with methionine biochemistry, we have pursued the incorporation of *L*-TFM into bacteriophage lambda lysozyme. This account constitutes the first characterization of an enzyme into which TFM has replaced methionine residues. There has previously been only indirect evidence of TFM incorporation into TCA-insoluble protein fractions from *S. cerevisiae* (Colombani et al., 1975) and no detailed characterization of a TFM-containing protein, nor its ^{19}F NMR spectrum has been reported to date. The partial sequencing of the N-terminus of TFM-LaL has indeed established that Met1 and Met14 have been substituted and taken together with analysis of TFM-LaL by ESMS and ^{19}F NMR is a clear indication that incorporation has occurred at each of the three Met positions (positions 1, 14 and 107) in LaL.

The trifluoromethylmercury moiety has been selectively introduced into cysteine residues of actin and myosin using (trifluoromethyl)mercuric bromide and ^{19}F NMR data on the chemically introduced protein label was presented (Barden et al., 1989). However, having prepared an effective system for expression of LaL directed by the methionine auxotroph B834(DE3) harboring pLR102, we were able to establish biosynthetic conditions to significantly increase and control the substitution of the methionines in LaL with TFM. We have prepared TFM-labelled LaL exhibiting wild type activity at high and low incorporation levels. High levels of incorporation are obtained by induction of cells in the presence of 1.0 mM *L*-TFM but is accompanied by low protein expression. Lower levels of incorporation with concomitant increase in protein yields are obtained when protein expression is conducted in media containing both *L*-TFM (1.0 mM) and *L*-Met (20 μM). In the application of *L*-TFM as a ^{19}F NMR probe for protein structure, it may be advantageous however to explore the effect of differential levels of incorporation of TFM into the particular protein of interest. The differences observed in the ^{19}F NMR spectra between high and low level TFM-LaL suggest that the structure of LaL is sensitive to the extent of TFM incorporation. Consequently, lower levels of incorporation may be more suitable to the study of less perturbed protein structures.

The noted differences predicted by the calculations on ethyl methyl sulfide and ethly trifluoromethyl sulfide suggest that some changes in enzyme function could result from the substitution of Met by TFM. Although the TFM labelled enzymes displayed similar activity to the wild type, we have noted in this study that the more extended sugar analogues ($(GlcNAc)_{4-6}$) may bind somewhat differently to TFM-LaL as indicated by their slightly reduced inhibitory efficiency. This effect may be the result of direct interactions of the larger CF_3-moiety with the saccharide. Alternatively, this effect could be due to

somewhat longer range perturbations of TFM residues that are not directly in contact with the sugar residues.

To date, the incorporation of fluorinated amino acids into proteins has focused primarily on aromatic amino acid analogues. We have described conditions to control the incorporation of TFM into a protein retaining full activity, and thereby make available a new non-aromatic amino acid probe to study protein structure and function. We have presented experimental evidence that labelling of LaL with TFM can be used to study its interactions with chitooligosaccharides as demonstrated by the changes in chemical shifts in the ^{19}F NMR spectra induced by $(GlcNAc)_5$. As exemplified by our current report, the utilization of TFM should be extremely useful in the study of protein-ligand and possibly protein-protein interactions by ^{19}F NMR. In proteins in which a specific function has been ascribed to particular methionine residues, the application of this approach may complement and offer additional insight on the roles and importance of this amino acid in protein structure and function.

Although the NMR characterization presented for the high and low level labelled enzymes has fittingly described the possible nature of the C/D double resonance, it would be of interest to examine the ^{19}F NMR spectra of TFM-LaL containing several other levels of incorporation. To this end, it would be appropriate and feasible to attempt to prepare a low level TFM-LaL in which the enzymes would contain either no or only a single incorporated TFM. This could be achieved by elevating the ratio of Met to TFM during protein induction. Although such an enzyme preparation would consist primarily of unlabelled (*wt*) LaL, those enzyme molecules that have incorporated only a single TFM at one of the three methionine positions would give rise to ^{19}F resonances that are indicative of the respective singularly labelled position. If our hypothesis regarding the nature of the C/D resonance is correct, then the spectra of such an enzyme sample should contain resonances corresponding to only A, B and C.

Conversely, but possibly less feasible, is to prepare TFM-LaL containing higher levels of TFM than the high level TFM-LaL already prepared. A possible approach could be to initially establish growth of *E. coli* B834(DE3)/pLR102 in medium containing a reduced *L*-Met supplement (a supplement of 0.02 mM is suggestive from Fig. 4.3) and to continue incubation of the cells in the medium for some time (2-3 hr) after cell growth has ceased. This would essentially result in the exhaustion of all sources of exogenous *L*-Met. Therefore, when the cells are collected and reintroduced into medium containing only TFM, there should be less "carry over" *L*-Met to compete with the TFM during protein induction. Furthermore, by reducing the induction time from 9.5 hr, this might also

decrease cellular recycling of *L*-Met resulting from salvage processes. It is expected that that such an approach would be accompanied with poor protein expression. However, this shortcoming might be circumventable by initially growing a large culture of cells (eg. 10-20 L) as described above but then reintroducing the methionine deprived cells into a smaller volume (eg. 1 L) of induction/labelling medium. If the ^{19}F NMR spectrum and ESMS characterization of this TFM-LaL are indicative of more complete incorporation generating primarily *tris*-labelled TFM-LaL, then it would be enticing to prepared adequate amounts of the *tris*-labelled enzyme for crystallographic studies.

Dr. E. Daub has also initiated the mutagenesis of the *R* gene contained in pLR102 to code for LaL in which leucine will replace the methionines at positions 14 and 107. We intend to prepare the M14L and M107L mutant proteins (and, if necessary, the double mutant) and incoroporate TFM into these mutants. It will then be possible to assign the respective A, B and C/D ^{19}F resonance to the specific Met positions. The mutants will also be useful towards the preparation of the respective [*methyl*-^{13}C]Met labelled proteins permitting the assignment of the ^{1}H and ^{13}C resonances observed in the HMQC spectrum of the *wt* ^{13}C-Met-LaL. The activity and inhibition of each mutant must be characterized. As was expressed in Chapter 3, it would be desirable to have access to a soluble, defined substrate for LaL for use in kinetic studies. It would indeed be advantageous to obtain the kinetic parameters of K_m and k_{cat} on such a successful substrate with the *wt*, TFM-labelled and mutant lysozymes to more fully examine the functional importance of the methionines in LaL.

Appendix A: Media and Buffer Preparation and Description

Materials. Bacto-agar, Bacto-casamino acids, Bacto-tryptone, Bacto-tryptose, Bacto-yeast extract, and Bacto-yeast nitrogen base are products of Difco and are distributed by Canlab (Toronto, ON). Trypticase® Peptone (which is a substitute for Bacto-tryptone) is from BBL and also distributed by Canlab. Bacto-tryptose was kindly donated by Dr. Barb Butler (Dept. of Biology, U. of Waterloo). Ammonium chloride (NH_4Cl, AnalaR), boric acid (H_3BO_3), citric acid ($C_6H_8O_7 \cdot H_2O$), glucose, magnesium chloride ($MgCl_2 \cdot 6H_2O$, AnalaR) magnesium sulfate ($MgSO_4 \cdot 7H_2O$), sodium ammonium phosphate ($NaNH_4HPO_4 \cdot 4H_2O$), sodium chloride (NaCl) and sucrose were obtained from BDH (Toronto, ON) and were of Assurance grade except where noted. Products obtained from Fischer Scientific (Don Mills, ON) included boric acid (H_3BO_3), calcium chloride ($CaCl_2$), glycerol, potassium phosphate monobasic (KH_2PO_4) and dibasic (K_2HPO_4) and sodium phosphate dibasic ($Na_2HPO_4 \cdot 7H_2O$). Ficoll (MW 400000) and sodium dodecylsulfate (SDS) were from ICN Biomedicals (Montreal, PQ). Bromophenol blue, 2-[N-morpholino]ethanesulfonic acid monohydrate (MES), PEG 8000, PIPES, and Triton X-100 were obtained from Sigma (St. Louis, MO). Ethylene-diaminetetraacetic acid (EDTA) and the disodium salt ($EDTA \cdot Na_2 \cdot 2H_2O$), Tris·HCl and Tris·base (electrophoretic grade) were products of Boehringer Manheim (Montreal, PQ). Nalgene sterilization filter units were obtained from Canadawide Scientific (Ottawa, ON).

A.1. Media

Media were sterilized by autoclaving for 20-40 min except where noted.

<u>Luria-Bertani Medium (LB)</u>

10 g Bacto-tryptone or Trypticase® Peptone
5 g Bacto-yeast extract
10 g NaCl
DDW to 1.0 L
adjust to pH 7 with 5 N NaOH
autoclave; cool; supplement with antibiotic

<u>Plates</u>: include 1.5 g agar (1.5% w/v) with the medium

<u>LT Medium</u>

LB medium supplemented with 0.0005% thiamine
- add 0.5 mL of 1% filter sterilized thiamine solution to 1.0 L LB medium

Supplemented M9 Minimal Medium

a. 10× M9 Salts

104 g $Na_2HPO_4 \cdot 7H_2O$
30 g KH_2PO_4
5 g NaCl
10 g NH_4Cl
DDW to 1.0 L

b. 1.0 M $MgSO_4$
Autoclave

c. 0.1 M $CaCl_2$
Autoclave

d. 30 or 40% Glucose (w/v)
Filter sterilize

add:
- 0.1 vol. (volume) solution a to 0.89 vol DDW; autoclave; cool
- 0.001 vol. of solution b (final conc. 1 mM)
- 0.001 vol. of solution c (final conc. 0.1 mM)
- 0.01 vol. solution d (final conc. 0.3 or 0.4%)
- supplement with antibiotic

Plates:
- prepare 3% agar; autoclave
- mix with 1 vol. of sterile 2× supplemented M9 medium
- supplement with antibiotic

Supplemented E Medium
(Ben-Basset et al., 1987)

a. 10× E Medium (Vogel & Bonner, 1956)

2 g $MgSO_4 \cdot 7H_2O$
20 g Citric Acid·H_2O
100 g K_2HPO_4
35 g $NaNH_4HPO_4 \cdot 4H_2O$

- add reagents in order to ca. 700 ml DDW until dissolved
- add additional DDW to 1.0 L

b. 2× Tryptose Broth plus 1% NaCl
40 g Bacto-Tryptose
10 g NaCl
DDW to 1.0 L
Autoclave

c. 10% Casamino Acids
10 g Bacto-casamino acids
DDW to 0.1 L
Autoclave

d. 10× Yeast Nitrogen Base
6.7 g Bacto-Yeast Nitrogen Base
DDW to 0.1 L
Filter sterilize
store at 4 °C

e. 40% Glucose (w/v)
Filter sterilize

add:
- 0.1 vol. solution a to 0.7375 vol. DDW; autoclave; cool
- 0.05 vol. each of solutions b, c and d
- 0.0125 vol. solution e (final conc. 0.5%)
- supplement with antibiotic

SOC Medium

		per 1.0 L
0.5%	Bacto-yeast extract	5 g
2%	Bacto-tryptone	20 g
10 mM	NaCl	0.585 g
2.5 mM	KCl	0.186 g
10 mM	$MgCl_2$	2.033 g ($MgCl_2 \cdot 6H_2O$)
20 mM	$MgSO_4$	4.929 g ($MgSO_4 \cdot 7H_2O$)
20 mM	glucose	3.603 g

A.2. Buffers

Potassium Phosphate Buffer (KPB)

a. 1.00 M Mono-basic potassium phosphate
   - dissolve 136.09 g KH_2PO_4 in MQW to 1.00 L in a volumetric flask

b. 1.00 M Di-basic potassium phosphate
   - dissolve 174.18 g K_2HPO_4 in MQW to 1.00 L in a volumetric flask

c. Stock 1.00 M KPB, pH 7.0
   - mix required amounts of solutions a and b at 25 °C to pH 7.0
   - filter and store at 4 °C; warm to 25 °C before use in preparation of solution d

d. X mM KPB, pH 7.0
   - appropriately dilute solution c with MQW to the desired concentration
   - eg. 50 mM KPB, pH 7.0
     - dilute 0.05 vol. of solution c with MQW to 1.00 vol. in a volumetric flask

Substrate Cell Buffer

		per 1.0 L
5 mM	$EDTA\text{-}Na_2$	1.86 g ($EDTA\text{-}Na_2 \cdot 2H_2O$)
60 mM	KPB, pH 7.0	60 mL of 1.00 M KPB, pH 7.0

1× STET Buffer

		per 100 m L
8%	Sucrose	8 g
0.5%	Triton X-100	0.5 mL
50 mM	EDTA	1.461 g
50 mM	Tris-HCl	0.788 g

adjust to pH 8.0 with conc. HCl

1x TE Buffer

		per 100 mL
10 mM	Tris-HCl	0.158 g
1 mM	EDTA	0.029 g

adjust to pH 8.0 with conc. HCl

5x TBE Buffer

		per 1.0 L
450 mM	Tris-base	54.495 g
450 mM	Boric acid	27.824 g
12.5 mM	EDTA	3.653 g

10x DNA Loading Buffer

		per 10 mL
20%	Ficoll 400	2 g
0.1 M	EDTA-Na_2	0.372 g (EDTA-Na_2-$2H_2O$)
1.0%	SDS	0.1 g
0.25%	Bromophenol Blue	25 mg

adjust to pH 8.0

$CaCl_2$ Buffer

		per 100 mL
60 mM	$CaCl_2$	0.666 g
15%	Glycerol	15 mL
10 mM	PIPES	0.274 g

adjust to pH 7.0; filter sterilize

5x PEG/NaCl Solution

		per 100 mL
15% (w/v)	PEG 8000	15 g
2.5 M	NaCl	14.613 g

Phage Buffer

		per 100 mL
100 mM	Tris-HCl	1.576 g
1 mM	EDTA	0.029 g
300 mM	NaCl	1.754 g

adjust to pH 8 with conc. HCl

Protein Loading Buffer

		per 100 mL
150 mM	Tris-base	1.818 g
2%	SDS	2 g
1%	2-mercaptoethanol	1 mL
10 %	Glycerol	10 mL
0.1%	Bromophenol Blue	0.1 g

adjust to pH 8 with conc. HCl and filter

MES Buffer

A. 0.5 M MES

- dissolve 2-[N-morpholino]ethane sulfonic acid monohydrate (10.663 g) into MQW (100.0 mL)

B. 100 mM MES, pH 5.8

- dilute 10.0 mL of A with ≈ 80 mL MQW
- adjust pH to 5.8 with 1 N NaOH and bring final volume to 100.0 mL

LaL Stabilizing Buffer

		per 100 mL
100 mM	KPB, pH 7.0	10.0 mL of 1.00 M KPB, pH 7.0
0.5 M	KCl	3.728 g
20% (v/v)	glycerol	20 mL

Appendix B: Crystal Screen™ Reagent Components

No.	SALT	BUFFER'	pH	PRECIPITANT
1.	0.02 M calcium chloride dihydrate	0.1 M Na acetate	4.6	30% v/v 2-methyl-2,4-pentanediol
2.	none	none		0.4 M K, Na tartrate tetrahydrate
3.	none	none		0.4 M ammonium dihydrogen phosphate
4.	none	0.1 M Tris HCl	8.5	2.0 M ammonium sulfate
5.	0.2 M tri-sodium citrate dihydrate	0.1 M Na Hepes	7.5	30% v/v 2-methyl-2,4-pentanediol
6.	0.2 M magnesium chloride hexahydrate	0.1 M Tris HCl	8.5	30% w/v PEG 4000
7.	none	0.1 M Na Cacodylate	6.5	1.4 M Na acetate trihydrate
8.	0.2 M tri-sodium citrate dihydrate	0.1 M Na Cacodylate	6.5	30% v/v 2-propanol
9.	0.2 M ammounium acetate	0.1 M Na citrate	5.6	30% w/v PEG 4000
10.	0.2 M ammounium acetate	0.1 M Na acetate	4.6	30% w/v PEG 4000
11.	none	0.1 M Na citrate	5.6	1.0 M ammonium dihydrogen phosphate
12.	0.2 M magnesium chloride hexahydrate	0.1 M Na Hepes	7.5	30% v/v 2-propanol
13.	0.2 M tri-sodium-citrate dihydrate	0.1 M Tris HCl	8.5	30% v/v PEG 400
14.	0.2 M calcium chloride dihydrate	0.1 M Na Hepes	7.5	28% v/v PEG 400
15.	0.2 M ammonium sulfate	0.1 M Na Cacodylate	6.5	30% w/v PEG 8000
16.	none	0.1 M Na Hepes	7.5	1.5 M Li sulfate monohydrate
17.	0.2 M lithium sulfate monohydrate	0.1 M Tris HCl	8.5	30% w/v PEG 4000
18.	0.2 M magnesium acetate tetrahydrate	0.1 M Na Cacodylate	6.5	20% w/v PEG 8000
19.	0.2 M ammonium acetate	0.1 M Tris HCl	8.5	30% v/v 2-propanol
20.	0.2 M ammonium sulfate	0.1 M Na acetate	4.6	25% w/v PEG 4000
21.	0.2 M magnesium acetate tetrahydrate	0.1 M Na Cacodylate	6.5	30% v/v 2-methyl-2,4-pentanediol
22.	0.2 M sodium acetate trihydrate	0.1 M Tris HCl	8.5	30% w/v PEG 4000
23.	0.2 M magnesium chloride hexahydrate	0.1 M Na Hepes	7.5	30% v/v PEG 400
24.	0.2 M calcium chloride dihydrate	0.1 M Na acetate	4.6	20% v/v 2-propanol
25.	none	0.1 M imidazole	6.5	1.0 M Na acetate trihydrate
26.	0.2 M ammonium acetate	0.1 M Na citrate	5.6	30% v/v 2-methyl-2,4-pentanediol
27.	0.2 M tri-sodium citrate dihydrate	0.1 M Na Hepes	7.5	20% v/v 2-propanol
28.	0.2 M sodium acetate trihydrate	0.1 M Na Cacodylate	6.5	30% w/v PEG 8000
29.	none	0.1 M Na Hepes	7.5	0.8 M K, Na tartrate tetrahydrate
30.	0.2 M ammonium sulfate	none		30% w/v PEG 8000
31.	0.2 M ammonium sulfate	none		30% w/v PEG 4000
32.	none	none		2.0 M ammonium sulfate
33.	none	none		4.0 M Na formate
34.	none	0.1 M Na acetate	4.6	2.0 M Na formate
35.	none	0.1 M Na Hepes	7.5	1.6 M Na, K phosphate monobasic
36.	none	0.1 M Tris HCl	8.5	8% w/v PEG 8000
37.	none	0.1 M Na acetate	4.6	8% w/v PEG 4000
38.	none	0.1 M Na Hepes	7.5	1.4 M Na citrate dihydrate
39.	none	0.1 M Na Hepes	7.5	2% v/v PEG 400 & 2.0 M ammonium sulfate
40.	none	0.1 M Na citrate	5.6	20% w/v PEG 4000 & 20% v/v 2-propanol
41.	none	0.1 M Na Hepes	7.5	20% w/v PEG 4000 & 10% v/v 2-propanol

No.	SALT	BUFFER[†]	pH	PRECIPITANT
42.	0.05 M potassium phosphate monobasic	none		20% w/v PEG 8000
43.	none	none		30% w/v PEG 1450
44.	none	none		0.2 M magnesium formate
45.	0.2 M zinc acetate dihydrate	0.1 M Na Cacodylate	6.5	18% w/v PEG 8000
46.	0.2 M calcium acetate	0.1 M Na Cacodylate	6.5	18% w/v PEG 8000
47.	none	0.1 M Na acetate	4.6	2.0 M ammonium sulfate
48.	none	0.1 M Tris HCl	8.5	2.0 M ammonium dihydrogen phosphate
49.	1.0 M lithium sulfate monohydrate	none		2% w/v PEG 8000
50.	0.5 M lithium sulfate monohydrate	none		15% w/v PEG 8000

[†] Buffer pH is that of a 1.0 M stock prior to dilution with the other reagents.

Appendix C: Sample Calculations

C.1. Residual Activity

All inhibition values and errors presented in figures or listed in tables were determined as follows. The following calculation including propagation of errors utilizes data taken from the inhibition of LaL by $(GlcNAc)_4$ at 5 mM. In the particular example, 4 μL of a 4 ng/μL solution of LaL (in stabilizing buffer) was the enzyme aliquot.

Sample	Order performed in Assay	Δ OD_{600} between 2 and 3 min	Average ± standard deviation
Blank	1	-0.002717	0.0042 ± 0.0022
	8	-0.005764	
16 ng LaL	2	-0.057773	0.0591 ± 0.0029
	4	-0.057119	
	6	-0.062466	
16 ng LaL and 5 mM $(GlcNAc)_4$	3	-0.010204	0.0101 ± 0.0009
	5	-0.009176	
	7	-0.010883	

$$\text{Residual Activity} = \frac{\text{activity with inhibitor - blank}}{\text{activity without inhibitor - blank}}$$

$$= \frac{(0.0101 \pm 0.0009) - (0.0042 \pm 0.0022)}{(0.0591 \pm 0.0029) - (0.0042 \pm 0.0022)}$$

$$= \frac{0.0059 \pm 0.0024^{a}}{0.0549 \pm 0.0036^{a}}$$, where ^{a} is the standard deviation of the sum

$$= \frac{0.0059 \pm 0.4068^{b}}{0.0549 \pm 0.0656^{b}}$$, where ^{b} are the individual relative errors

$= 0.107 \pm 0.4121^{c}$, where ^{c} is the sum of the individual relative errors

$= 0.107 \pm 0.044^{d}$, where ^{d} is the absolute standard deviation of the result

Propagation of Errors (Skoog & West, 1982)

i) Summation: the standard deviation of a sum (see errors with "^{a}" above) is

$s_y = (s_a^2 + s_b^2)^{1/2}$, where s_y is the standard deviation of the sum and s_a and s_b are the standard deviations of the two terms in the sum

eg. s_y for activity with inhibitor - blank $= [\,(\pm 0.0009)^2 + (\pm 0.0022)^2\,]^{1/2}$
$= \pm 0.0024$

ii) Division:

- the relative errors (see errors with "b "above) are first determined by

$$(s_a)_r = \frac{\pm 0.0024}{0.0059} = \pm 0.4068$$

$$(s_b)_r = \frac{\pm 0.0036}{0.0549} = \pm 0.0656$$

- the relative error of the result (see error "c " above) is

$$s_y = [\,(s_a)_r^2 + (s_b)_r^2)\,]^{½}$$
$$= [\,(\pm 0.4068)^2 + (\pm 0.0656)^2\,]$$
$$= \pm 0.4121$$

- the absolute standard deviation of the result is (see error "d " above)

$$= 0.107 \times (\pm 0.4121)$$
$$= \pm 0.044$$

The percentage residual activity is expressed by multiplication of the result and absolute standard deviation by 100%; i.e. $[0.107 \pm 0.044] \times 100\% = 10.7 \pm 4.4\%$ or more appropriately, $11 \pm 4\%$. Percentage inhibition is simply $100\% - (11 \pm 4\%) = 89 \pm 4\%$.

C.2. Preparation of Protein Samples for DSC

The sample calculation given below is based on the same LaL preparation that was described in section 2.2.10.2 (p. 108). The values required are:

Total Mass of protein/buffer powder:	46.05 mg
Mass Contribution from LaL:	39.86 mg (86.56% of the powder)
Wet Mass of Sample obtained from dialysis vs. 5 mM KPB, pH 7.0:	7.25930 g

For efficient loading of the calorimetric cells, 2.30 mL of sample was required. The following calculations are necessary:

1. Amount of LaL required = 2.3 mL × 3.5 mg/mL = 8.05 mg
2. Amount of powder required = $\frac{8.05\ \text{mg}}{86.56\%}$ = 9.30 mg
3. Dissolving 9.30 mg of the protein/buffer powder into 2.30 mL would result in:

$$\frac{9.30\ \text{mg}}{46.05\ \text{mg}} \times \frac{7.25930\ \text{mL}}{2.30\ \text{mL}} \times 5.0\ \text{mM} = 3.187\ \text{mM}$$

   i.e. a 3.187 mM KPB solution would result from the buffer salts in the powder
4. Additional buffer required = 50 mM - 3.187 mM = 46.813 mM
5. Amount of 1.00 M KPB, pH 7.0 required = $\frac{46.813\ \text{mM}}{1.00\ \text{M}} \times 2.3\ \text{mL} = 0.1077\ \text{mL}$

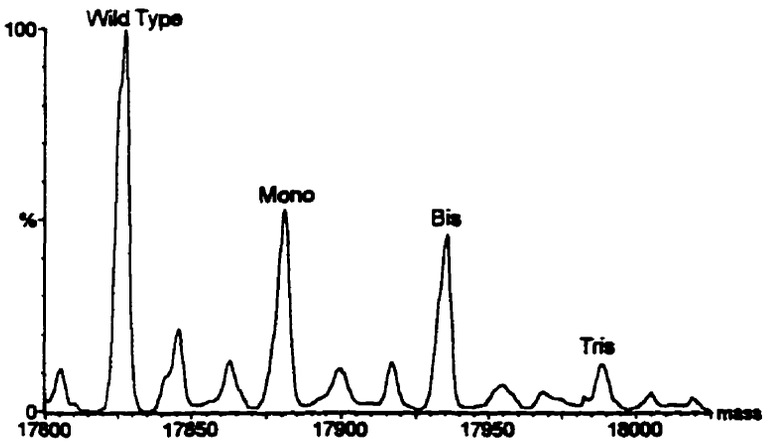
Therefore, to prepare a 3.5 mg/mL solution of LaL in 50 mM KPB, pH 7.0 using this particular LaL preparation, 9.30 mg of the protein/buffer mixture was dissolved into 2.192 mL of MQW and 108 μL mL of 1.00 M KPB, pH 7.0 was added. Similar calculations were used to prepare samples of LaL at other concentrations and from different protein/buffer powders obtained from various purifications.

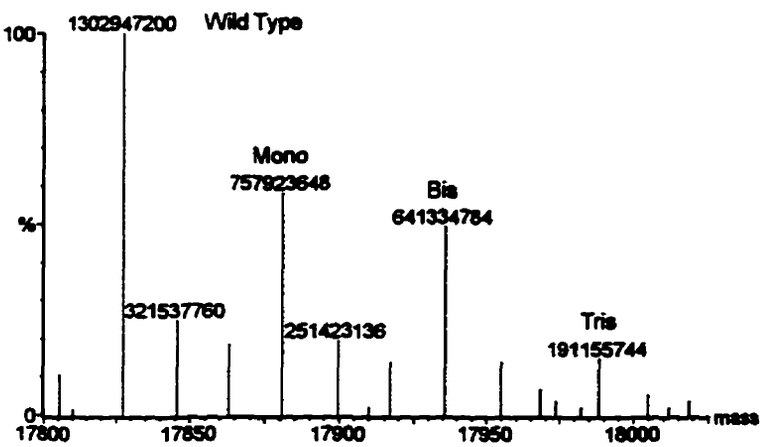
Appendix D: Determination of the Level of TFM Incorporation

The level of TFM incorporation into LaL prepared under low and high incorporation conditions were estimated from their ESMS spectra. This was achieved by producing a bar spectrum with heights proportional to component concentrations in the MaxEnt spectrum. The procedure is as follows. Using the MassLynx software, the desired MaxEnt spectrum is recovered. Centre is chosen from the Process menu. Under the centre method options, "top" is selected and the minimum peak width at half height is set to 2 (in channels). Under the centered spectrum options, the "areas" option is selected and the centred spectrum is created as a bar spectrum proportional to the respective component areas. The method assumes that the ionization and detection of the individual components are the same and can therefore be quantitated.

1) Example for low level TFM-LaL

Shown below are the MaxEnt and centred spectra for low level TFM-LaL with the values corresponding to area of the component peaks given in the centered spectrum.





The incorporation is then calculated with the following data.

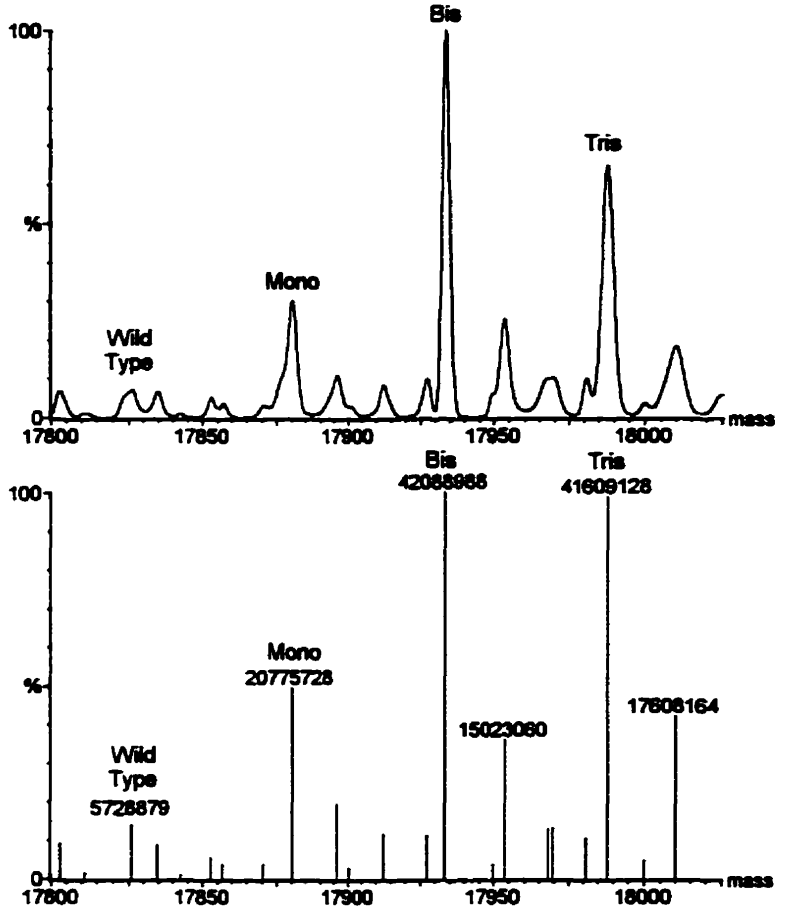
Enzyme Species	Area	Relative Area	Relative Presence		% TFM Present		Contribution
wild type	1302947200	1	0.450	×	0	=	0
Mono	757923648	0.5817	0.262	×	33.33	=	8.7 %
Bis	641334784	0.4922	0.222	×	66.67	=	14.8 %
Tris	191155744	0.1467	0.066	×	100.00	=	6.6 %
Total		2.2206	1.000				30.1 %

The relative area is found by dividing the area for each enzyme species by the wild type area. The relative presence is found by dividing the relative area by the sum of the relative areas (2.2206 in this case). Multiplication of the relative presence by the % TFM in each species (i.e. *wt* LaL has no TFM residues, a mono-species has 1 out of 3 positions substituted or 33.33%, a bis-species has 2 out of 3 positions substituted or 66.67%, and the tris-labelled enzyme has 100% substitution of the Met residues by TFM) gives the contribution to incorporation from each species. The sum of the individual contributions then gives the total level of incorporation and in this case, is 30.1%.

Three individually acquired MaxEnt spectra from the sample of low level TFM-LaL were obtained and subjected to the analysis described above. Similar calculations on the other two spectra gave values of 30.2 and 31.3%, and including the 30.1% value calculated above, gives an average of 30.5 ± 0.7%.

An example of the data obtained for one of the two samples of high level TFM-LaL prepared is given below. In this case, the total incorporation was 69.6%. Replica measurements were not performed on this sample. The MaxEnt and centered spectra are shown on the following page. For the other sample of high level TFM-LaL prepared, the relative presence of wild type, mono, bis, and tris were 0.053, 0.125, 0.397, 0.434 respectively which equates to 74.0% incorporation.

Enzyme Species	Area	Relative Area	Relative Presence		% TFM Present		Contribution
wild type	5728879	0.1361	0.052	×	0	=	0
Mono	20775728	0.4936	0.189	×	33.33	=	6.3 %
Bis	42088988	1	0.382	×	66.67	=	25.5 %
Tris	41609128	0.9886	0.378	×	100.00	=	37.8 %
Total		2.3024	1.00				69.6 %


Bis
Tris
Mono
Wild
Type
100
%
0
17800
17850
17900
17950
18000
mass
Bis
42068968
Tris
41609128
Mono
20775728
15023060
17606164
Wild
Type
5728879

REFERENCES

Abe, Y., Ueda, T., Iwashita, H., Hashimoto, Y., Motoshima, H., Tanaka, Y., & Imoto, T. (1995) *J. Biochem.*, **118**, 946-952

Acharya, K. R., Ren, J., Stuart, D. I., Phillips, D. C., & Fenna, R. E. (1991) *J. Mol. Biol.*, **221**, 571-581

Adam, S. (1992) *Bioorg. Med. Chem. Lett.*, **2**, 571-574

Agrawal, P. K. (1992) *Phytochemistry*, **31**, 3307-3330

Alber, T., & Matthews, B. W. (1987) *Methods Enzymol.*, **154**, 511-533

Anand, N. N., Stephen, E. R., & Narang, S. A. (1988) *Biochem. Biophys. Res. Commun.*, **153**, 862-868

Anderson, W. F., Grütter, M. G., Remington, S. J., Weaver, L. H., & Matthews, B. W. (1981) *J. Mol. Biol.*, **147**, 523-543

Araki, Y., Nakatani, T., Makino, R., Hayashi, H., & Ito, E. (1971) *Biochem. Biophys. Res. Commun.*, **42**, 684-690

Arendt, E. K., Daly, C., Fitzgerald, G. F., & van de Guchte, M. (1994) *Appl. Environ. Microbiol.*, **60**, 1875-1883

Arnheim, N. (1974) in *Lysozyme* (Osserman, E. F., Canfield, R. E. & Beychok, S., eds.) pp. 153-168, Academic Press, New York

Arnheim, N., Hindenburg, A., Begg, G. S., & Morgan, F. J. (1973a) *J. Biol. Chem.*, **248**, 8036-8042

Arnheim, N., Inouye, M., Law, L., & Laudin, A. (1973b) *J. Biol. Chem.*, **248**, 233-236

Arnold, G. E., Manchester, J. I., Townsend, B. D., & Ornstein, R. L. (1994) *J. Biomolec. Struct. Dyn.*, **12**, 457-474

Artymiuk, P. J., & Blake, C. C. F. (1981) *J. Mol. Biol.*, **152**, 737-762

Artymiuk, P. J., Blake, C. C. F., & Sippel, A. E. (1981) *Nature*, **290**, 287-288

Aschattenburg, R., Blake, C. C. F., Dickie, H. M. Gayen, S. K., Keegan, R., & Sen, A. (1980) *Biochim. Biophys. Acta*, **625**, 64-71

Audy, P., Grenier, H., & Asselin, A. (1989) *Comp. Biochem. Physiol. B.*, **92**, 523-527

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K. (1989) in *Short Protocols in Molecular Biology*, John Wiley & Sons, New York

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K. (1991) *in Current Protocols in Molecular Biology*, John Wiley & Sons, New York

Avila, L. Z., Chu, Y.-H., Blossey, E. C., & Whitesides, G. M. (1993) *J. Med. Chem.*, **36**, 126-133

Azuma, I., Thomas, D. W., Adam, A., Ghuysen, J.-M., Bonaly, R., Petit, J. F., & Lederer, E. (1970) *Biochim. Biophys. Acta*, **208**, 444-451

Bahr, G. M., & Chedid, L. (1986) *FASEB J.*, **45**, 2541-2544

Ballardie, F. W., Capon, B., Cuthbert, M. W., & Dearie, W. M. (1977) *Bioorganic Chem.*, **6**, 483-509

Banerjee, S. K., & Rupley, J. A. (1973a) *J. Biol. Chem.*, **248**, 2117-2124

Banerjee, S. K., & Rupley, J. A. (1973b) *Arch. Biochem. Biophys.*, **155**, 19-23

Banerjee, S. K., Kregar, I., Turk, V., & Rupley, J. A. (1973) *J. Biol. Chem.*, **248**, 4786-4792

Banerjee, S. K., Vandenhoff, G. E., & Rupley, J. A. (1974) *J. Biol. Chem.*, **249**, 1439-1444

Banerjee, S. K., Holler, E., Hess, G. P., & Rupley, J. A. (1975) *J. Biol. Chem.*, **250**, 4355-4367

Banoub, J., Boullanger, P., & Lafont, D. (1992) *Chem. Rev.*, **92**, 1167-1195

Barden, J. A., & Kemp, B. E. (1993) *Biochemistry*, **32**, 7126-7132

Barden, J. A., Phillips, L., Cornell, B. A., & dos Remedios, C. G. (1989) *Biochemistry*, **28**, 5895-5901

Barnickel, G., Labischinski, H., Bradaczek, H., & Giesbrecht, P. (1979) *Eur. J. Biochem.*, **95**, 157-165

Baschang, G. (1989) *Tetrahedron*, **45**, 6331-6360

Baumstark, B. R., Spremulli, L. L., RajBhandary, U. L., & Brown, G. M. (1977) *J. Bacteriol.*, **129**, 457-471

Bax, A., Ikura, I., Kay, L. E., Torchia, D. A., & Tschudin, R., (1990) *J. Magn. Res.*, **86**, 304-318

Beddell, C. R., Moult, J., & Phillips, D. C. (1970) in *Molecular Properties of Drug Receptors* (Porter, R. & O' Connor, M., eds.) p 85, Churchill, London

Ben-Bassat, A., Bauer, K., Chang, S. Y., Myambo, K., Boosman, A., & Chang, S. (1987) *J. Bacteriol.*, **169**, 751-757

Bennet, A. J., & Sinnott, M. L. (1986) *J. Am. Chem. Soc.*, **108**, 7287-7294

Berger, L. R., & Weiser, R. S. (1957) *Biochim. Biophys. Acta*, **26**, 517-521

Bernard, A. R., Wells, T. N. C., Cleasby, A., Borlat, F., Payton, M. A., & Proudfoot, A. E. I. (1995) *Eur. J. Biochem.*, **230**, 111-118

Bernstein, H. D., Poritz, M. A., Strub, K., Hoben, P. J., Brenner, S., & Walter, P. (1989) *Nature*, **340**, 482-486

Beveridge, T. J. (1988) *Can. J. Microbiol.*, **34**, 363-372

Beveridge, T. J., & Graham, L. L. (1991) *Microbiol. Rev.*, **55**, 684-705

Biberstine, K. J., & Rosenthal, R. S. (1995) *Infect. Immun.*, **62**, 3276-3281

Bienkowska, K., & Taylor, A. (1979) *Eur. J. Biochem.*, **96**, 581-584

Bienkowska-Szewczyk, K., & Taylor, A. (1980) *Biochim. Biophys. Acta*, **615**, 489-496

Bienkowska-Szewczyk, K., Lipinska, B., & Taylor, A. (1981) *Mol. Gen. Genet.*, **184**, 111-114

Birdsell, D. C., & Cota-Robles, E. H. (1967) *J. Bacteriol.*, **93**, 427-437

Black, L. W., & Hogness, D. S. (1969a) *J. Biol. Chem.*, **244**, 1968-1975

Black, L. W., & Hogness, D. S. (1969b) *J. Biol. Chem.*, **244**, 1976-1981

Black, L. W., & Hogness, D. S. (1969c) *J. Biol. Chem.*, **244**, 1982-1987

Blackwell, J. (1988) *Methods Enzymol.*, **161**, 435-442

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T, Phillips, D. C., & Sarma, V. R. (1965) *Nature (Lond.)*, **206**, 757-761

Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T, Phillips, D. C., & Sarma, V. R. (1967) *Proc. R. Soc. London Ser B*, **167**, 378-388

Bodor, N., Gabanyi, Z., & Wong, C-K. (1989) *J. Amer. Chem. Soc.*, **111**, 3783-3786

Boizet, B., Lahbib-Mansais, Y., Dupont, L., Ritzenthaler, P., & Mata, M. (1990) *Gene*, **94**, 61-67

Boles, J. O., Cisneros, R. J., Weir, M. S., Odom, J. D., Villafranca, J. E., & Dunlap, R. B. (1991) *Biochemistry*, **30**, 11073-11080

Boles, J. O., Lewinski, K., Kunke, M., Odom, J. D., Dunlap, B. R., Lebioda, L., & Hatada, M. (1994) *Nat. Struct. Biol.*, **1**, 283-284

Bonilla, C. A. (1970) *J. Chromatogr.*, **47**, 499-501

Bonovich, M. T., & Young, R. Y. (1991) *J. Bacteriol.*, **173**, 2897-2905

Bradford, M. M. (1976) *Anal. Biochem.*, **72**, 248-254

Brandts, J. F., Hu, C. Q., Lin, L. N., & Mas, M. T. (1989) *Biochemistry*, **28**, 8588-8596

Braun, V. (1975) *Biochim. Biophys. Acta*, **415**, 335-377

Braun, V., & Bosch, V. (1972) *Eur. J. Biochem.*, **28**, 51-56

Braun, V., & Rehn, K. (1969) *Eur. J. Biochem.*, **10**, 426-438

Braun, V., Rehn, K., & Wolff, H. (1970) *Biochemistry*, **9**, 5041-5049

Brennan, R. G., Wozniak, J., Faber, R., & Matthews, B. W. (1988) *J. Crystal Growth*, **90**, 160-167

Brot, N., & Weissbach, H. (1991) *Biofactors*, **3**, 91-96

Brot, N., Fliss, H., Coleman, T., & Weissbach, H. (1984) *Meth. Enzymol.*, **107**, 352-360

Bruix, M., Jiménez, M. A., Santoro, J., González, C., Colilla, F. J., Méndez, E., & Rico, M. (1993) *Biochemistry*, **32**, 715-724

Brunie, S., Zelwer, C., & Risler, J. L. (1990) *J. Mol. Biol.*, **216**, 411-424

Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J., & Huber, R. (1995) *Eur. J. Biochem.*, **230**, 788-796

Bugg, T. D. H., & Walsh, C. T. (1992) *Nat. Prod. Rep.*, 199-215

Burge, R. E., Fowler, A. G., & Reaveley, D. A. (1977) *J. Mol. Biol.*, **117**, 927-953

Byers, D. A., & Leaback, D. H. (1988) *Analytica Chimica Acta*, **205**, 89-95

Caldentey, J., & Bamford, D. H. (1992) *Biochim. Biophys. Acta*, **1159**, 44-50

Caldentey, J., Hänninen, A.-L., & Bamford, D. H. (1994) *Eur. J. Biochem.*, **225**, 341-346

Campbell, A. (1961) *Virology*, **14**, 22-32

Canfield, R. E. (1963) *J. Biol. Chem.*, **238**, 2698-2707

Capon, B., & Dearie, W. M. (1974) *J. Chem. Soc. D*, 371-374

Carpenter, F. H., & Shiigi, S. M. (1974) *Biochemistry*, **13**, 5159-5164

Carraway, K. L., & Koshland, D. E. (1968) *Biochim. Biophys. Acta*, **160**, 272-274

Carraway, K. L., & Triplett, R. B. (1970) *Biochim. Biophys. Acta*, **200**, 564-566

Carter, C. W., Jr. (1992) in *Crystallization of Nucleic Acids and Proteins - A Practical Approach* (Ducruix, A., & Giegé, R. eds.), pp. 64, IRL Press, Oxford

Carter, T. H., & Miller, C. G. (1984) *J. Bacteriol.*, **159**, 453-459

Chan, V. W. F., Jorgensen, A. M., Borders, C. L. (1988) *Biochem. Biophys. Res. Commun.*, **151**, 709-716

Chang, J.-J., Holladay, L. A., & Hash, J. H. (1979) *J. Biol. Chem.*, **254**, 7772-7777

Chang, C. Y., Nam, K., & Young, R. Y. (1995) *J. Bacteriol.*, **177**, 3283-3294

Cheetham, J. C., Artymiuk, P. J., & Phillips, D. C. (1992) *J. Mol. Biol.*, **224**, 613-628

Chenault, H. K., Dahmer, J., & Whitesides, G. M. (1989) *J. Am. Chem. Soc.*, **111**, 6354-6364

Cheng, X., Zhang, X., Pflugrath, J. W., & Studier, F. W. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 4034-4038

Chipman, D. M., & Sharon, N. (1969) *Science*, **165**, 454-465

Chipman, D. M., Grisaro, V., & Sharon, N. (1967) *J. Biol. Chem.*, **242**, 4388-4392

Chitarra, V., Alzari, P. M., Bentley, G. A., Bhat, T. N., Eiselé, J. L., Houdusse, A., Lescar, J., Souchon, H., & Poljak, R. J. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 7711-7715

Chou, T. C., & Talalay, P. (1972) *Biochemistry*, **11**, 1065-1073

Chow, S., Daub, E., & Murialdo, H. (1987) *Gene*, **60**, 277-289

Chu, Y.-H., Avila, L. Z., Biebuyck, H. A., & Whitesides, G. M. (1992) *J. Med. Chem.*, **35**, 2915-2917

Chu, Y.-H., Lees, W. J., Stassinopoulos, A., & Walsh, C. T. (1994) *Biochemistry*, **33**, 10616-10621

Chu, Y.-H., & Whitesides, G. M. (1992) *J. Org. Chem.*, **57**, 3524-3525

Clarke, A. J. (1993) *J. Bacteriol.*, **175**, 4550-4553

Clarke, A. J., & Dupont, C. (1992) *Can. J. Microbiol.*, **38**, 85-91

Cohen, S. N., Chang, A. C. Y., Hsu, L. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 2110-2114

Collyer, C. A., Guss, J. M., Sugimura, Y., Yoshizaki, F., & Freeman, H. C. (1990) *J. Mol. Biol.*, **211**, 617-632

Colombani, F., Cherest, H., & de Robichon-Szulmajster, H. (1975) *J. Bacteriol.*, **122**, 375-385

Colon, M., Staveski, M. M., & Davis, J. T. (1991) *Tetrahedron Lett.*, **32**, 4447-4450

Compton, S. J., & Jones, C. G. (1985) *Anal. Biochem.*, **151**, 369-374

Cookson, B. T., Cho, H.-L., Herwaldt, L. A., & Goldman, W. E. (1989) *Infect. Immun.*, **57**, 2223-2229

Croux, C., & García, J. L. (1991) *Gene*, **104**, 25-31

Cummins, C. S., & Harris, H. (1956) *J. Gen. Microbiol.*, 14, 583-600

Dahlquist, F. W., Jao, L., & Raftery, M. (1966) *Proc. Natl. Acad. Sci. USA*, **56**, 26-30

Dahlquist, F. W., Rand-Meir, T., & Raftery, M. A. (1969a) *Biochemistry*, **8**, 4214-4221

Dahlquist, F. W., Borders, C. L., Jacobson, G., & Raftery, M. A. (1969b) *Biochemistry*, **8**, 694-700

Danielson, M. A., & Falke, J. J. (1996) *Annu. Rev. Biophys. Biomol. Struct.*, **25**, 163-195

Dannley, R. L., & Taborsky, R. G. (1957) *J. Org. Chem.*, **22**, 1275-1276

Dao-Pin, S., Liao, D., & Remington, S. J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5361-5365

Dao-Pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W., & Matthews, B. W. (1991) *Biochemistry*, **30**, 11521-11529

Davies, R. C., Neuberger, A., & Wilson, B. M. (1969) *Biochim. Biophys. Acta*, **178**, 294-503

Dawson, R. M. C., Elliott, D. C., Elliott, W. H., & Jones, K. M. (1986) in *Data for Biochemical Research 3rd ed.*, p. 543, Oxford University Press, Oxford

de Boer, H. A., & Hui, A. S. (1990) *Methods Enzymol.*, **185**, 103-114

De Jonge, B. L. M., Sidow, T., Chang, Y.-S., Labischinski, H., Berger-Bächi, B., Gage, D., & Tomasz, A. (1993) *J. Bacteriol.*, **175**, 2779-2782

Dell, A. (1986) in *Practical Protein Chemistry - A Handbook* (Darbre, A., Ed.) pp 531-532, John Wiley & Sons

Delmotte, F. M., Privat, J. P. D., & Monsigny, M. L. P. (1975) *Carbohydr. Res.*, **40**, 353-364

Dewar, M. J. S., Zoebisch, E. G., Healy, E. F., & Stewart, J. J. P. (1985) *J. Amer. Chem. Soc.*, **107**, 3902-3909

Diaz, E., López, R., & García, J. L. (1991) *J. Biol. Chem.*, **266**, 5464-5471

Dijkstra, B. W., & Thunnissen, A.-M. W H. (1994) *Curr. Opin. Struct. Biol.*, **4**, 810-813

Dill, K. A., Alonso, D. O. V., & Hutchinson, K. (1989) *Biochemistry*, **28**, 5439-5449

Dillon, J. R. (1985) in *Recombinant DNA Methodology* (Dillon, J. R., Nasim, A., & Nestmann, E. R., eds) John Wiley & Sons, New York

Dobson, C. M. (1990) *Nature*, **348**, 198-199

Dokter, W. H. A., Dijkstra, A. J., Koopmans, S. B., Stulp, B. K., Keck, W., Halie, M. R., & Vellenga, E. (1994a) *J. Biol. Chem.*, **269**, 4201-4206

Dokter, W. H. A., Dijkstra, A. J., Koopmans, S. B., Mulder, A. B., Stulp, B. K., Halie, M. R., Keck, W., & Vellenga, E. (1994b) *Infect. Immun.*, **62**, 2953-2957

Doolittle, R. F. (1989) in *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G. D., ed.) pp. 599-623, Plenum Press, New York

Douzou, P., & Petsko, G. A. (1984) *Adv. Protein Chem.*, **36**, 245-359

Doyen, N., & Lapresle, C. (1979) *Biochem. J.*, **177**, 251-254

Ducruix, A., & Giegé, R. (1992) in *Crystallization of Nucleic Acids and Proteins - A Practical Approach* (Ducruix, A., & Giegé, R. eds.), pp. 73-97, IRL Press, Oxford

Dukor, P., Tarcsay, L., & Baschang, G. (1979) *Annu. Rep. Med. Chem.*, **14**, 146-167

Durette, P. L., Meitzner, E. P., & Shen, T. Y. (1979) *Carbohydr. Res.*, **77**, C1-C4

Durig, J. R., Rollins, M. S., & Phan, H. V. (1991) *J. Molec. Struct.*, **263**, 95-122

Dykes, D. F., Creighton, T. E., & Sheppard, R. C. (1974) *Nature*, **247**, 202-204

Dziarski, R. (1986) in *Biological Properties of Peptidoglycan* (Seidl, P. H., & Schleifer, K. H., eds.) pp. 229-247, Walter de Gruyter, Berlin

Dziarski, R. (1991a) *J. Biol. Chem.*, **266**, 4713-4718

Dziarski, R. (1991b) *J. Biol. Chem.*, **266**, 4719-4725

Dziarski, R. (1994) *J. Biol. Chem.*, **269**, 20431-20436

Edelhoch, H. (1967) *Biochemistry*, **6**, 1948-1954

Eipper, B. A., Quon, A. S. W., Mains, R. E., Boswell, J. S., & Blackburn, N. J. (1995) *Biochemistry*, **34**, 2857-2865

Ellouz, F., Adam, A., Ciorbaru, R., & Lederer, E., (1974) *Biochem. Biophys. Res. Commun.*, **59**, 1317-1325

Engel, H., Kazemier, B., & Keck, W. (1991) *J. Bacteriol.*, **173**, 6773-6782

Engel, H., Smink, A. J., van Wijngaarden, L., & Keck, W. (1992a) *J. Bacteriol.*, **174**, 6394-6403

Engel, H., van Leeuwen, A., Dijkstra, A., & Keck, W. (1992b) *Appl. Microbiol. Biotechnol.*, **37**, 772-783

Eshdat, Y., McKelvy, J. F., & Sharon, N. (1973) *J. Biol. Chem.*, **248**, 5892-5898

Eshdat, Y., Dunn, A., & Sharon, N. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 1658-1662

Faber, H. R., & Matthews, B. W. (1990) *Nature*, **348**, 263-266

Feeney, J., McCormick, J. E., Bauer, C. J., Birdsall, B., Moody, C. M., Starkmann, B. A., Young, D. W., Francis, P., Havlin, R. H., Arnold, W. D., & Oldfield, E. (1996) *J. Am. Chem. Soc.*, **118**, 8700-8706

Fersht, A. R. (1974) *Proc. R. Soc. London Ser B*, **187**, 397-407

Fields, G. B., & Noble, R. L. (1990) *Int. J. Peptide Protein Res.*, **35**, 161-214

Fink, A. L., Horner, R., & Weber, J. P. (1980) *Biochemistry*, **19**, 811-820

Fleet, G. W. J. (1985) *Tetrahedron Lett.*, **26**, 5073-5076

Fleming, A. (1922) *Proc. R. Soc. London Ser B*, **93**, 306-317

Fletcher, H. G. (1963) *Methods Carbohydr. Chem.*, **II**, 307-308

Flowers, H. M., & Jeanloz, R. W. (1963) *J. Org. Chem.*, **28**, 2983-2986

Fontana, A., & Gross, E. (1986) in *Practical Protein Chemistry - A Handbook* (Darbre, A., Ed.) pp 67-120, John Wiley & Sons

Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C., & Tjian, R. (1974) *J. Mol. Biol.*, **88**, 349-371

Formanek, J., Formanek, S., & Wawra, H. (1974) *Eur. J. Biochem.*, **46**, 279-294

Fouche, P. B., & Hash, J. H. (1978) *J. Biol. Chem.*, **253**, 6787-6793

Frank , R. W. (1992) *Bioorg. Chem.*, **20**, 77-88

Freifelder, D. (1982) in *Physical Biochemistry 2nd edn.*, pp. 537-572, W. H. Freeman & Company, San Francisco

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. G., Robb, M. A., Cheeseman, J. R., Keith, T., Petersson, G. A., Montgomery, J. A., Raghavachari, K., Al-Laham, M. A., Zakrzewski, V. G., Ortiz, J. V., Foresman, J. B., Cioslowski, J., Stefanov, B. B., Nanayakkara, A., Challacombe, M., Peng, C. Y., Ayala, P. Y., Chen, W., Wong, M. W., Andres, J. L., Replogle, E. S., Gomperts, R., Martin, R. L., Fox, D. J., Binkley, J. S., Defrees, D. J., Baker, J., Stewart, J. P., Head-Gordon, M., Gonzalez, C., & Pople, J. A., (1995) *Gaussian 94, Revision C.3*, Gaussian, Inc., Pittsburgh, PA

Fukada, H., Tanimoto, T., & Yamaha, T. (1985) *Chem. Pharm. Bull.*, **33**, 3375-3380

Fukamizo, T., Torikata, T., Nagayama, T., Minematsu, T., & Hayashi, K. (1983) *J. Biochem.*, **94**, 115-122

Fukamizo, T., Ikeda, Y., Torikata, T., Araki, T., Kuramoto, M., & Goto, S. (1991) *J. Biochem.*, **110**, 997-1003

Fukamizo, T., Ikeda, Y., Ohkawa, T., & Goto, S. (1992) *Eur. J. Biochem.*, **210**, 351-357

Fukuhara, K.-I., Tsuji, T., Toi, K., Takao, T., & Shimonishi, Y. (1985) *J. Biol. Chem.*, **260**, 10487-10494

Ganem, B., Li, Y.-T., & Henion, J. D. (1991) *J. Am. Chem. Soc.*, **113**, 7818-7819

Garcia-Bustos, J. F., Chait, B. T., & Tomasz, A. (1988) *J. Bacteriol.*, **170**, 2143-2147

García, J. L., García, E., Sánchez-Puelles, J. M., & López, R. (1988a) *FEMS Microbiol. Lett.*, **52**, 133-138

García, E., García, J. L., García, P., Arrarás, A., Sánchez-Puelles, J. M., & López, R. (1988b) *Proc. Natl. Acad. Sci. USA*, **85**, 914-918

García, P., García, J. L., García, E., Sánchez-Puelles, J. M., & López, R. (1990) *Gene*, **86**, 81-88

Garrett, J.-M., & Young, R. Y. (1982) *J. Virol.*, **44**, 886-892

Garvey, K. J., Saedi, M. S., & Ito, J. (1986) *Nucl. Acids Res.*, **14**, 10001-10008

Gavezzotti, A. (1983) *J. Amer. Chem. Soc.*, **105**, 5220-5225

Gebler, J. C., Aebersold, R., & Withers, S. G. (1992) *J. Biol. Chem.*, **267**, 11126-11130

Gellman, S. H. (1991) *Biochemistry*, **30**, 6633-6636

Gerig, J. T. (1989) *Methods Enzymol.*, **177B**, 3-23

Gerig, J. T. (1994) *Prog. NMR Spectrosc.*, **26**, 293-370

Ghuysen, J.-M. (1968) *Bacteriol. Rev.*, **32**, 425-464

Ghuysen, J.-M. (1974) in *Lysozyme* (Osserman, E. F., Canfield, R. E. & Beychok, S., eds.) pp. 185-193, Academic Press, New York

Ghuysen, J.-M., & Strominger, J. L. (1963) *Biochemistry*, **2**, 1119-1125

Ghuysen, J.-M., Tipper, D. J., & Strominger, J. L. (1966) *Methods Enzymol.*, **8**, 685-699

Gigg, R., Carroll, P. M., & Warren, C. D. (1965) *J. Chem. Soc.*, 2975-2977

Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem.*, **182**, 319-326

Gisin, B. F. (1973) *Helv. Chim. Acta.*, **56**, 1476-1482

Glauner, B. (1988) *Anal. Biochem.*, **172**, 451-464

Glauner, B., Höltje, J.-V., & Schwarz, U. (1988) *J. Biol. Chem.*, **263**, 10088-10095

Godden, J. W., Turley, S., Teller, D. C., Adman, E. T., Liu, M. Y., Payne, W. J., & LeGall, J. (1991) *Science*, **253**, 438-442

Goff, S. A., Casson, L. P., & Goldberg, A. L. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6647-6651

Goff, S. A., & Goldberg, A. L. (1985) *Cell*, **41**, 587-595

Gomez, F. A., Chen, J. K., Tanaka, A., Schreiber, S. L., & Whitesides, G. M. (1994) *J. Org. Chem.*, **59**, 2885-2886

Goodell, (1985) *J. Bacteriol.*, **163**, 305-310

Goodlett, D. R., Armstrong, F. B., Creech, R. J., & van Breemen, R. B. (1990) *Anal. Biochem.*, **186**, 116-120

Gorin, G., Wang, S.-F., & Papapavlou, L. (1971) *Anal. Biochem.*, **39**, 113-127

Gram, C. (1884) *Fortschr. Med.*, **2**, 185-189

Grassetti, D. R., & Murray, J. F. (1967) *Arch. Biochem. Biophys.*, **119**, 41-49

Gregory, D. H., & Gerig, J. T. (1991) *J. Comp. Chem.*, **12**, 180-185

Gross, E. (1967) *Methods Enzymol.*, **11**, 238-255

Gross, E., & Witkop, B. (1961) *J. Am. Chem. Soc.*, **83**, 1510-1511

Gross, E., & Witkop, B. (1962) *J. Biol. Chem.*, **237**, 1856-1860

Grütter, M. G., & Matthews, B. W. (1982) *J. Mol. Biol.*, **154**, 525-535

Grütter, M. G., Rine, K. L., & Matthews, B. W. (1979) *J. Mol. Biol.*, **135**, 1029-1032

Grütter, M. G., Weaver, L. H., & Matthews, B. W. (1983) *Nature*, **303**, 828-831

Guckert, J. A., Lowery, M. D., & Solomon, E. I. (1995) *J. Am. Chem. Soc.*, **117**, 2817-2844

Guillon, J. M., Mechulam, Y., Schmitter, J. M., Blanquet, S., & Fayat, G. (1992) *J. Bacteriol.*, **174**, 4294-4301

Hadfield, A. T., Harvey, D. J., Archer, D. B., MacKenzie, D. A., Jeenes, D. J., Radford, S. E., Lowe, G., Dobson, C. M., & Johnson, L. N. (1994) *J. Mol. Biol.*, **243**, 856-872

Hancock, R. E. W., Karunaratne, D. N., & Bernegger-Egli, C. (1994) in *Bacterial Cell Wall. New Comprehensive Biochemistry, vol 27* (Ghuysen, J.-M. & Hakenbeck, R., eds.) pp. 263-279, Elsevier Science B. V., Amsterdam

Hanych, B., Kedzierska, S., Walderich, B., Uznanski, B., & Taylor, A. (1993) *Gene*, **129**, 1-8

Harada, S., Sarma, R., Kakudo, M., Hara, S., & Ikenaka, T. (1981) *J. Biol. Chem.*, **256**, 11600-11602

Hardy, L. W., & Poteete, A. R. (1991) *Biochemistry*, **30**, 9457-9463

Harvey, R. J. (1973) *J. Bacteriol.*, **114**, 309-322

Harz, H., Burgdorf, K., Hõltje, J.-V. (1990) *Anal. Biochem.*, **190**, 120-128

Hash, J. H. (1974) in *Lysozyme* (Osserman, E. F., Canfield, R. E. & Beychok, S., eds.) pp. 95-103, Academic Press, New York

Hayashida, M., Watanabe, K., Muramatsu, T., & Goto, M.-A. (1987) *J. Gen. Microbiol.*, **133**, 1343-1349

Hehre, W. J., Radom, L., Schleyer, P. V., & Pople, J. A. (1986) *Ab-initio Molecular Orbital Theory*, Wiley-Interscience; New York, NY

Hendrickson, W. A., Horton, J. R., & LeMaster, D. M. (1990) *EMBO J.*, **9**, 1665-1672

Henze, U., Sidow, T., Wecke, J., Labischinski, H., & Berger-Bächi, B. (1993) *J. Bacteriol.*, **175**, 1612-1620

Heukeshoven, J., & Dernick, R. (1988) *Electrophoresis*, **9**, 28-32

Hidenori, M., Kawagishi, H., & Usui, T. (1990) *Biochim. Biophys. Acta*, **1035**, 90-96

Hill, R. L., & Brew, K. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol.*, **43**, 411-490

Hinz, H. J. (1983) *Ann. Rev. Biophys. Bioeng.*, **12**, 285-317

Hitoshi, K., Emori, Y., Iba, H., & Okada, Y. (1978) *Proc. Japan Acad. Ser. A*, **54**, 337-340

Hoijer, M. A., Melief, M. J., Van Heldenmeeusen, C. G., Eulderink, F., & Hazenberg, M. P (1995) *Infect. Immun.*, **63**, 1652-1657

Holler, E., Rupley, J. A., & Hess, G. P. (1975a) *Biochemistry*, **14**, 1088-1094

Holler, E., Rupley, J. A., & Hess, G. P. (1975b) *Biochemistry*, **14**, 2377-2385

Höltje, J.-V., & Glauner, B. (1990) *Res. Microbiol.*, **141**, 75-89

Höltje, J.-V., & Tuomanen, E. I. (1991) *J. Gen. Microbiol.*, **137**, 441-454

Höltje, J.-V., Mirelman, D., Sharon, N., & Schwarz, U. (1975) *J. Bacteriol.*, **124**, 1067-1076

Horn, M. J., & Laursen, R. A. (1973) *FEBS Lett.*, **36**, 285-288

Horton, D. (1972) in *Methods in Carbohydrate Chemistry Vol. VI* (Whistler, R. L., & BeMiller, J. N., eds.) pp. 282-285, Academic Press, New York

Hoshino, O., Zehavi, U., Sinaÿ, P., & Jeanloz, R. W. (1972) *J. Biol. Chem.*, **247**, 381-387

Houston, M. E. (1992) Ph.D. Dissertation, University of Waterloo, Ontario, Canada

Houston, M. E., & Honek, J. F. (1989) *J. Chem. Soc. Chem. Commun.*, 761-762

Houston, M. E., Vander Jagt, D. L., & Honek, J. F. (1991) *Bioorg. Med. Chem. Lett.*, **1**, 623-628

Howell, P. L., Almo, S. A., Parsons, M. R., Hajdu, J., & Petsko, G. A. (1992) *Acta Crystallogr. B.*, **48**, 200-207

Hoyle, B. D., & Beveridge, T. J. (1984) *Can. J. Microbiol.*, **30**, 204-211

Ikeda, K., & Hamaguchi, K. (1976) *J. Biochem.*, **79**, 237-247

Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., & Bax, A. (1992) *Science*, **256**, 632-638

Imada, M., & Tsugita, A. (1971) *Nature New Biol.*, **233**, 230-231

Imoto, T., Forster, L. S., Rupley, J. A., & Tanaka, F. (1971) *Proc. Natl. Acad. Sci. USA*, **69**, 1151-1155

Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C., & Rupley, J. A. (1972) in *The Enzymes* (Boyer, P. D., ed.), 3rd ed., **Vol. 7**, pp 665-868, Academic Press, New York

Inglis, A. S., & Edman, P. (1970) *Anal. Biochem.*, **37**, 73-80

Inoue, M., Yamada, H., Yasukochi, T., Miki, T., Horiuchi, T., & Imoto, T. (1992a) *Biochemistry*, **31**, 10322-10330

Inoue, M., Yamada, H., Yasukochi, T., Kuroki, R., Miki, T., Horiuchi, T., & Imoto, T. (1992b) *Biochemistry*, **31**, 5545-5553

Inouye, M., Arnheim, N., & Sternglanz, R. (1973) *J. Biol. Chem.*, **248**, 7247-7252

Isaacs, N. W., Machin, K. J., & Masakuni, M. (1985) *Aus. J. Biol. Sci.*, **38**, 13-22

Ishido, Y., Matsuba, T., Hosono, A., Fujii, K., Tanaka, H., Iwabuchi, K., Isome, S., Maruyama, A., Kikuchi, Y., & Sato, T. (1965) *Bull. Chem. Soc. Jpn.*, **38**, 2019

Ishido, Y., Matsuba, T., Hosono, A., Fujii, K., & Sato, T. (1967) *Bull. Chem. Soc. Jpn.*, **40**, 1007-1009

Jacob, F., & Fuerst, C. R. (1958) *J. Gen. Microbiol.*, **18**, 518-526

Jacobs, C., Huang, L., Bartowsky, E., Normark, S., & Park, J. T. (1994) *EMBO J.*, **13**, 4684-4694

Jacobson, R. H., Zhang, X-J., DuBose, R. F., & Matthews, B. W. (1994) *Nature*, **369**, 761-766

Janin, J., Wodak, S., Levitt, M., & Maigret, B. (1978) *J. Mol. Biol.*, **125**, 357-386

Jeanloz, R. W., Walker, E., & Sinaÿ, P. (1968) *Carbohydr. Res.*, **6**, 184-196

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys.*, **71**, 4546-4553

Jencks, W. P. (1975) *Adv. Enz. Relat. Areas. Mol. Biol.*, **43**, 219-410

Jensen, H. B., & Kleppe, K. (1972) *Eur. J. Biochem.*, **26**, 305-312

Jensen, H. B., Kleppe, G., Schindler, M., & Mirelman, D. (1976) *Eur. J. Biochem.*, **66**, 319-325

Jespers, L., Sonveaux, E., Fastrez, J., Phanapoulos, A., & Davison, J. (1991) *Protein Eng.*, **4**, 485-492

Jespers, L., Sonveaux, E., & Fastrez, J. (1992) *J. Mol. Biol.*, **228**, 529-538

Johnson, L. N. (1966) *Acta Crystallogr.*, **21**, 885-891

Jollès, P., & Jollès, J. (1984) *Mol. Cell. Biochem.*, **63**, 165-189

Jollès, J., Jauregui-Adell, J., Bernier, I., & Jollès, P. (1963) *Biochim. Biophys. Acta*, **78**, 668-689

Jollès, P., Saint-Blancard, J., Charlemagne, D., Dianoux, A.-C., Jollès, J., & Le Baron, J. L. (1968) *Biochim. Biophys. Acta*, **151**, 532-534

Jollès, P., Bernier, I., Berthou, J., Charlemagne, D., Faure, A., Hermann, J., Jollès, J., Périn, J.-P., & Saint-Blancard, J. (1974) in *Lysozyme* (Osserman, E. F., Canfield, R. E. & Beychok, S., eds.) pp. 31-54, Academic Press, New York

Jung, A., Sippel, A. E., Grez, M., & Schütz, G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5759-5763

Kachurin, A. M., Golubev, A. M., Geisow, M. M., Veselkina, O. S., Isaevaivanova, L. S., & Neustroev, K. N. (1995) *Biochem. J.*, **308**, 955-964

Kadner, R. J. (1977) *J. Bacteriol.*, **129**, 207-216

Kaestle, K., Anwer, M. K., Audhya, T. K., & Goldstein, G. (1991) *Tetrahedron Lett.*, **32**, 327-330

Kaiser, E., Colescott, R. L., Bossinger, C. D., & Cook, P. I. (1970) *Anal. Biochem.*, **34**, 595-598

Kang, B. H., Tan, S., & Ho, K. K. (1992) *J. Liq. Chromatogr.*, **15**, 2325-2339

Karlsen, S., Eliassen, B. E., Hansen, L. K., Larsen, R. L., Riise, B. W., Smalas, A. O., Hough, E., & Grinde, B. (1995) *Acta Crystallogr. D.*, **51**, 354-367

Karnbrock, W., Weyher, E., Budisa, N., Huber, R., & Moroder, L. (1996) *J. Amer. Chem. Soc.*, **118**, 913-914

Katchalsky, E., Bichowsky-Slomnitzki, L., & Volcani, B. E. (1953) *Biochem. J.*, **55**, 671-680

Kelly, J. A., Sielecki, A. R., Sykes, B. D., James, M. N. G., & Phillips, D. C. (1979) *Nature (London)*, **282**, 875-878

Kent, L. H. (1957) *Biochem. J.*, 67, 5P

Kikuchi, M., Yamamoto, Y., Taniyama, Y., Ishimaru, K., Yoshikawa, W., Kaisho, Y., & Ikehara, M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 9411-9415

Kim, H. Y., Ghosh, G., Schulman, L. H., Brunie, S., & Jakubowski, H. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 11553-11557

Kimber, B. J., Griffiths, D. V., Birdsall, B., King, R. W., Scudder, P., Feeney, J., Roberts, G. C. K., & Burgen, A. S. V. (1977) *Biochemistry*, **16**, 3492-3500

Kirby, A. J. (1987) *CRC Crit. Rev. Biochem.*, **22**, 923-925

Kirby, A. J. (1995) *Nature Struct. Biol.*, **2**, 1007-1011

Knight, J. A., Hardy, L. W., Rennell, D., Herrick, D., & Poteete, A. R. (1987) *J. Bacteriol.*, **169**, 4630-4636

Koch, A. L. (1983) *Adv. Microb. Physiol.*, **24**, 301-367

Koonin, E. V., & Rudd, K. E. (1994) *Trends Biochem. Sci.*, **19**, 106-107

Koshland, D. E. (1953) *Biol. Rev.*, **28**, 416-436

Kotani, S., Tsujimoto, M., Koga, T., Nagao, S., Tanaka, A., & Kawata, S. (1986) *FASEB J.*, **45**, 2534-2540

Koteswara Rao, G. R., & Burma, D. P. (1971) *J. Biol. Chem.*, **246**, 6474-6479

Koul, A. K., Wasserman, G. F., & Warme, P. K. (1979) *Biochem. Biophys. Res. Commun.*, **89**, 1253-1259

Kraak, J. C., Busch, S., & Poppe, H. (1992) *J. Chromatogr.*, **608**, 257-264

Kravchenko, N. A. (1967) *Proc. R. Soc. London Ser B*, **167**, 429-430

Krishnamoorthy, G., & Prabhananda, B. S. (1982a) *Biochim. Biophys. Acta*, **709**, 53-57

Krishnamoorthy, G., & Prabhananda, B. S. (1982b) *Biochim. Biophys. Acta*, **709**, 234-246

Kronman, M. J. (1989) *CRC Crit. Rev. Biochem. Mol. Biol.*, **24**, 565-667

Krueger, J. M., Bacsik, J., & Garcia-Arrarás, J. (1980) *Am. J. Physiol.*, **238**, E116-E123

Krueger, J. M., Karnovsky, M. L., Martin, S. A., Pappenheimer, J. R., Walter, J., & Biemann, K. (1984) *J. Biol. Chem.*, **259**, 12659-12662

Kumagai, I., Takeda, S., & Miura, K. (1992a) *Proc. Natl. Acad. Sci. USA*, **89**, 5887-5891

Kumagai, I., Sunada, F., Takeda, S., & Miura, K. (1992b) *J. Biol. Chem.*, **267**, 4608-4612

Kumagai, I., Maenaka, K., Sunada, F., Takeda, S., & Miura, K. (1993) *Eur. J. Biochem.*, **212**, 151-156

Kurachi, K., Sieker, L. C., & Jensen, L. H. (1975) *J. Biol. Chem.*, **250**, 7663-7667

Kuramitsu, S., Ikeda, K., Hamaguchi, K., Fujio, H., Amano, T., Miwa, S., & Nishina, T. (1974) *J. Biochem.*, **76**, 671-683

Kuroki, R., Yamada, H., Moriyama, T., & Imoto, T. (1986) *J. Biol. Chem.*, **261**, 13571-13574

Kuroki, R., Taniyama, Y., Seko, C., Nakamura, H., Kikuchi, M., Ikehara, M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6903-6907

Kuroki, R., Weaver, L. H., & Matthews, B. W. (1993) Science, **262**, 2030-2033

Kuroki, R., Weaver, L. H., & Matthews, B. W. (1995) *Nature Struct. Biol.*, **2**, 1007-1011

Kusumoto, S., Imoto, M., Ogiku, T., & Shiba, T. (1986) *Bull. Chem. Soc. Jpn.*, **59**, 1419-1423

Kusumoto, S., Yamamoto, K., Imoto, M., Inage, M., Tsujimoto, M., Kotani, S., & Shiba, T. (1986) *Bull. Chem. Soc. Jpn.*, **59**, 1411-1417

Labischinski, J., Barnickel, G., Bradaczek, H., & Giesbrecht, P. (1979) *Eur. J. Biochem.*, **95**, 147-155

Labischinski, J., & Johannsen, L. (1986) in *Biological Properties of Petidoglycan* (Seidl, P. H., & Schleifer, K. H., eds.), pp. 37-42, Walter de Gruyter, Berlin

Labischinski, H., & Maidhof, H. (1994) in *Bacterial Cell Wall. New Comprehensive Biochemistry, vol 27* (Ghuysen, J.-M. & Hakenbeck, R., eds.) pp. 23-38, Elsevier Science B. V., Amsterdam

Laemmli, U. K. (1970) *Nature*, **227**, 680-685

Leaback, D. H., & Walker, P. G. (1957) *J. Chem. Soc.*, 4754-4760

Leclerc, D., & Asselin, A. (1989) *Can. J. Microbiol.*, **35**, 749-753

Lederer, E. (1980) *J. Med. Chem.*, **23**, 819-825

Lee, T. D., & Shively, J. E. (1990) *Methods Enzymol.*, **193**, 361-374

Lee, C. P., Seong, B. L., & RajBhandary, U. L. (1991) *J. Biol. Chem.*, **266**, 18012-18017

Lefrancier, P., & Lederer, E. (1987) *Pure & Appl. Chem.*, **59**, 449-454

Lefrancier, P., Derrien, M., Lederman, I., Nief, F., Choay, J., & Lederer, E. (1978) *Int. J. Peptide Protein Res.*, **11**, 289-296

Lefrancier, P., & Lederer, E. (1981) in *Progress in the Chemistry of Organic Natural Products* (Herz, W., Grisebach, H., & Kirby, G. W., eds.) pp. 1-47, Springer-Verlag Wien, New York

Legler, G., Sinnott, M. L., & Withers, S. G. (1980) *J. Chem. Soc. Perkin Trans. II*, 1376-1383

Lehrer, S. S., & Fasman, G. D. (1966) *Biochem. Biophys. Res. Commun.*, **23**, 133-138

Lehrer, S. S., & Fasman, G. D. (1967) *J. Biol. Chem.*, 242, 4644-4645

Lescar, J., Souchon, H., & Alzari, P. M. (1994) *Protein Sci.*, **3**, 788-798

Levitt, M., Sander, C., & Stern, P. S. (1985) *J. Mol. Biol.*, **181**, 423-447

Levy, H. M., Leber, P. D., & Ryan, E. M. (1963) *J. Biol. Chem.*, **238**, 3654-3659

Li, E., Qian, S., Nader, L., Yang, N. C., d'Avignon, A., Sacchettini, J. C., & Gordon, J. I. (1989) *J. Biol. Chem.*, **264**, 17041-17048

Lian, C., Le, H., Montez, B., Patterson, J., Harrell, S., Laws, D., Matsumura, I., Pearson, J., & Oldfield, E. (1994) *Biochemistry*, **33**, 5238-5245

Lichenstein, H. S., Hastings, A. E., Langley, K. E., Mendiaz, E. A., Rohde, M. F., Elmore, R., & Zukowski, M. M. (1990) *Gene*, **88**, 81-86

Lin, X., Krudy, G. A., Howarth, J., Brito, R. M. M., Rosevear, P. R., & Putkey, J. A. (1994) *Biochemistry*, **33**, 14434-14442

Link, T. P., & Stark, G. R. (1968) *J. Biol. Chem.*, **243**, 1082-1086

Liras, J. L., & Anslyn, E. V. (1994) *J. Am. Chem. Soc.*, **116**, 2645-2646

Liu, T. Y., & Gotschlich, E. C. (1967) *J. Biol. Chem.*, **242**, 471-476

Loessner, M. J., Wendlinger, G., & Scherer, S. (1995) *Molecular Microbiology*, **16**, 1231-1241

Lowe, G. (1967) *Proc. R. Soc. London Ser B*, **167**, 431-434

Luck, L. A., & Falke, J. J. (1991a) *Biochemistry*, **30**, 4248-4256

Luck, L. A., & Falke, J. J. (1991b) *Biochemistry*, **30**, 4257-4261

Luck, L. A., & Falke, J. J. (1991c) *Biochemistry*, **30**, 6484-6490

Lumb, K. J., & Dobson, C. M. (1992) *J. Mol. Biol.*, **227**, 9-14

Lumb, K. J., Aplin, R. T., Radford, S. E., Archer, D. B., Jeenes, D. J., Lambert, N., MacKenzie, D. A., Dobson, C. M., & Lowe, G. (1992) *FEBS Lett.*, **296**, 153-157

Lumb, K. J., Cheetham, J. C., & Dobson, C. M. (1994) *J. Mol. Biol.*, **235**, 1072-1087

Lutcke, H., High, S., Romisch, K., Ashford, A. J., & Dobberstein, B. (1992) *EMBO J.*, **11**, 1543-1551

Lyne, J. E., Carter, D. C., He, X., Stubbs, G., & Hash, J. H. (1990) *J. Biol. Chem.*, **265**, 6298-6930

Macquaire, F., Baleux, F., Huynh-Dinh, T., Rouge, D., Neuman, J.-M., & Sanson, A. (1993) *Biochemistry*, **32**, 7244-7254

Madhusudan, & Vijayan, M. (1992) *Protein Eng.*, **5**, 339-404

Maeda, H. (1980) *J. Biochem.*, **88**, 1185-1191

Maenaka, K., Kawai, G., Watanabe, K., Sunada, F., & Kumagai, I. (1994) *J. Biol. Chem.*, **269**, 7070-7075

Malcolm, B. A., Rosenberg, S., Corey, M. J., Allen, J. S., de Baetselier, A., & Kirsch, J. F. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 133-137

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Manojloviv-Muir, L., Muir, K. W., & Solomun, T. (1977) *Inorg. Chim. Acta*, **22**, 69-74

Markham, G. D., & Bock, C. W. (1995) *J. Phys. Chem.*, **99**, 10118-10129

Martin, S. A., Karnovsky, M. L., Krueger, J. M., Pappenheimer, J. R., & Biemann, K. (1984) *J. Biol. Chem.*, **259**, 12652-12658

Matsuhashi, M. (1994) in *Bacterial Cell Wall. New Comprehensive Biochemistry, vol 27* (Ghuysen, J.-M. & Hakenbeck, R., eds.) pp. 55-71, Elsevier Science B. V., Amsterdam

Matsui, H., Tanaka, Y., Brewer, C. F., Blanchard, J. S., & Hehre, E. J. (1993) *Carbohydr. Res.*, **250**, 45-56

Matsumura, I., & Kirsch, J. F. (1996a) *Biochemistry*, **35**, 1881-1889

Matsumura, I., & Kirsch, J. F. (1996b) *Biochemistry*, **35**, 1890-1896

Matsushima, Y., & Park, J. T. (1962) *J. Org. Chem.*, **27**, 3581-3584

Matta, M. S., Henderson, P. A., & Patrick, T. B. (1981) *J. Biol. Chem.*, **256**, 4172-4174

Matthews, B. W. (1993) *Curr. Opin. Struct. Biol.*, **3**, 589-593

Matthews, B. W., & Remington, S. J. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 4178-4182

Matthews, B. W., Grütter, M. G., Anderson, W. F., & Remington, S. J. (1981a) *Nature*, **290**, 334-335

Matthews, B. W., Remington, S. J., Grütter, M. G., & Anderson, W. F. (1981b) *J. Mol. Biol.*, **147**, 545-558

Maurel, P., & Douzou, P. (1976) *J. Mol. Biol.*, **102**, 253-264

McCammon, J. A., Gelin, B. R., Karplus, M., & Wolynes, P. G. (1976) *Nature*, **262**, 325-326

McKenzie, H. A., & White, F. H. (1987) *Biochem. Int.*, **14**, 347-356

McMacken, R., Mantei, N., Butler, B., Joyner, A., & Echols, H. (1970) *J. Mol. Biol.*, **49**, 639-655

McPherson, A. (1982) *Preparation and Analysis of Protein Crystals*, Krieger Publishing

Meador, W. E., Means, A. R., & Quiocho, F. A. (1992) *Science*, **257**, 1251-1255

Meinnel, T., & Blanquet, S. (1995) *J. Bacteriol.*, **177**, 1883-1887

Melly, M. A., McGee, Z. A., & Rosenthal, R. S. (1984) *J. Infect. Dis.*, **149**, 378-386

Mendelson, N. H., & Thwaites, J. J. (1989) *J. Bacteriol.*, **171**, 1055-1062

Merrifield, R. B. (1966) *J. Am. Chem. Soc.*, **88**, 2149-2154

Merser, C., Sinaÿ, P., & Adam, A. (1975) *Biochem. Biophys. Res. Commun.*, **66**, 1316-1322

Merten, H., & Brossmer, R. (1989) *Carbohydr. Res.*, **191**, 144-149

Messerschmidt, A., Ladenstein, R., Huber, R., Bolognesi, M., Avigliano, L., Petruzzelli, R., Rossi, A., & Finazzi-Agro, A. (1992) *J. Mol. Biol.*, **224**, 179-205

Miles, A. M., & Smith, R. L. (1993) *Biochemistry*, **32**, 8168-8178

Mirelman, D., Bracha, R., & Sharon, N. (1974) *Biochemistry*, **13**, 5045-5053

Mintz, G., Herbold, D. R., & Glaser, L. (1975) *Anal. Biochem.*, **66**, 272-278

Mirelman, D. (1979) in *Bacterial Outer Membranes*, pp. 115-166, John Wiley & Sons, New York

Mirelman, D., Kleppe, G., & Jensen, H. B. (1975) *Eur. J. Biochem.*, **55**, 369-373

Mohd. Salleh, H. (1994) Ph.D. Dissertation, University of Waterloo, Ontario, Canada

Moir, A., & Smith, D. A. (1990) *Ann. Rev. Microbiol.*, **44**, 531-553

Morita, T., Hara, S., & Matsushima, Y. (1978) *J. Biochem.*, **83**, 893-903

Morton, A., & Matthews, B. W. (1995) *Biochemistry*, **34**, 8576-8588

Moser, I. Pittner, F., Dworsky, P. (1988) *J. Biochem. Biophys. Methods*, **17**, 249-252

Mosig, G., Lin., G. W., Franklin, J., Fan, & W.-H. (1989) *New Biologist*, **1**, 171-179

Moskovitz, J., Rahman, M. A., Strassman, J., Yancey, S. O., Kushner, S. R., Brot, N., & Weissbach, H. (1995) *J. Bacteriol.*, **177**, 502-507

Mott, J. E., Grant, R. A., Ho, Y., & Platt, T. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 88-92

Moult, J., Eshdat, Y., & Sharon, N. (1973) *J. Mol. Biol.*, **75**, 1-4

Mulvey, R. S., Gualtieri, R. J., & Beychok, S. (1974) in *Lysozyme* (Osserman, E. F., Canfield, R. E. & Beychok, S., eds.) pp. 281-300 , Academic Press, New York

Muraki, M., Morikawa, M., Jigami, Y., & Tanaka, H. (1989) *Eur. J. Biochem.*, **179**, 573-579

Muraki, M., Harata, K., Hayashi, Y., Machida, M., & Jigami, Y. (1991) *Biochim. Biophys. Acta*, **1079**, 229-237

Muraki, M., Harata, K., Sugita, N., & Sato, K.-I. (1996) *Biochemistry*, **35**, 13562-13567

Murialdo, H., Davidson, A., Chow, S., & Gold, M. (1987) *Nucl. Acids Res.*, **15**, 119-140

Murphy, L. M., Strange, R. W., Karlsson, B. G., Lundberg, L. G., Pascher, T., Reinhammar, B., & Hasnain, S. S. (1993) *Biochemistry*, **32**, 1965-1975

Nakabayashi, S., Warren, C. D., & Jeanloz, R. W. (1986) *Carbohydr. Res.*, **150**, C7-C10

Nanjo, F., Sakai, K., Usui, T., Takai, I., & Ishido, Y. (1988a) *J. Carbohydr. Chem.*, **7**, 67-82

Nanjo, F., Sakai, K., & Usui, T. (1988b) *J. Biochem.*, **104**, 255-258

Naumann, D., Barnickel, G., Bradaczek, H., Labischinski, H., & Giesbrecht, P. (1982) *Eur. J. Biochem.*, **125**, 505-515

Nitta, K., & Sugai, S. (1989) *Eur. J. Biochem.*, **182**, 111-118

Noller, E. C., & Hartsell, S. E. (1961) *J. Bacteriol.*, **81**, 482-491

Nogami, N., Sugeta, H., & Miyazawa, T. (1975) *Bull. Chem. Soc. Jpn.*, **48**, 3573-3575

Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., & Schultz, P. G. (1989) *Science*, **244**, 182-188

Nozaki, Y. (1972) *Methods Enzymol.*, **26**, 43-50

Ogawa, T., Beppu, K., & Nakabayashi, S. (1981) *Carbohydr. Res.*, **93**, C6-C9

Ohkubo, T., Taniyama, Y., Kikuchi, M. (1991) *J. Biochem.*, **110**, 1022-1029

Ohno, N., & Morrison, D. C. (1989a) *J. Biol. Chem.*, **264**, 4434-4441

Ohno, N., & Morrison, D. C. (1989b) *Eur. J. Biochem.*, **186**, 621-627

Ohno, N., Tanida, N., Yadomae, T. (1991) *Carbohydr. Res.*, **214**, 115-130

Oldmixon, E. J., Glauser, S., & Higgins, M. L. (1974) *Biopolymers*, **13**, 2037-2060

Osawa, T. (1966) *Carbohydr. Res.*, **1**, 435-443

Osawa, T., & Jeanloz, R. W. (1965) *J. Org. Chem.*, **30**, 448-450

Osawa, T., & Nakazawa, Y. (1966) *Biochim. Biophys. Acta*, **130**, 56-63

Osserman, E. F., Canfield, R. E. & Beychok, S. (eds.) (1974) in *Lysozyme*, Academic Press, New York

Oyanagi, K., & Kuchitsu, K. (1978) *Bull. Chem. Soc. Jpn.*, **51**, 2243-2248

Pace, C. N., Shirley, B. A., & Thomson, J. A. (1990) in *Protein Structure - A Practical Approach* (Creighton, T. E., ed.) pp. 311-330, IRL Press, Oxford

Pakula, T. M., Savilahti, H., & Bamford, D. H. (1989) *Eur. J. Biochem.*, **180**, 149-152

Park, J. T. (1993) *J. Bacteriol.*, **175**, 7-11

Park, J. T., & Hancock, R. (1960) *J. Gen. Microbiol.*, **22**, 249-258

Park, J. T., & Johnson, M. J. (1949) *J. Biol. Chem.*, **181**, 149-151

Parsons, J. F., & Armstrong, R. N. (1996) *J. Am. Chem. Soc.*, **118**, 2295-2296

Parsons, S. M., & Raftery, M. A. (1972a) *Biochemistry*, **11**, 1623-16291633-1638

Parsons, S. M., & Raftery, M. A. (1972b) *Biochemistry*, **11**, 1633-1638

Patt, S. L., Dolphin, D., & Sykes, B. D. (1974) in *Lysozyme* (Osserman, E. F., Canfield, R. E., & Beychok, S., eds.) pp. 229-237, Academic Press, New York

Patt, S. L., Baldo, J. H., Boekelheide, K., Weisz, G., & Sykes, B. D. (1978) *Can. J. Biochem.*, **56**, 624-629

Paulsen, H. (1982) *Angew. Chem. Int. Ed. Engl.*, **21**, 155-224

Pecht, I., Teichberg, V. I., & Sharon, N. (1970) *FEBS Letters*, **10**, 241-245

Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennent, G. A., Soutar, A. K., Totty, N., Nguyen, O., Blake, C. C. F., Terry, C. J., Feest, T. G., Zalin, A. M., & Hsuan, J. J. (1993) *Nature*, **362**, 553-557

Perkins, S. J., Johnson, L. N., Machin, P. A., & Phillips, D. C. (1979) *Biochem. J.*, **181**, 21-36

Perkins, S. J., Johnson, L. N., Phillips, D. C., & Dwek, R. A. (1981a) *Biochem. J.*, **193**, 553-572

Perkins, S. J., Johnson, L. N., Phillips, D. C., & Dwek, R. A. (1981b) *Biochem. J.*, **193**, 553-572

Phillips, D. C. (1966) *Sci. Am.*, **215**, 78-90

Phillips, D. C. (1967) *Proc. Natl. Acad. Sci. USA*, **57**, 484-495

Phillips, D. C. (1974) in *Lysozyme* (Osserman, E. F., Canfield, R. E. & Beychok, S., eds.) pp. 9-30, Academic Press, New York

Pincus, M. R., & Scheraga, H. A. (1979) *Macromolecules*, **12**, 633-644

Pincus, M. R., & Scheraga, H. A. (1981) *Biochemistry*, **20**, 3960-3965

Piszkiewicz, D., & Bruice, T. C. (1968a) *J. Am. Chem. Soc.*, **89**, 6237-6243

Piszkiewicz, D., & Bruice, T. C. (1968b) *J. Am. Chem. Soc.*, **90**, 2156-2163

Piszkiewicz, D., & Bruice, T. C. (1968c) *J. Am. Chem. Soc.*, **90**, 5844-5848

Pollock, J. J., & Sharon, N. (1970) *Carbohydr. Res.*, **13**, 211-224

Pooley, H. M., & Karamata, D. (1994) in *Bacterial Cell Wall. New Comprehensive Biochemistry, vol 27* (Ghuysen, J.-M. & Hakenbeck, R., eds.) pp. 187-198, Elsevier Science B. V., Amsterdam

Post, C. B., & Karplus, M. (1986) *J. Am. Chem. Soc.*, **108**, 1317-1319

Post, C. B., Brooks, B. R., Karplus, M., Dobson, C. M., Artymiuk, P. J., Cheetham, J. C., & Phillips, D. C. (1986) *J. Mol. Biol.*, **190**, 455-479

Poteete, A. R., Rennell, D., & Bouvier, S. E. (1992) *Proteins: Struct. Funct. Genet.*, **13**, 38-40

Prager, E. M., Wilson, A. C., & Arnheim, N. (1974) *J. Biol. Chem.*, **249**, 7295-7297

Price, J. A. R., & Pethig, R. (1986) *Biochim. Biophys. Acta*, **889**, 128-135

Primosigh, J., Pelzer, H., Maas, D., & Weidel, W. (1961) *Biochim. Biophys. Acta*, **46**, 68-80

Privalov, P. L. (1982) *Adv. Protein Chem.*, **35**, 1-104

Privalov, P. L. (1989) *Annu. Rev. Biophys. Biophys. Chem.*, **18**, 47-69

Quiocho, F. A. (1986) *Annu. Rev. Biochem.*, **55**, 287-315

Quiocho, F. A. (1988) *Curr. Topics Microbiol. Immunol.*, **126**, 135-148

Quiocho, F. A., Wilson, D.K., & Vyas, N. K. (1989) *Nature*, **340**, 404-407

Quintela, J. C., Pittenauer, E., Allmaier, G., Arán, V., & de Pedro, M. A. (1995a) *J. Bacteriol.*, **177**, 4947-4962

Quintela, J. C., Caparrós, M., & de Pedro, M. A. (1995b) *FEMS Microbiol. Lett.*, **125**, 95-100

Radford, S. E., Woolfson, D. N., Martin, S., Lowe, G., & Dobson, C. M. (1991) *Biochem. J.*, **273**, 211-217

Raftery, M. A., Rand-Meir, T., Dahlquist, F. W., Parsons, S. M., Borders, C. L., Jr., Wolcott, R. G., Beranek, W., Jr., & Jao, L. (1969) *Anal. Biochem.*, **30**, 427-435

RajBhandary, U. L. (1994) *J. Bacteriol.*, **176**, 547-552

Rand-Meir, T., Dahlquist, F. W., & Raftery, M. A. (1969) *Biochemistry*, **8**, 4206-4214

Reedy, B. J., & Blackburn, N. J. (1994) *J. Am. Chem. Soc.*, **116**, 1924-1931

Remington, S. J., & Matthews, B.W. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 2180-2184

Rennell, D., & Poteete, A. R. (1985) *Virology*, **143**, 280-289

Rennell, D., & Poteete, A. R. (1989) *Genetics*, **123**, 431-440

Rennell, D., Bouvier, S. E., Hardy, L. W., & Poteete, A. R. (1991) *J. Mol. Biol.*, **222**, 67-87

Richard, C., Demarly, A., Han, K.-K., & Dautrevaux, M. (1984) *Int. J. Biochem.*, **16**, 727-732

Robinson, M., Lilley, R., Little, S., Emtage, J. S., Yarranton, G., Stephens, P., Milligan, A., Eaton, M., & Humphreys, G. (1984) *Nucleic Acids Res.*, **12**, 6663-6671

Roderick, S. L., & Matthews, B. W. (1993) *Biochemistry*, **32**, 3907-3912

Rogers, H. J., Perkins, H. R., & Ward, J. B. (1980) in *Microbial Cell Walls and Membranes*, Ch. 6, 8, 11, Chapman & Hall, London

Romeis, T., Vollmer, W., & Höltje, J.-V. (1993) *FEMS Microbiol. Lett.*, **111**, 141-146

Romera, A., López, R., & García, P. (1990) *J. Virol.*, **64**, 137-142

Rosenberg, S., & Kirsch, J. F. (1981) *Biochemistry*, **20**, 3196-3204

Rossi, G. L., Holler, E., Kumar, S., Rupley, J. A., & Hess, G. P. (1969) *Biochem. Biophys. Res. Commun.*, **37**, 757-766

Rossmann, M. G., & Argos, P. (1976) *J. Mol. Biol.*, **105**, 75-95

Roten, C.-A. H., & Karamata, D. (1992) *Experientia*, **48**, 921-931

Rupley, J. A. (1964) *Biochim. Biophys. Acta*, **83**, 245-255

Rupley, J. A. (1967) *Proc. R. Soc. London, Ser B*, **167**, 416-430

Rupley, J. A., & Gates, V. (1967) *Proc. Natl. Acad. Sci. USA*, **57**, 496-510

Rupley, J. A., Gates, V., & Bilbrey, R. (1968) *J. Am. Chem. Soc.*, **90**, 5633-5635

Saedi, M. S., Garvey, K. J., Ito, J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 955-958

Sailer, M., Helms, G. L., Henkel, T., Niemczura, W. P., Stiles, M. E., & Vederas, J. C. (1993) *Biochemistry*, **32**, 310-318

Salgado, J., Jimenez, H. R., Moratal, J. M., Kroes, S., Warmerdam, G. C. M., & Canters, G. W. (1996) *Biochemistry*, **35**, 1810-1819

Salton, M. R. J. (1952a) Nature, **170**, 746-747

Salton, M. R. J. (1952b) *Biochim. Biophys. Acta*, **8**, 510-519

Salton, M. R. J. (1953) *Biochim. Biophys. Acta*, **10**, 512-523

Salton, M. R. J. (1963) *J. Gen. Microbiol.*, **30**, 223-235

Salton, M. R. J. (1994) in *Bacterial Cell Wall. New Comprehensive Biochemistry, vol 27* (Ghuysen, J.-M. & Hakenbeck, R., eds.) pp. 1-22 , Elsevier Science B. V., Amsterdam

Salton, M. R. J.. & Ghuysen, J.-M. (1960) *Biochim. Biophys. Acta*, **45**, 355-363

Sambrook, T., Fritsch, E., & Maniatis, T. (1989) in *Molecular Cloning - A Laboratory Manual 2nd ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Sanger, F., Coulson, A. R., Hong, G. F., & Peterson, G. B. (1982) *J. Mol. Biol.*, **162**, 729-773

Sanz, J. M., & García, J. L. (1990) *Eur. J. Biochem.*, **187**, 409-416

Sanz, J. M., García, P., & García, J. L. (1992) *Biochemistry*, **31**, 8495-8499

Sanz, J. M., García, J. L., Laynez, J., Usobiaga, P., & Menédez, M. (1993) *J. Biol. Chem.*, **268**, 6125-6130

Sarin, V. K., Kent, S. B. H., Tam, J. P., & Merrifield, R. B. (1981) *Anal. Biochem.*, **117**, 147-157

Sarma, R., & Bott, R. (1977) *J. Mol. Biol.*, **113**, 555-565

Sarma, R., Harada, S., Tanaka, N., Kakudo, M., Hara, S., & Ikenaka, T. (1979) *J. Biochem.*, **86**, 1765-1771

Satishchandran, C., Taylor, J. C., & Markham, G. D. (1990) *J. Bacteriol.*, **172**, 4489-4496

Schindler, M., & Sharon, N. (1976) *J. Biol. Chem.*, **251**, 4330-4335

Schindler, M., Assaf, Y., Sharon, N., & Chipman, D. M. (1977a) *Biochemistry*, **16**, 423-431

Schindler, M., Mirelman, D., & Sharon, N. (1977b) *Biochim. Biophys. Acta*, **482**, 386-392

Schleifer, K. H., & Kandler, O. (1972) *Bacteriol. Rev.*, **36**, 407-477

Schmid, F. X. (1990) *in Protein Structure - A Practical Approach* (Creighton, T. E., ed.) pp. 251-285, IRL Press, Oxford

Schmidt, R. R. (1986) *Angew. Chem. Int. Ed. Engl.*, **25**, 212-235

Schoner, B. E., Belagaje, R. M., & Schoner, R. G. (1987) *Methods Enzymol.*, **153**, 401-416

Schoner, B. E., Belagaje, R. M., & Schoner, R. G. (1990) *Methods Enzymol.*, **185**, 94-103

Secemski, I. I., & Lienhard, G. E. (1974) *J. Biol. Chem.*, **249**, 2932-2938

Secemski, I. I., Lehrer, S. S., & Lienhard, G. E. (1972) *J. Biol. Chem.*, **247**, 4740-4748

Seidl, P. H., & Schleifer, K. H. (1986) in *Biological Properties of Petidoglycan* (Seidl, P. H., & Schleifer, K. H., eds.), pp. 1-20, Walter de Gruyter, Berlin

Sekiguchi, M., & Cohen, S. S. (1964) *J. Mol. Biol.*, **8**, 638-659

Sen, Z., & Karnovsky, M. L. (1984) *Infect. Immun.*, **43**, 937-941

Senn, H., Guerlesquin, F., Bruschi, M., & Wüthrich, K. (1983) *Biochim. Biophys. Acta*, **748**, 194-204

Sharon, N. (1967) *Proc. R. Soc. London Ser B*, **167**, 402-415

Sharon, N., & Jeanloz, R. W. (1964) *Experientia*, **20**, 253-254

Sharon, N., Eshdat, Y., Maoz, I., Bernstein, Y., Prager, E. M., & Wilson, A. C. (1974) *Isr. J. Chem.*, **12**, 591-603

Shearman, C. A., Jury, K. L., & Gasson, M. J. (1994) *Appl. Environ. Microbiol.*, **60**, 3063-3073

Sheppard, R. C. (1979) in *Peptides: Structure and Biological Function* (Gross, E. & Meienhofer, J., eds.) pp. 577-585, Pierce, Rockford, Illinois

Shirota, F. N., Nagasawa, H. T., & Elberling, J. A. (1977a) *J. Med. Chem.*, **20**, 1176-1181

Shirota, F. N., Nagasawa, H. T., & Elberling, J. A. (1977b) *J. Med. Chem.*, **20**, 1623-1627

Shively, J. E., Hawke, D., & Jones, B. N. (1982) *Anal. Biochem.*, **120**, 312-322

Shockman, G. D., & Höltje, J.-V. (1994) in *Bacterial Cell Wall. New Comprehensive Biochemistry, vol 27* (Ghuysen, J.-M. & Hakenbeck, R., eds.) pp. 131-166, Elsevier Science B. V., Amsterdam

Shugar, D. (1952) *Biochim. Biophys. Acta*, **8**, 302-309

Sinha, R. K., & Rosenthal, R. S. (1980) *Infect. Immun.*, **29**, 914-925

Sinnott, M. L. (1987) in *Enzyme Mechanisms* (Page, M. I., & Williams, A., eds.) pp 259-301, Royal Society of Chemistry, London

Sinnott, M. L. (1990) *Chem. Rev.*, **90**, 1171-1202

Sinnott, M. L. (1993) *Bioorg. Chem.*, **21**, 34-40

Sinnott, M. L., & Jencks, W. P. (1980) *J. Am. Chem. Soc.*, **102**, 2026-2032

Skoog, D. A., & West, D. M. (1982) in *Fundamentals of Analytical Chemistry 4th Ed.*, pp. 73-77, Saunders College Publishing, New York

Smith, L. E. H., Mohr, L. H., & Raftery, M. A. (1973) *J. Am. Chem. Soc.*, **95**, 7497-7500

Smith-Gill, S. J., Rupley, J. A., Pincus, M. R., Carty, R. P., & Scheraga, H. A. (1984) *Biochemistry*, **23**, 993-997

Song, H., Inaka, K., Maenaka, K., & Matsushima, M. (1994) *J. Mol. Biol.*, **244**, 522-540

Soumillion, P., & Fastrez, J. (1992) *Biochem. J.*, **286**, 187-191

Soumillion, P., Jespers, L, Vervoort, J., & Fastrez, J. (1995) *Protein Eng.*, **8**, 451-456

Spande, T. F., Witkop, B., Degani, Y., & Patchornik, A. (1970) *Adv. Protein. Chem.*, **24**, 97-260

Sparrow, J. T. (1976) *J. Org. Chem.*, **41**, 1350-1353

Spassov, V. Z., Karshikov, A. D., & Atanasov, B. P. (1989) *Biochim. Biophys. Acta*, **999**, 1-6

Spurlino, J. C., Lu, G. Y., & Quiocho, F. A. (1991) *J. Biol. Chem.*, **266**, 5202-5219

Stark, M. J. R. (1987) *Gene*, **51**, 255-267

Sternberg, M. J. E., Grace, D. E. P., & Phillips, D. C. (1979) *J. Mol. Biol.*, **130**, 231-253

Stewart, J. J. P. (1989) *J. Comput. Chem.*, **10**, 209-220; 221-264

Stewart, J. M., & Young, J. D. (1984) in *Solid Phase Peptide Synthesis 2nd Ed.*, Pierce Chemical Company, Rockford, Illinois

Still, W. C., Kahn, M., & Mitra, A. (1978) *J. Org. Chem.*, **43**, 2923-2925

Stormo, G. D., Schneider, T. D., Gold. L., & Ehrenfeucht, A. (1982) *Nucleic Acids Res.*, **10**, 2997-3011

Stoscheck, C. M. (1990) *Methods Enzymol.*, **182**, 50-68

Strange, R. E., & Kent, L. H. (1959) *Biochem. J.*, **71**, 333-339

Street, I. P., Kempton, J. B., & Withers, S. G. (1992) *Biochemistry*, **31**, 9970-9978

Streisinger, G., Mukai, F., Dreyer, W. J., Miller, B., & Horiuchi, S. (1961) *Cold Spring Harbor Symp. Quant. Biol.*, **26**, 25-30

Strominger, J. L., & Tipper, D. J. (1974) in *Lysozyme* (Osserman, E. F., Canfield, R. E. & Beychok, S., eds.) pp. 169-184, Academic Press, New York

Strynadka, N. C., & James, M. N. G. (1991) *J. Mol. Biol.*, **220**, 401-424

Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol.*, **185**, 60-89

Sykes, B. D., & Hull, W. E. (1978) *Methods Enzymol.*, **49**, 270-295

Sykes, B. D., & Weiner, J. H. (1980) *Magn. Reson. Biol.*, **1**, 171-196

Szilágyi, L., & Forgó, P. (1991) *Biophys. Chem.*, **40**, 89-96

Tabor, C. W., & Tabor, H. (1976) *Annu. Rev. Biochem*, **45**, 285-306

Tam, J. P., Heath, W. F., & Merrifield, R. B. (1986) *J. Am. Chem. Soc.*, **108**, 5242-5251

Tarr, G. E., & Crabb, J. W. (1983) *Anal. Biochem.*, **131**, 99-107

Taylor, A. (1970) *Biochem. Biophys. Res. Commun.*, **41**, 16-24

Taylor, A. (1971) *Nature New Biol.*, **234**, 144-145

Taylor, A., & Gorazdowska, M. (1974) *Biochim. Biophys. Acta*, **342**, 133-136

Taylor, A., Das, B. C., & Van Heijenoort, J. (1975) *Eur. J. Biochem.*, **53**, 47-54

Taylor, H. C., Richardson, D. C., Richardson, J. S., Wlodawer, A., Komoriya, A., & Chaiken, I. M. (1981) *J. Mol. Biol.*, **149**, 313-317

Taylor, H. C., Komoriya, A., & Chaiken, I. M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6423-6426

Teahan, C. G., McKenzie, H. A., Shaw, D. C., & Griffiths, M. (1991) *Biochem. Int.*, **24**, 85-95

Templin, M. F., Edwards, D. H., & Hõltje, J.-V. (1992) *J. Biol. Chem.*, **267**, 20039-20043

Tenforde, T., Fawwaz. R. A., Freeman, N. K., & Castagnoli, N. (1972) *J. Org. Chem.*, **37**, 3373-3374

Termin, A., & Schmidt, R. R. (1992) *Liebigs Ann. Chem.*, 527-533

Tesser, G. I., & Nivard, J. F. (1964) *Rec. Trav. Chim.*, **83**, 53-66

Thompson, A., & Wolfrom, M. L. (1963) in *Methods in Carbohydrate Chemistry Vol. II* (Whistler, R. L., & Wolfrom, M. L., eds.) pp. 215-220, Academic Press, New York

Thunnissen, A.-M. W. H., Dijkstra, A. J., Kalk, K. H., Rozeboom, H. J., Engel, H., Keck, W., & Kijkstra, B. W. (1994) *Nature*, **367**, 750-753

Thunnissen, A.-M. W. H., Rozeboom, H. J., Kalk, K. H., & Dijkstra, B. W. (1995a) *Biochemistry*, **34**, 12729-12737

Thunnissen, A.-M. W. H., Isaacs, N. W., & Dijkstra, B. W. (1995b) *Proteins: Struct. Funct. Genet.*, **22**, 245-258

Timasheff, S.N., & Arakawa, T. (1990) in *Protein Structure-A Practical Approach* (Creighton, T.E., ed.), pp. 331-345, IRL Press, Oxford

Timkovich, R. (1977) *Anal. Biochem.*, **79**, 135-143]

Tipper, D. J., & Strominger, J. L. (1966) *Biochem. Biophys. Res. Commun.*, **22**, 48-56

Tipper, D. J., Strominger, J. L., & Enxign, J. C. (1967) *Biochemistry*, **6**, 906-920

Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A., & Neurath, H. (1972) *Biochemistry*, **11**, 2427-2435

Travis, J., & Salvesen, G. S. (1983) *Ann. Rev. Biochem.*, **52**, 655-709

Trupin, J., Dickerman, H., Nirenberg, M., & Weissbach, H. (1966) *Biochem. Biophys. Res. Commun.*, **24**, 50-55

Tsugita, A., & Inouye, M. (1968) *J. Mol. Biol.*, **37**, 202-212

Tsugita, A., Inouye, M., Terzaghi, E., & Streisinger, G. (1968) *J. Biol. Chem.*, **243**, 391-397

Turner, M. A., & Howell, P. L. (1995) *Protein Sci.*, **4**, 442-449

Ubbink, M., Campos, A. P., Teixeira, M., Hunt, N. I., Hill, H. A. O., & Canters, G. W. (1994) *Biochemistry*, **33**, 10051-10059

Ursinus, A., & Höltje, J.-V. (1994) *J. Bacteriol.*, **176**, 338-343

Usui, T., Hayashi, Y., Nanjo, F., Sakai, K, & Ishido, Y. (1987) *Biochim. Biophys. Acta*, **923**, 302-309

Usui, T., Hayashi, Y., Nanjo, F., & Ishido, Y. (1988) *Biochim. Biophys. Acta*, **953**, 179-184

Vanderwinkel, E., Depauw, P., Philipp, D., Tenhave, J. P., & Bainter, K. (1995) *Biochem. Molec. Med.*, **54**, 26-32

Van Eikeren, P., & Chipman, D. M. (1972) *J. Am. Chem. Soc.*, **94**, 4788-4790

Van Eikeren, P., & McLaughlin, H. (1977) *Anal. Biochem.*, 77, 513-522

Van Heijenoort, J. (1994) in *Bacterial Cell Wall. New Comprehensive Biochemistry, vol 27* (Ghuysen, J.-M. & Hakenbeck, R., eds.) pp. 39-54, Elsevier Science B. V., Amsterdam

Van Landshoot, A., Loontiens, F. G., Clegg, R. M., Sharon, N., & De Bruyne, C. K. (1977) *Eur. J. Biochem.*, **79**, 275-283

Vårum, K. M., Anthonsen, M. W., Grasdalen, H., & Smidsrød, O. (1991) *Carbohydr. Res.*, **217**, 19-27

Vasala, A., Välkkilä, M., Caldentey, J., & Alatossava, T. (1995) *Appl. Environ. Microbiol.*, **61**, 4004-4011

Vega-Pérez, J. M., Espartero, J. L., Ruiz, F., & Alcudia, F. (1992) *Carbohydr. Res.*, **232**, 235-247

Verma, M. (1986) *Current Microbiology*, **13**, 299-301

Verma, M., & Siddiqui, J. Z. (1986) *Biochem. Int.*, **12**, 267-277

Vermesch, P. S., Lemon, D. D., Tesmer, J. G., & Quicho, F. A. (1991) *Biochemistry*, **30**, 6861-6866

Vernon, C. A. (1967) *Proc. R. Soc. London Ser B*, **167**, 389-401

Verschueren, K. H.G., Seljée, F., Rozeboom, H. J., Kalk, K. H., & Dijkstra, B. W. (1993) *Nature*, **363**, 693-698

Viguera, A. R., & Serrano, L. (1995) *Biochemistry*, **34**, 8771-8779

Viswanadhan, V. N., Ghose, A. K., Revankar, G. R., & Robins, R. K. (1989) *J. Chem. Inf. Comput. Sci.*, **29**, 163-172

Vogel, H. J., & Bonner, D. M. (1956) *J. Biol. Chem.*, **218**, 97-106

Vogt, W. (1995) *Free Rad. Biol. Medic.*, **18**, 93-105

Vyas, N. K. (1991) *Curr. Opin. Struct. Biol.*, **1**, 732-740

Wagner, G., & Wüthrich, K. (1986) *Methods Enzymol.*, **131**, 307-326

Wallace, C. J. A., & Clark-Lewis, I. (1992) *J. Biol. Chem.*, **267**, 3852-3861

Ward, J. B. (1973) *Biochem. J.*, **133**, 395-398

Warshel, A. (1981) *Acc. Chem. Res.*, **14**, 284-290

Warshel, A., & Levitt, M. (1976) *J. Mol. Biol.*, **103**, 227-249

Watt, S. R., & Clarke, A. J. (1994) *FEMS Microbiol. Lett.*, **124**, 113-119

Waxmann, L., & Goldberg, A. L. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4833-4887

Weaver, L. H., & Matthews, B. W. (1987) *J. Mol. Biol.*, **193**, 189-199

Weaver, L. H., Grütter, M. G., Remington, S. J., Gray, T. M., Isaacs, N. W., & Matthews, B. W. (1985a) *J. Mol. Evol.*, **21**, 97-111

Weaver, L. H., Rennell, D., Poteete, A. R., & Matthews, B. W. (1985b) *J. Mol. Biol.*, **184**, 739-741

Weaver, L. H., Grütter, M. G., & Matthews, B. W. (1995) *J. Mol. Biol.*, **245**, 54-68

Wheat, R. W. (1966) *Methods Enzymol.*, **8**, 60-78

Whitaker, J. R. (1963) *Anal. Chem.*, **35**, 1950-1953

White, F. H., Balkwill, D. L., Meeter, D. A., & Merchant, K. K. (1993) *Anal. Biochem.*, **212**, 263-268

Williams, R. M., & Yuan, C. G. (1994) *J. Org. Chem.*, **59**, 6190-6193

Wiseman, A, (1978) in *Topics in Enzyme and Fermentation Biotechnology* (Wiseman, A., ed.), Vol. 2, pp. 280-303, Ellis Horword, Chichester

Wolfrom, M. L., & Thompson, A. (1963) in *Methods in Carbohydrate Chemistry Vol. II* (Whistler, R. L., & Wolfrom, M. L., eds.) pp. 211-215, Academic Press, New York

Wolinski, K., Hinton, J. F., & Pulay, P. (1990) *J. Amer. Chem. Soc.*, **112**, 8251-8260

Work, E. (1957) *Nature*, 179, 841-847

Yamada, H., & Imoto, T. (1981) *Carbohydr. Res.*, **92**, 160-162

Yamada, H., Imoto, T., Fujita, K., Okazaki, K., & Motomura, M. (1981) *Biochemistry*, **20**, 4836-4842

Yamada, H., Imoto, T., & Noshita, S. (1982) *Biochemistry*, **21**, 2187-2192

Yamada, H., Kanaya, E., Inaka, K., Ueno, Y., Ikehara, M., & Kikuchi, M. (1994) *Biol. Pharm. Bull.*, 17, 192-196

Yamamoto, K., Fujinaga, M., & Matsushima, Y. (1963) *Bull. Chem. Soc. Japan*, 36, 1275-1277

Yang, Y., & Hamaguchi, K. (1980) *J. Biochem.*, **87**, 1003-1014

Yao, M., Tanaka, I., Hikichi, K., & Nitta, K. (1992) *J. Biochem.*, **111**, 1-3

Yao, Y., Yin, D., Jas, G. S., Kuczera, K., Williams, T. D., Schoneich, C., & Squier, T. C. (1996) *Biochemistry*, **35**, 2767-2787

Yoakum, G. M., Yeung, A. T., Mattes, W. B., & Grossman, L. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 1766-1770

Yong Chang, K., & Carr, C. W. (1971) *Biochim. Biophys. Acta*, **229**, 496-503

Young, R. Y. (1992) *Microbiol. Rev.*, **56**, 430-481

Young, R. Y., & Bläsi, U. (1995) *FEMS Microbiol. Rev.*, **17**, 191-205

Young, P. R., & Jencks, W. P. (1977) *J. Am. Chem. Soc.*, **99**, 8238-8248

Young, R. Y., Way, J., Way, S., Yin, J., & Syvanen, M. (1979) *J. Mol. Biol.*, **132**, 307-322

Young, J. F., Desselberger, U., Palese, P., Ferguson, B., & Shatzman, A. R. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6105-3-6109

Zhang, T., Bertelsen, E., Benvegnu, D., & Alber, T. (1993) *Biochemistry*, **32**, 12311-12318

Zhang, M., & Vogel, H. J. (1994) *J. Mol. Biol.*, **239**, 545-554

Ziermann, R., Bartlett, B., Calender, R., & Christie, G. E. (1994) *J. Bacteriol.*, **176**, 4974-4984

Zygmunt, W. A., & Tavormina, P. A. (1966) *Can. J. Microbiol.*, **12**, 143-148