Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

STUDIES TOWARDS THE CHEMICAL SYNTHESIS OF THE 18 KDA ANTIGENIC PROTEIN FROM *MYCOBACTERIUM LEPRAE*.

A thesis presented in partial fulfilment
of the requirements for the
degree of Ph.D. in
Chemistry at
Massey University

STEPHEN GEORGE LOVE
1991

ABSTRACT

This thesis describes the solid phase synthesis of a series of peptides from the antigenic *Mycobacterium leprae* 18 kDa protein and the total synthesis of the 148 amino acid protein.

The peptides were synthesised using an Applied Biosystems 430A Solid Phase Peptide Synthesiser usually modified by the removal of in-line filters to the reaction vessel to allow the synthesis of the peptides using programs evolved by S.B.H. Kent. The peptides were cleaved from the peptide resin using liquid HF and purified by reverse phase HPLC.

The first series of peptides to be synthesised revolved around a monoclonal antibody binding site. These peptides were SLP-1, (101-115, RILASYQEGVLKLSI), SLP-2, (111-124, LKLSIPVAERAKPRK), SLP-3, (121-134, AKPRKISVDRGNNG) and SLP-4 (109-125 GVLKLSIPVAERAKPRK). The other peptides were synthesised as a series of overlapping 20 mers covering the entire 148 amino acid sequence. These peptides are SLP-5 (1-20, MLMRTDPFRELDRFAEQVLG), SLP-6 (16-35, EQVLGTS ARPAVMPMDAWRE), SLP-7 (31-50, DAWREGEFVVEFDLPGIKA), SLP-8 (46-65, PGIKADSLDIDIERNVVTVR), SLP-9 (61-80, VVTVRAERPGVDPDREMLAA), SLP-10 (76-95, EMLAAERPRGLFNRQLVLGE), SLP-11 (91-110, LVLGENLDTERIL ASYQEGV), SLP-12 (106-125, YQEGVLKLSIPVAERAKPRK), SLP-13 (121-140, AKPRKISVDRGNNGHQTINK) and SLP-14 (136-148, QTINKTAEHEIIDA). Two other peptides SLP-15 (101-125, RILASYQEGVLKLSIPVAERAKPRK) and SLP-16 (91-115, LVLGEIVLDTERILASVQEGVLKLSI) were also synthesised.

The peptides were used in immunological studies that determined the location of the L-5 monoclonal antibody binding site and showed where T-cell stimulation sites are located on the 18 kDa protein in murine systems.

The synthesis of larger 50 amino acid fragments of the 18 kDa protein, peptides SLP-17 (101-148), SLP-18 (50-100) and SLP-19 (1-50) was carried out in order to determine if it was possible to synthesise the total 18 kDa protein. These fragments were synthesised in a similar manner to the previous peptides. From the synthesis of these peptides it was decided that the total synthesis of the 148 amino acid protein was possible.

The synthesis of the 18 kDa protein was carried out using Kent's protocols as a single step process. The synthesis was monitored up to the 100th amino acid by ninhydrin assay with no failed couplings detected. The coupling percentages for all of the amino

acids was achieved by peptide resin sequencing where the average percentage couplings were shown to be 99.49% with an overall yield for the protein on the resin of 49%.

After Lo-Hi HF cleavage purification of the protein was hampered by the formation of aggregated products which proved initially to be inseparable from the protein. By sequencing some partially purified protein it was shown that under certain cleavage conditions the benzyl ether side chain protecting group was present on the threonine amino acids at positions 5 and 21. A further treatment of the protein using a hard acid-soft base mechanism with trimethylsilyl trifluorosulphonate/thioanisole was used in these cases to remove the remaining benzyl protecting groups.

An attempt to overcome the aggregation of the protein involved the addition of SDS to the HF cleavage vessel. After purification the protein showed no signs of the aggregation products.

A 90 amino acid fragment removed during the course of the synthesis of the 18 kDa proteien was cleaved and dissolved in 6M guanidine.HCl. After gel filtration on a Sephadex G-50 column, preparative HPLC was carried out on the isolated protein peak. The protein was then gel filtered again on a Sephadex G-50 column using 6M guanidine.HCl which separated the aggregated product to give a pure 90 amino acid protein after dialysis.

The full synthetic 18 kDa protein was purified in a similar manner to the 90 amino acid fragment with the second gel filtration being carried out using a Sephadex G-100 column or a Pharmacia Superose 12 column. This provided pure synthetic 18 kDa protein in an estimated 1.8 % yield or 47 mg from this synthesis of a final 2.46 g of protein resin based on the starting protein on the resin.

Acknowledgements

The author wishes to acknowledge Dr Dave Harding and Dr Chris Moore for their advice and encouragement throughout the time spent collating this work. The Medical Research Council is thanked for a postgraduate scholarship and the Leprosy Trust Board for the cost of the chemicals used for the total synthesis of the 18 kDa protein.

I am grateful to Professor J.D. Watson without whom this work would never of been initiated and to Dr R.L. Prestidge, Dr R Booth, Mark Doherty, Dr D. Harris and Thomas Backstrom of the Department of Molecular Medicine, Auckland Medical School, who carried out the immunological assays and provided many enlightening discussions on aspects of immunology.

Trish Heaton is thanked for my initiation into solid phase peptide synthesis. I am grateful to Darren Engelbretson, Simon Burton, Jenny Gibson and Dr N.Haggarty and other members of the SSU for their help and companionship inside and outside the lab for the past three years. Dick Poll's help and advice with HPLC and FPLC chromatography has been invaluable. Marcia Baker who did not wince too often at the price of the chemicals. Staff members and other students in the Chemistry and Biochemistry Department, Massey University are thanked for their inspiration and for helpful discussions.

Dr G. Midwinter, Mr Julian Reed performed the many amino acid analyses and contributed helpful discussion to my sequencing problems particularly in the sequencing of the peptide resin samples. Dr J Ploughman was helpful with the frustrating Circular Dichroism instrument at the Dairy Research Institute. Dr Brian Anderson set up the CONTIN program to run on the micro Vax for me.

Last but not least to my parents George and Margaret, my two brothers Jonathan and Vaughan, I am very grateful for all their support and encouragement in the last four years and to Louise Holden for her support and friendship.

Finally this thesis is dedicated to my grandmother, who provided one of the driving forces for the completion of this thesis and to whom I will never be able to complete a promise.

TABLE OF CONTENTS.

Abstract.	i
Acknowledgements.	iii
Table of Contents.	iv
Table of Figures.	viii
Table of Tables.	xiii

CHAPTER 1. A REVIEW OF LEPROSY.

1.1 Leprosy the Disease.	1
1.2 Chemotherapy.	1
1.3 Immunological Responses and Possible Vaccines.	3
1.4 Non Protein Antigens.	5
1.5 Protein Antigens.	6
1.6 Summary.	9

CHAPTER 2. PEPTIDE SYNTHESIS.

2.1 Introduction.	10
2.1.1 Peptide Synthesis.	10
2.1.1.1 Introduction.	10
2.1.1.2 Solid Phase Peptide Synthesis (SPPS).	12
2.1.1.3 FMOC Peptide Synthesis.	17
2.1.1.4 Summary of Peptide Synthesis.	18
2.1.1.5 Analysis and Purification of Synthetic Peptides.	19
2.1.2 Immunology: Antibody and T-Cell Site Recognition.	22
2.1.2.1 Introduction.	22
2.1.2.2 Development of Antibody Responses.	24
2.1.2.3 Antigenicity of Proteins.	24
2.1.2.4 T-Cell Antigens.	25
2.1.2.5 Summary.	27
2.2 Experimental	28
2.2.1 General Experimental Methods.	28
2.2.1.1 Peptide Synthesis.	28

2.2.1.2 Ninhydrin Analysis. 29
2.2.1.3 Cleavage and Peptide Purification. 30
2.2.1.4 Removal of the DNP Group from Histidine. 31
2.2.1.4.1 Removal of the DNP Group Prior to the HF Cleavage 31
2.2.1.4.2 Removal of the DNP Group After the HF Cleavage. 31
2.2.1.5 High Performance Liquid Chromatography. 31
2.2.1.6 Peptide and Protein Sequencing. 32
2.2.1.7 Amino Acid Analysis. 32

2.2.2 Peptides SLP1-to SLP-16. 33

2.2.3 Immunology Assay Experimental. 36
2.2.3.1 Animals. 36
2.2.3.2 Enzyme Linked Immunoabsorbance Assay. 36
2.2.3.3 Data Analysis. 37
2.2.3.4 Purification of the 18 kDa Protein Antigen. 37
2.2.3.5 Immunisation. 37
2.2.3.4 T-Cell Proliferation Assay. 38

2.3 Results and Discussion. 39
2.3.1 Peptide Synthesis. 39
2.3.1.1 Peptides Synthesised Prior to the Commencement of this Work. 39
2.3.1.2 Peptides Synthesised. 40
2.3.1.2.1 Peptide SLP-1 to SLP-4. 40
2.3.1.2.2 SLP-5 (1-20). 40
2.3.1.2.3 SLP-6 (16-35). 41
2.3.1.2.4 SLP-7 (31-50). 41
2.3.1.2.5 SLP-8 (46-65). 43
2.3.1.2.6 SLP-9 (61-80). 46
2.3.1.2.7 SLP-10 (76-95). 46
2.3.1.2.8 SLP-11 (91-110). 46
2.3.1.2.9 SLP-12 (106-125). 48
2.3.1.2.10 SLP-13 (121-140). 48
2.3.1.2.11 SLP-14 (136-148). 50
2.3.1.2.12 SLP-15 (101-125). 51
2.3.1.2.13 SLP-16 (91-115). 51

2.3.2 Immunology Assays. 54
2.3.2.1 RLP Peptides and SLP-1 to SLP-4. 54
2.3.2.2 Peptides SLP-5 to SLP-16. 56

2.3.3 Summary of Chapter 2. 60

CHAPTER 3 PROTEIN SYNTHESIS.

3.1 Introduction. 63

3.2 Experimental. 71
3.2.1 General Experimental. 71
3.2.1.1 Gel Filtration. 71
3.2.1.2 Ion Exchange Chromatography. 71
3.2.1.3 Polyacrylamide Gel Electrophoresis (PAGE). 71
3.2.1.4 Polyacrylamide Gel Electrophoresis using the Phast System 72
3.2.1.5 Mass Spectrometry 72
3.2.1.6 Tryptic Digests. 73
3.2.2 Peptides SLP-17, SLP-18, SLP-19. 73
3.2.2.1 Peptide 101-148.SLP-17. 73
3.2.2.2 Peptide 50-100 SLP-18. 74
3.2.2.3 Peptide 1-50, SLP-19. 75
3.2.3. Protein Synthesis. 76
3.2.3.1 The Synthesis of the 148 Amino Acid 18 kDa Protein. 76
3.2.3.2 Cleavage #1. 78
3.2.3.3 Cleavge #2. 79
3.2.3.4 Cleavage #3. 79
3.2.3.5 Removal of the Residual Protecting Groups using Hard Acid Soft Base. 80
3.2.3.6 Cleavage #4. 82
3.2.3.7 Cleavage #5. 83
3.2.3.8 90 mer protein (Arg^{58}-Ala^{148}). 84
3.2.3.9 Cleavage #6. 86
3.2.3.10 Cleavage #4 Samples Purified Using the Methods Developed in Cleavage #6. 88
3.2.3.11 Western Blots of the Synthetic 18kDa Protein. 88

3.3 Results and Discussion. 90
3.3.1 Synthesis of Large Peptides SLP-17,SLP-18 and SLP-19. 90
3.3.1.1 C-Terminal PeptideSLP-17 101-148. 90
3.3.1.2 Mid Region Peptide SLP-18 50-100. 95
3.3.1.3 N-Terminal Peptide SLP-19 1-50. 95
3.3.1.4 Asp-Pro Bond Acid Stability. 96
3.3.1.5 Summary. 98
3.3.2 Protein Synthesis. 100
3.3.3 Protein Cleavage and Purification. 104
3.3.3.1 Cleavage #1. 104
3.3.3.2 Cleavage #2. 106
3.3.3.3 Cleavage #3. 106
3.3.3.4 Cleavge #4. 108
3.3.3.5 Cleavage #5 with SDS. 111
3.3.3.6 Cleavage of the 90 Amino Acid 18 kDa Peptide (Arg^{48}-Ala^{148}). 114
3.3.3.7 Cleavage #6. 120
3.3.3.8 Cleavage of Stored 18 kDa Protein from Cleavage # 4. 124
3.3.3.9 Protein Assays. 129
3.3.3.10 Conclusion. 132

APPENDIX

A.1.1 Introduction. 137
A.1.2 Experimental. 138
A.1.3 Results and Discussion. 139
A.1.4 Conclusion. 144

REFERENCES. 145

TABLE OF FIGURES.

CHAPTER 1.

Figure 1.1 Disease Spectrum of Leprosy showing the extremes of the disease. 5

CHAPTER 2.

Figure 2.1.1. A scheme for DCC mediated coupling reactions with amino acids. 11
Figure 2.1.2. Merrifield solid phase peptide synthesis for Boc chemistry 13
Figure 2.1.3. PAM handle attached to the resin. 14
Figure 2.1.4. FMOC peptide synthesis showing the coupling of an FMOC protected amino acid with the peptide chain. 18
Figure 2.1.5. Demonstration of preview analysis in sequencing peptide resin samples. 20
Figure 2.3.1. Ninhydrin results for peptide 1-20 41
Figure 2.3.2. Analytical HPLC of crude and purified peptide1-20 41
Figure 2.3.3. Ninhydrin results for peptide 16-35. 42
Figure 2.3.4. Analytical HPLC of crude and purified peptide 16-35. 42
Figure 2.3.5. UV spectra of the purified peptide 16-35 42
Figure 2.3.6. Ninhydrin results for peptide 31-50. 43
Figure 2.3.7. Analytical HPLC of crude and purified peptide 31-50. 43
Figure 2.3.8. UV spectra of the purified peptide 31-50 43
Figure 2.3.9. Ninhydrin results for peptide 46-65. 44
Figure 2.3.10. Analytical HPLC of crude and purified peptide 46-65. 44
Figure 2.3.11. Hopp and Woods Hydrophilicity plot for the 18kDa protein. 45
Figure 2.3.12. Kyte and Doolittle Hydropathic Index plot for the 18 kDa protein 45
Figure 2.3.13. Predicted secondary structure plot using Garnier-Osguthorpe-Robinson prediction program. 45
Figure 2.3.14. Ninhydrin analysis for peptide 61-80. 46
Figure 2.3.15. Analytical HPLC of the crude and purified peptide 61-80 46
Figure 2.3.16. Ninhydrin analysis for peptide 76-95. 47
Figure 2.3.17. Analytical HPLC of the crude and purified peptide 76-95 47
Figure 2.3.18. Ninhydrin analysis for peptide 91-110. 47
Figure 2.3.19. Analytical HPLC of the crude and purified peptide 91-110. 47
Figure 2.3.20. Ninhydrin analysis for peptide 106-125. 48
Figure 2.3.21. Analytical HPLC of the crude and purified peptide 106-125. 48
Figure 2.3.22. Ninhydrin analysis for peptide 121-140. 49

Figure 2.3.23. Analytical HPLC of the crude and purified peptide 121-140. 49
Figure 2.3.24. Analytical HPLC for the peptide 136-148 50
Figure 2.3.25. Ninhydrin analysis for peptide 101-125. 51
Figure 2.3.26. Analytical HPLC of the crude and purified peptide 101-125. 51
Figure 2.3.27. Ninhydrin analysis for peptide 91-115 for synthesis #1 and #2 52
Figure 2.3.28. Analytical HPLC traces for peptide 91-115. 52
Figure 2.3.29. Protein sequence showing the peptides synthesised in the first phase of this work (RLP peptides) and the earlier SLP peptides SLP-1 to SLP-4. 54
Figure 2.3.30. A comparison of L5 Binding to synthetic peptides in ELISA assays. 54
Figure 2.3.31. Proliferative response lymph node cells from mice immunised with 18 kDa protein and challenged in culture with synthetic peptides. 55
Figure 2.3.32. Proliferative response of lymph node cells from mice immunised with synthetic peptide then challenged in culture with 18 kDa protein. 55
Figure 2.3.33. Protein sequence showing the overlapping 20 amino acid peptides and the two 25 amino acid peptides. 56
Figure 2.3.34. IgG from BALB/cJ and B10.BR mice immunised and subsequently challenged with 18 kDa protein 56
Figure 2.3.35. IgG titre to the 18kDa protein From BALB/cJ and B10.BR mice and challenged with the 20 mer synthetic peptides. 57
Figure 2.3.38. IgG titre to the immunising peptide from BALB/cJ and B10.BR mice immunised and with the synthetic peptides 57
Figure 2.3.37. Proliferative T-cell responses to lymph node cells from mice immunised with 20 mer peptides and challenged *in-vitro* with 18 kDa protein. 59

CHAPTER 3

Figure 3.3.1. Protein sequence showing the peptides 1-50, 50-100, 101-148, the 90 amino acid fragment 59-148 and the full length 148 amino acid 18 kDa protein synthesied in this chapter. 90
Figure 3.3.2. Ninhydrin analysis for peptide 101-148 synthesis #2. 91
Figure 3.3.3. Analytical HPLC of the crude peptide 101-148 from synthesis 2 91

Figure 3.3.4. Analytical HPLC of crude and purified peptide 101-148 isolated from the peptide resin 92

Figure 3.3.5. SDS-PAGE gel showing the HPLC purified 50 amino acid peptides. 93

Figure 3.3.6. Ninhydrin analysis for peptide 50-100. 94

Figure 3.3.7. Analytical HPLC of purified 50-101. 95

Figure 3.3.8. Ninhydrin analysis for peptide 1-50. 96

Figure 3.3.9. Analytical HPLC traces of purified peptide 1-50 and peptide 1-50 hydrolysed with 0.015N HCl. 97

Figure 3.3.10. Analytical HPLC of peptide 1-50 hydrolysed for 48, 96 and 120 hours respectively with 10% acetic acid. 97

Figure 3.3.11. Analytical HPLC of purified 50-100 and of peptide 50-100 hydrolysed with 0.015N HCl 98

Figure 3.3.12. Analytical HPLC of peptide 50-100 hydrolysed with 10% actetic acid at pH 2.5 for 48, 96 and 120 hours. 98

Figure 3.3.13. Ninhydrin assay for the total protein synthesis. 101

Figure 3.3.14. Protein resin sequence analysis. 102

Figure 3.3.15. Analytical HPLC of the recombinant 18 kDa protein as supplied by the Department of Molecular Medicine Auckland Medical School. 104

Figure 3.3.16. SDS-PAGE of synthetic 18 kDa protein from cleavage #1. purified by preparative HPLC. 105

Figure 3.3.17. SDS-PAGE gel Phast System of synthetic 18 kDa protein from cleavage 3 after treatment with hard acid soft base and purification by preparative HPLC 107

Figure 3.3.18. Analytical HPLC of crude and gel filtered 18 kDa protein from cleavage 4. 109

Figure 3.3.19. Analytical HPLC of crude and gel filtered 18 kDa protein from cleavage 4. 110

Figure 3.3.20. Tryptic map of gel filtered (fraction 1) synthetic 18 kDa protein from cleavage 4. 110

Figure 3.3.21. Protein sequence showing the identified tryptic fragments from digest Figure 3.3.20. 110

Figure 3.3.22. Analytical HPLC of crude synthetic 18 kDa protein from cleavage 5 112

Figure 3.3.23. Gel filtration on G-100 column of SDS protein complex and analytical HPLC of the resulting eluted protein. 112

Figure 3.3.24. Analytical HPLC of ultrafiltered material from cleavage 5 after gel filtration. 112

Figure 3.3.25. SDS-PAGE gel of synthetic 18 kDa from synthesis 5. 113

Figure 3.3.26. SDS-Page gel of gel filtered and dialysed synthetic 18 kDa protein from cleavage 5. 113

Figure 3.3.27. Gel filtration on G-50 column in 6M guanidine.HCl of crude 90 amino acid protein. 115

Figure 3.3.28. Analytical HPLC of crude 90 amino acid protein and 90 amino acid protein after gel filtration. 115

Figure 3.3.29. Preparative HPLC of 90 amino acid protein after gel filtration 116

Figure 3.3.30. Gel filtration of 90 amino acid protein after preparative HPLC 116

Figure 3.3.31. SDS-PAGE gel of 90 amino acid material isloated after gel filtration (figure 3.3.30) and dialysis. 116

Figure 3.3..32. SDS-PAGE of repeat purification of 90 amino acid protein. 116

Figure 3.3.33. Analytical HPLC of stored purified 90 amino acid protein. 118

Figure 3.3.34. SDS-PAGE gel of protein peak trapped from the analytical HPLC (figure 3.3.33) showing the intact 90 amino acid fragment and the two fragments at 6 and 4 kDa. 119

Figure 3.3.35. SDS-PAGE gel of the 90 amino acid protein after gel filtration from figure 3.3.33 and 34. 119

Figure 3.3.36. Gel filtration of crude 18 kDa on a G-50 column in 6M guanidine.HCl. 121

Figure 3.3.37. Preparative HPLC of 18 kDa protein. 121

Figure 3.3.38. SDS-PAGE of purified synthetic 18 kDa protein from the first fraction isolated from the preparative HPLC and gel filtered 121

Figure 3.3.39. Gel filtration on a G-100 column in 6M guanidine.HCl of one of the 18 kDa fractions from the preparative HPLC (figure 3.3.37). 122

Figure 3.3.40. SDS-PAGE of fractions from the gel filtration (figure 3.3.38) after dialysis. 122

Figure 3.3.41. a) Streak material from figure 3.3.39, and 3.3.40 subject to treatment with hard acid soft base and purified by analytical HPLC b) SDS-PAGE of material after HPLC. 123

Figure 3.3.42. Gel filtration in 6M guanidine.HCl of stored synthetic 18 kDa from cleavage 4, fraction 2. 125

Figure 3.3.43. SDS-PAGE of material from gel filtration fraction 1 (figure 3.3.42) after preparative HPLC. 125

Figure 3.3.44. Preparative HPLC of material from gel filtration fraction 2, figure 3.3.42. 126

Figure 3.3.45. SDS-PAGE of synthetic 18 kDa from the second gel filtration after preparative HPLC (figure 3.3.43). 126

Figure 3.3.46. UV spectra of a) cloned 18 kDa protein. b) synthetic 18 kDa. 127

Figure 3.3.47. Tryptic digests maps of a) synthetic 18 kDa protein. b) recombinant 18 kDa protein. c) material that streaks. 128
Figure 3.3.48. T-cell proliferation response to immunisation with synthetic 18 kDa protein compared to the proliferation response due to immunisation with recombinant protein. 130
Figure 3.3.49. Western blot blot stained to detect protein. 130
Figure 3.3.50. Western blots comparing the synthetic and the recombinant 18 kDa protein to antibody L5, and antipeptide antibodies α1-20 and α76-95. 131
Figure 3.3.51. Western blots comparing the synthetic and the recombinant 18 kDa protein to antipeptide antibodies α23 (RLP-23), α1-20 and α76-95. 131

APPENDIX.

Figure A.1 Profiles of the respective forms of secondary structure of polylysine in the random, α-Helical and β-sheet conformations. 137
Figure A.2. CD spectra of peptide 61-80 140
Figure A.3. CD spectra of peptide 91-110. 141
Figure A.4. a) CD spectra of peptide 1-20 in 80% TFE. b) CD spectra of peptide 91-115 in 80% TFE 142
Figure A.5. CD spectra of peptides 101-148 and 50-100 143
Figure A.6. CD spectra of recombinant 18 kDa protein. 143

TABLE OF TABLES.

CHAPTER 2.

Table 2.1.1. Common Side Chain Protecting Groups used for Boc Amino Acid Chemistry. 16
Table 2.1.2. Alternative α-Amino Groups Cleaved by Milder Chemical Conditions. 17
Table 2.2.1. Experimental detail for peptides SLP-1 to SLP-16. 34

Table 2.3.1. AAA of peptide SLP-1 (101-115). 40
Table 2.3.2. AAA of peptide SLP-2 (110-125). 40
Table 2.3.3. AAA of peptide SLP-3 (115-134). 40
Table 2.3.4. AAA of peptide SLP-5 (1-20). 41
Table 2.3.5. AAA of peptide SLP-6 (16-35). 42
Table 2.3.6. AAA of peptide SLP-7 (31-50). 42
Table 2.3.7. AAA of peptide SLP-8 (46-65). 44
Table 2.3.8. AAA of peptide SLP-9 (61-80). 46
Table 2.3.9 AAA of peptide SLP-10 (76-95). 47
Table 2.3.10. AAA of peptide SLP-11 (91-110). 47
Table 2.3.11. AAA of peptide SLP-12 (106-125). 48
Table 2.3.12. AAA of peptide SLP-13 (121-140). 49
Table 2.3.13. AAA of peptide SLP-14 (136-148). 50
Table 2.3.14. AAA of peptide SLP-15 (101-125). 51
Table 2.3.15. AAA of peptide SLP-16 (91-115). 52

CHAPTER 3.

Table 3.1. Proteins synthesised using solid phase chemistry developed by Merrifield. 64
Table 3.3.1. Amino acid analysis table for peptide 101-148, synthesis #2. 90
Table 3.3.2. Amino acid analysis of peptide 101-148 isolated from the total protein synthesis. 92
Table 3.3.3. Amino acid analysis of peptide 50-100. 94
Table 3.3.4. Amino acid analysis of peptide 1-50. 96
Table 3.3.5. Peptide resin hydrolysis of samples removed during the protein synthesis. 100

Table 3.3.6. Sequence analysis of peptide resin samples from the protein synthesis. 103
Table 3.3.7. Amino acid analysis of 18 kDa protein from cleavage #1. 105
Table 3.3.8. Amino acid analysis of 18 kDa protein from cleavage #4. 109
Table 3.3.9. Amino acid analysis of the 90 amino acid protein fragment. 118
Table 3.3.10. Amino acid analysis of the 18 kDa protein from cleavage 6. 122
Table 3.3.11. Amino acid analysis of the cloned 18 kDa protein, synthetic 18 kDa protein pure and of the streak material from the synthetic 18 kDa protein. 124

Chapter 1.

A Review of Leprosy.

LEPROSY.

1.1 LEPROSY THE DISEASE.

Leprosy remains a major communicable disease, particularly in the developing countries where it is estimated that there are an estimated 10-12 million people infected with leprosy and there is an at risk population of close to 1 billion people [1]. The current technologies for diagnosing and treating leprosy have not had any significant impact on disease control, with the number of registered cases staying relatively constant at around 5.1 million [2]. This is despite the success of chemotherapy programs in populations already treated.

Mycobacterium Leprae was one of the first human bacterial pathogens to be identified (Hanson 1873) [3], but the fact that it has not been cultivated *in-vitro*, has hampered investigations into the disease. It was not until 1960 that the first reproducible transmission of leprosy in the mouse footpad was reported by Shepard [4]. Finally after investigation into some 80 species [5] reproducible progressive *M.Leprae* was found to occur in the Nine Banded Armadillo (*Dasypus novemcintus linn*) (Kichheimmer and Storrs) [6]. There are still draw backs in using the Nine Banded Armadillo. Only 50-80% of the indigenous animals are susceptible to *M.Leprae* infection which does not manifest itself in the same way as in humans. Also armadillos do not breed in captivity. Despite this the mouse footpad assay and the Nine Banded Armidillo represent the two main animal models used for studying leprosy.

It is possible to induce leprosy infection in the primates, Rhesus, Mangabes and African Green monkeys. The infection resembles the human disease, providing the hope that these monkeys may provide the best long term animal model for comparative study of the disease in humans [7].

1.2 CHEMOTHERAPY.

The Multidrug chemotherapy program being used throughout the world appears to be impracticable due to the number of patients, the size of the at risk population and the requirement for drug taking supervision of patients. The Multidrug Therapy (MDT) combines three drugs depending on the stage of the disease. Dapsone was the most common drug to be used in the last forty years and takes three to six months to effect clinical improvement in patients [8]. For

patients with paucibacillary leprosy*. Dapsone is now prescribed with Rifampicin, a faster acting drug which takes about five weeks for clinical improvement. For patients with multibacillary leprosy Dapsone, Rifampicin and Clofazimine are prescribed [1]. Dapsone monotherapy can cause drug resistant strains of *M.Leprae* to occur. Resistance is thought to be due to irregular small doses of Dapsone [9]. There are also reports of resistance occurring for Rifampicin and Clofazimine [10,11].

There are several reasons why chemotherapy appears to be ineffective in eradication of leprosy on a world wide scale. *M.Leprae* is non-toxic and well tolerated by the human body either not developing into the clinical disease or taking five to fifteen years or more for the symptoms of leprosy to develop [1]. The detection of people infected with the disease is a major undertaking particularly in the third world countries such as those in the Leprosy Trust Board of New Zealand's targeted area of the Pacific Basin. At present there is no highly specific test for leprosy that can easily be carried out in the field where the major method of detection of leprosy sufferers relies on physical signs such as lesions, patches on the skin, swollen nerves and gross physical damage. Therefore leprosy can be confused with other diseases such as Kaposi's Sarcoma, late syphillis,psoriasis and tinea. However an assay is currently being developed that may be carried out in the field that is highly specific for *M.Leprae* [12,13]. This assay is based on Phenolic Glycolipid I (PGL-I) and consists of a dip stick dot enzyme immunoassay using the disaccharide terminal sugars of PGL-I as the antigen. The dipstick consists of two nitrocellulose pads, on one of which the antigen is dotted. A positive result gives a blue colour. These antigen dipsticks are stable for up to four months and the assay can be completed in 30 minutes. Field tests are being carried out testing this method and another similar method in India.

The appearance of drug resistant strains of *M.Leprae* if not controlled may begin to affect the MDT program, which is at present the only effective method available to control leprosy. Ultimately there is the stigma attached to leprosy that causes sufferers to hide their condition, particularly in the third world. This in turn encourages spread of the disease through the immediate population. These complications have led to investigations into new drugs and started a search for a universally applicable vaccine.

---

* Multibaccilary Leprosy includes lepromatous (L) and borderline (B) leprosy in the Madrid classification and polar lepromatous (LL), borderline lepromatous (BL) and borderline leprosy (BL) in the Jopling classification with bacterial index ≥ 2 according to Ridley scale at any one site.

Paucibaccilary Leprosy includes indeterminate (I) and turberculoid (T) leprosy in the Madrid classification and as indeterminate (I) polar turberculoid (TT) and borderline turberculoid (BT) leprosy in the Ridley Jopling classification with bacteriological index,2 according to the Ridley scale at any one site.

Investigations are being carried out on drugs such as Rifamicin derivatives and fluorinated quinolones which have comparable bactericidal effect to Rifamicin [14]. Dihydrofolate reductase inhibitors such as trimethoprim show a hundred fold greater activity against *M.Leprae* compared to bovine reductases and show great promise as another possible drug if the activity discrimination is as great between *M.leprae* and human reductases[15]. The major drawback for new drug and vaccine development is the lack of an *in-vitro* assay to measure their effect on *M.Leprae*. The mouse footpad assay remains the most important drug screening method despite the use of other assays such as uptake of radioactive thymidine and hypoxanthaine [16,17].

1.3 IMMUNOLOGICAL RESPONSES and POSSIBLE VACCINES.

Immunoprophylaxis is becoming a more viable alternative for developing a vaccine to control leprosy. It is only since the discovery of the Nine Banded Armadillo in 1971 by Kichheimmer and Storrs along with the discovery of cultivatable mycobacteria and the more recent advances in molecular biology, that significant investigations into this area have been possible. The search for a vaccine to combat the spread of leprosy is being propagated under the guidance of the Scientific Working Group (SWG) on Immunology of Leprosy (IMMLEP) which has defined the priority areas as

1) Primary preventive tools, ie vaccine.
2) Immunodiagnostic tools that would lead to a better understanding of the epidemiology of leprosy as well as early diagnosis.
3) Immunopathological methods to understand the immuneunresponsiveness and immunological reactions in certain forms of leprosy so that more effective interventions are possible.

Under the guidance of the SWG, the World Health Organisation (WHO) Special Programme for Research and Training in Tropical Diseases is testing a candidate vaccine comprising of BCG (*Bacilla Calmette-Guerin*) and killed *M.Leprae* from the Nine Banded Armadillo, in Venezuela, Malawi and India [1,2]. Killed *M.Leprae* given alone does not provide immunopotentiation for lepromatous leprosy patients, hence the necessity to combine it with another immunostimulant such as BCG[18]. BCG, an attenuated cultivatable mycobacteria derived from *Mycobacterium Bovis*, has been used as a vaccine against tuberculosis with some evidence that it could be protective against *M.Leprae* [19]. BCG has been found to share common protein antigens with *M.Leprae* and *M.Tuberculosis* [20]. However a major drawback to a vaccine based on killed *M.Leprae* is the high cost and the availability of a long term supply in large enough quantities for a general world wide vaccine.

Other vaccines being investigated are based on a) *M.Leprae* antigens; b) cultivatable mycobacteria that share common cell mediated immunity function antigens with *M.Leprae* ie *Mycobacterium w* [21] and ICRC baccillus [22]. There are also investigations into using BCG as a multivaccine vehicle where BCG expresses a set of recombinant genes encoding protective antigens of a specific pathogen ie *M.Leprae* [23]. Recent advances have also shown that *Mycobacterium Habana* TMC 5135 (now*Mycobacterium simiae*), another cultivatable mycobacteria could offer the best hope for a vaccine against leprosy. There are reports that mice have been vaccinated against *M.Leprae* infection and that *M.Habana* shares common key antigens with *M.Leprae* [24,25].

The search for a vaccine for leprosy requires an understanding of the human immune responses to *M.Leprae* as there are a variety of clinical manifestations to the disease [26]. Only a small proportion of the people exposed to *M.Leprae* ever contract leprosy. Most people are naturally immune or have developed effective immunity which prevents the growth of the bacterium from a subclinical stage. At one end of the spectrum in tuberculoid leprosy patients there is a high *M.Leprae* specific cell mediated immunity that destroys the bacilli within infected cells but results in immunological damage to nerves (figure 1). At the other end of the spectrum in lepromatous leprosy patients there are enormous amounts of bacilli present (10^{10} bacilli / g of infected tissue) and very poor cell mediated immune responses [27].

The majority of cases exist in the borderline region between the two extremes. If patients are left untreated at this stage the disease migrates rapidly to either end of the spectrum [2]. Antibodies against *M.Leprae* are present across the whole spectrum of the disease suggesting B-lymphocytes and antibodies do not affect *M.Leprae,* indicating that the humoral immunity is not impaired. Cell mediated responses rather than humoral responses are required for immunological protection against *M.Leprae.* Cell mediated responses involve activation of macrophages and T-cell lymphocytes by presentation of antigens on the surface of antigen presenting cells in association with molecules belonging to the Major Histocompatability Complex (MHC). This response is present in tuberculoid leprosy patients but not in lepromatous leprosy patients. In order to favour the induction of protective immunity it is necessary to evaluate antigen specificity and the function of the T-cell populations involved.

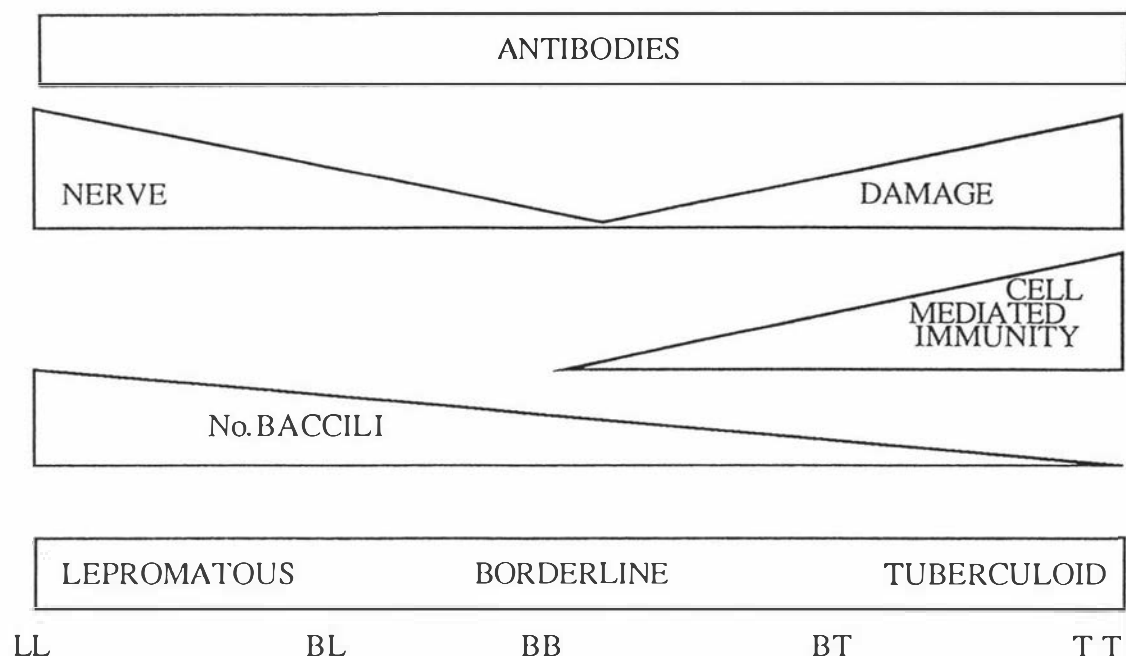


Figure 1. Disease Spectrum of Leprosy showing the extremes of the disease. The antibody titre is constant across the spectrum with nerve damage occuring in the most advanced stages of the disease. Tuberculoid patients are able to clear the bacilli from their bodies with the associated cell mediated immunity where as lepromatous patients have high levels of bacilli[27].

1.4 NON PROTEIN ANTIGENS.

The discovery and structural detemination of the first specific antigen from *M.Leprae*, which elicits significant binding of antibodies from sera of patients, was the phenolic-glycolipid (PGL-I), which is unique to *M.Leprae* [28]. The immunodominant region of PGL-I has been determined as the terminal disaccharide residue, 3,6-Di-D-methyl-ß-D-glucopyranose, which has been prepared by chemical synthesis [29]. An enzyme immunoassay exists based on PGL-I which will detect active Lepromatous Leprosy (LL) and Borderline Lepromatous Leprosy (BL) where there are high levels of bacilli. The assay can detect very low levels of antibody in the urine of LL and BL patients [12]. PGL-I, a highly immunogenic carbohydrate lipoarabinomannan and other carbohydrate antigens are being investigated to determine if they could be used for serodiagnosis, in monitoring leprosy and how they protect mycobacteria against host cell defence mechanisms [30,31]. *M.Leprae* grows in mononuclear phagocytes, cells most heavily armed for destruction of pathogens. It has been discovered that PGL-I suppresses mitogenic responses of lymphocytes with suppressor T-cells recognizing the tri-saccharide moiety of PGL-I [32]. There is some evidence that PGL-I acts as a scavanging system *in-vitro* for

hydroxyl radicals (OH·) and superoxide anions (O_2-), which are thought to be the major cytotoxides in activated macrophages against intracellular parasites. This maybe the factor that permits *M.leprae's* survival in human mononuclear phagocytes [33,34].

T-cell antigen recognition appears to involve mainly protein antigens. Recognition of PGL-I and carbohydrate components of *M.Leprae* has been reported[32]. There is, however, doubt about the ability of carbohydrate or glycolipid antigens alone to form functional interactions with MHC molecules in order to allow T-cell recognition.

1.5 PROTEIN ANTIGENS.

In 1984 a workshop organised by the WHO was held to examine monoclonal antibody reagents to antigenic components of the leprosy bacilli [35]. These monoclonal antibodies raised from the spleen cells of mice in a variety of laboratories around the world were examined for antigenic specificity and cross reactivity on a panel of mycobacterial strains. A range of immunodominant proteins for *M.Leprae* were identified in terms of their molecular weight, 65, 36, 28, 18 and 12 kD. There is now a panel of monoclonal antibodies that permits the identification of the leprosy bacillus highlighting features that are unique to *M.Leprae*.

In 1985 R.A. Young, *et al* used the monoclonal antibodies to isolate genes encoding for the five most immunogenic protein antigens of the leprosy bacillus [36]. From isolating the total *M.Leprae* genomic DNA, randomly shearing the DNA into small fragments then ligating these fragments to Λgt11, a bacteriophage vector is created that can drive the expression of foreign DNA with *E.coli* transcription and translation signals. By infecting *E.coli* cells with the bacteriophage, Young and co-workers built a DNA library that covered the entire genomic DNA of *M.Leprae*. Using the monoclonal antibodies it was possible to isolate the DNA fragments that encoded the antigenic proteins. These fragments have been sequenced to provide either partial or total amino acid sequences of these proteins.

Since the discovery of the five key immunodominant proteins others such as the 31/30 kD complex [37] and a 70 kD protein [38] have also been identified. The discovery of these antigenic proteins makes possible the identification of the epitopes responsible for antibody recognition and those epitopes that act as T-cell stimulators or suppressors. A major feature of these proteins is their similarity with stress or heat shock proteins which could offer some insight into the immunological activity of these proteins [39]. It is also hoped that these antigens could be used to immunise against other mycobacterial pathogens.

The 70 kD protein shows homology with a 71 kD *M.Tb* protein and 55% homology with the C-terminal end of the dnaK gene of *E.coli* and therefore the 70 kD protein is a member of a family of heat shock proteins [38]. There has been one B-cell epitope identified on the *M.Leprae* 70 kD protein with murine monoclonal antibodies L7 and L27, but no other B-or T-cell epitopes have been identified [40]. The 70 kD protein stimulates *in-vitro* T-cell proliferation with some variation between patients in the proliferation response [41]. Synthetic peptides from a region of the protein that has multiple alpha helices (a typical protein secondary structure that is found to be involved in T-cell responses) were not reactive.

The 65 kD protein is a frequently targeted protein of monoclonal antibodies raised during their immunisation of mice. There is also a high frequency of human T-cell responses to this protein [35]. The 65 kD protein also shows homology with a highly conserved family of bacterial proteins notably a heat shock stress protein *E.coli* groEL [39] and 95% homology with *M.Tb* and *M.Bovis* 65 kD proteins [20]. The 65 kD protein epitopes recognized by Mabs have been mapped using subclones and synthetic peptides including the species specific antibody lllE9 whose specificity is determined by lysine^{426} and threonine^{430} in the *M.Leprae* sequence [42]. There are however differences between the mouse monoclonal antibodies and the human leprosy patient reactivity to the 65 kD antigens [43]. Analysis of epitopes recognized by human T-cells has led to the isolation of a *M.Leprae* specific Human T-cell clone which overlaps the lllE9 antibody site [44]. The 65 kD antigen has been used in an attempt to induce immunisation but results seem to indicate there is no enhanced protection against mycobacterial infection. In an effort to induce T-cell growth and increase antibody response levels in mice, a synthetic peptide covering a T-cell stimulation region of the 65 kD protein has been used as a carrier of non responsive regions of the Foot and Mouth disease virus and to regions on the 65 kD protein. The induced antibody and T-cell responses has confirmed that the 20 amino acid 65-85 epitope of the 65 kD protein plays a role in T-cell activity [45].

The 36 kD antigen was originally identified in human patient sera [46]. This antigen is a protein of 249 amino acids with a high frequency of proline rich areas. The 36 kD protein has been shown to react with human patient T-cell clones, one of which is specific to *M.Leprae*, and others that cross react with other mycobacterial diseases [47]. The mapping of B and T-cell epitopes is currently being investigated. There is an assay available that uses primers from the 36 kD gene sequence to allow specific detection of one leprosy bacillus [48].

The 31/30 kD antigen is a secreted antigen originally found in *M.Tb* and *M.Bovis* culture fluids. Cross reactive monoclonal antibodies to human patient sera indicate *M.Leprae* probably expresses a homologous protein [37]. The antigen may not be recognised during screening of *M.Leprae* extracts. Cloning of the gene from *M.Bovis* has provided evidence that the protein is a secreted protein which means it could play a part in the early immunological response to *M.Leprae*. The protein also binds specifically to fibronectin which could be involved in mediating the *M.Leprae* and host cell interaction during infection [49]. There are strong B-and T-cell responses, but so far no detailed mapping has been carried out.

The 28 kD antigen has been cloned, but as yet no B-or T-cell epitopes have been identified [50]. It is suspected that due to the cross reactivity of some of the antibodies to this protein with the *M.Tb* 23 kD antigen, the 28 kD antigen is also secreted in infected tissue.

The 18 kDa protein was originally identified by the use of monoclonal antibodies [35]. It was found that this protein encoded in crude lamba gt11 phage lysates of *E.coli*, stimulated nearly half of the *M.Leprae* -specific T-cell clones [51]. The gene for this protein has been isolated cloned and sequenced to give a 148 amino acid protein [52]. There are very few cross reactive monoclonal antibodies to 18kDa with other mycobacterial pathogens, but there is some relation with a family of small plant heat shock proteins [53]. This observation offered some hope for the development of a relatively easy assay that could be used for the specific detection of *M.Leprae*. However a cross reactive 18 kD protein similar to *M.Leprae* 18kDa protein has also been found in *Mycobacterium Habana* TMC5135 (now *M.simiae*). *M.simiae* is a cultivatable mycobacterium which has been used to vaccinate mice against infection by *M.Leprae* [24]. At this stage it is not known if the presence of the 18 kD protein antigen in *M.simiae* contributes to its ability to immunise mice. The gene in *M.simiae* for the 18 kD antigen is now being sequenced to determine the degree of homology [24].

Synthetic peptides and cloned fragments have been used to map B-cell and T-cell epitopes of the 18 kD protein, leading to the isolation of the antibody binding site for the L5 monoclonal antibody [54] and T-cell response sites that overlap the binding site[55]. There has been a recent report that human T-cells recognise the region involving residues 38-50 of the 18 kDa protein[56]. Further discussion on the results of investigations into this protein will be discussed in Chapter 2.

Further work on isolating other protein antigens has resulted in the identification of some highly immunogenic cell wall associated proteins with apparent molecular weights

of 7-16 and 28 kD. These proteins were tested with human T-cell wall clones to identify immunodominant proteins in *M.Leprae* cell walls [57,58].

1.6 SUMMARY.

Despite isolating the antigens expressed by *M.Leprae* there are still problems with elucidating their role in protective immunity. It will therefore be some time before a vaccine is developed for *M.Leprae* . One of the most unsatisfactory aspects of studying leprosy (as dicussed above) is the lack of available models for investigating *M.Leprae.* As demonstrated by the isolation of the epitopes to the 65 kD protein and the difference in effects between the murine and human sera, there is some difficulty in extrapolating the results from animal to human models.

Development of highly specific probes to *M.Leprae* should arise out of the cloning of protein antigens. These tests involve detecting the genomic DNA of *M.Leprae* that allows detection to 40 or less bacteria in biopsy material. The continuing development of a wide variety of serodiagnosis tests for leprosy, until the development of an effective vaccine, offers the best hope of preventing the spread of the leprosy by enabling the early detection of infected people and rapid treatment by current methods [48,59].

Chapter 2.

Peptide Synthesis.

2.1 INTRODUCTION.

2.1.1 PEPTIDE SYNTHESIS.

2.1.1.1 INTRODUCTION.

In 1906 Emil Fisher defined the peptide bond as an "amide type" linkage between amino acids.[60] Since this discovery synthetic peptides have played an increasingly important role in research into the structure, function and metabolism of biological compounds. The synthesis of peptides and their derivatives has been carried out since 1881[61] for many purposes including the confirmation of general structures, the delineation of antigenic structures and the provision of usable quantities of naturally occurring peptides that are often in trace amounts from natural sources. The major force for much of the current research in synthetic polypeptides is the possibility of using synthetic peptides to construct vaccines. Currently some of the diseases under investigation include Foot and Mouth Disease[62] Hepatitis B [63], Cholera[64], Diphtheria[65], Streptococcus infection[66], Influenza[67], Malaria[68], Polio[69], and AIDS (Human Immunodefficency Virus).[70] The synthesis of polypeptides can be extended to the study of proteins which have been defined as polypeptides of 20 or more amino acids and are classified on the basis of their *in vivo* biological function.[71]

There is now a vast array of techniques for the chemical synthesis of peptides[72], where the principle reaction involved in the formation of a peptide bond between two amino acids, is the acylation of an α-amino group by the α-carboxyl group of an adjacent amino acid. The need to preserve the optical integrity of the chiral α–carbon and the functional side chain groups on many of the amino acids complicates this coupling step. The prevention of racemization is therefore a major concern since biologically active mammalian proteins and peptides are composed of L-amino acids. Racemisation can occur via oxazolone formation or by direct hydrogen abstraction from the α-carbon, and may be minimised in modern peptide synthesis using urethanes to protect the amino group.

During peptide synthesis the amine function of the incoming amino acid is temporarily blocked. The carboxyl function of the carboxy terminus of the peptide is usually blocked throughout the synthesis. The reactive side chains of some of the amino acids are also protected throughout the synthesis by semi-permanent protective groups. However there are cases where the reactive side chains have been left unprotected particularly in solution synthesis where a large number of protecting groups decreases

the solubility of peptides[73] and increases the cost substantially especially for producing gram amounts of peptides. For amino acids to couple, the carboxyl function of one of the amino acids is activated. Figure 2.1.1 shows a popular method for activation where dicyclohexylcarbodiimide (DCC) is used to form a symetrical anhydride species (1) or to form an active ester (2) (eg 1-hydroxybenzotriazole, HOBt). In the coupling process the activated species are reacted rapidly with the amino terminus of the peptide with few side reactions, and often high yields. Once the entire peptide is formed, the semi permanent and temporary protecting groups are removed.

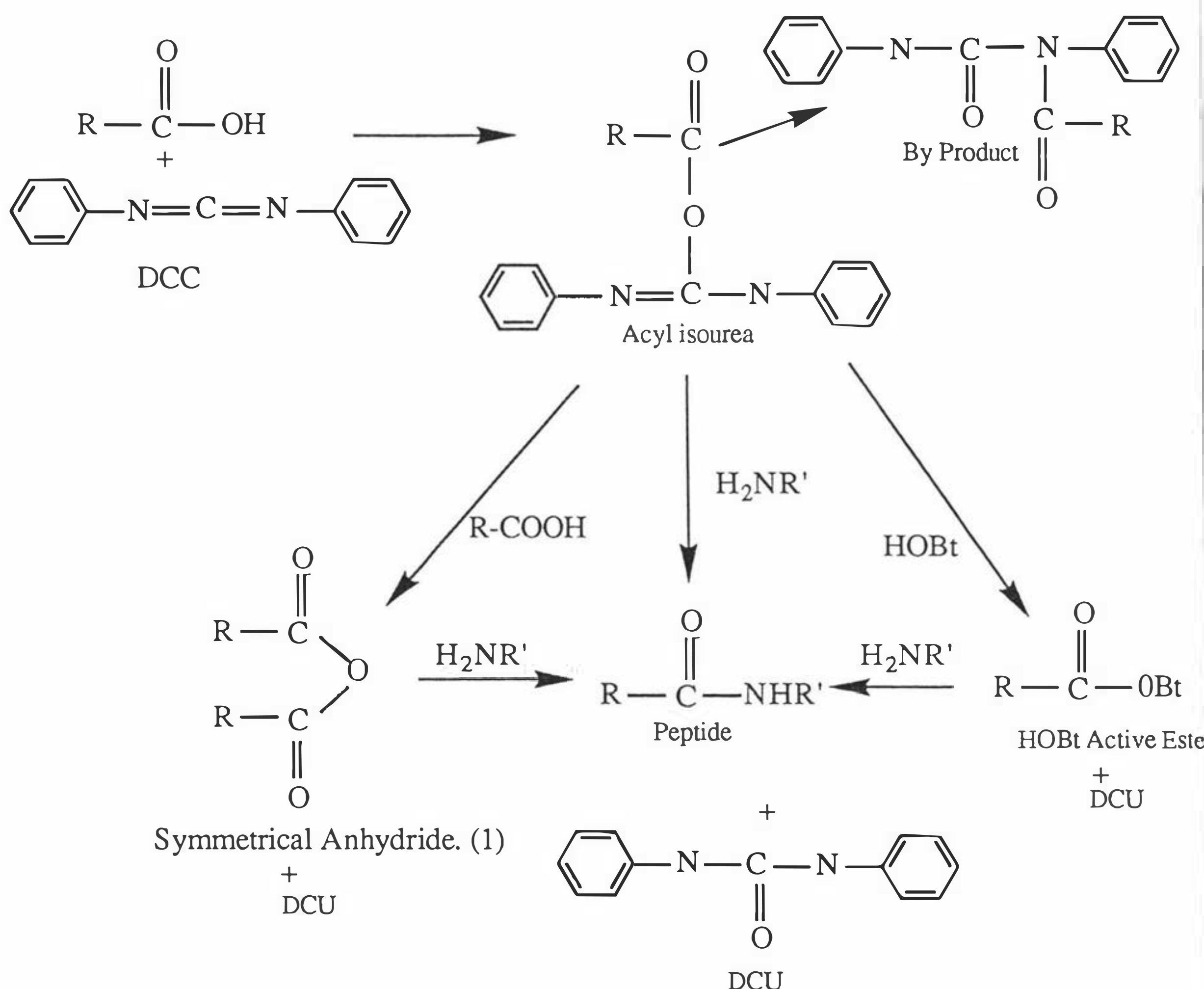


Figure 2.1.1: A scheme showing DCC mediated coupling of amino acids showing the use of symmetrical anhydrides and HOBt active esters in the formation of a peptide bond.

Using solution syntheses and fragment condensation it is possible to synthesise peptides up to 15-20 amino acids reliably. Longer peptide chains may be formed by segment condensation, where smaller peptides are coupled together to form larger peptides. At each step purification and characterisation of the peptide is carried out to ensure the integrity of the growing peptide chain. However the synthesis of peptides in solution is demanding in labour, time and skill. This is mainly due to the difficulty in solubilising the protected peptide fragments and slow coupling speeds. Despite this, in Japan a large number of peptides are synthesised in this manner, including Human Growth Hormone Releasing Factor (1-29)[74] and small hormone peptides such as Endotheli ns[75].

2.1.1.2 SOLID PHASE PEPTIDE SYNTHESIS (SPPS).

In 1963 Merrifield reported the chemical synthesis of a tetrapeptide leucine-alanine-glycine-valine which was covalently bound through the carboxy terminus to a solid insoluble support of chloromethylated polystyrene.[76] Since then Solid Phase Peptide Synthesis (SPPS) has progressed rapidly and has become the major method of peptide synthesis. The idea of attaching the peptide to a solid support has lent itself to a number of other biochemical techniques such as affinity chromatography[77], enzyme immobilisation[78], oligonucleotide and oligosaccharide synthesis[79]. The scope of the original discovery led to the presentation of the Nobel Prize to R.B. Merrifield in 1984.

The basic Merrifield procedure for solid phase synthesis outlined in Figure 2.1.2, is to anchor the growing peptide chain to an insoluble matrix or resin. This allows excess reagents to be easily applied and readily removed, repeatedly if necessary, to provide high yields without mechanical losses. The insolubility problems associated with solution synthesis are prevented and the whole process is capable of being automated. Today standardised processes allow for high speed, repetitive production of peptides. The major drawback of SPPS is the microheterogeneity of the final product resulting from repeated incomplete coupling reactions. This is attributed to the lack of purification (in the solution method sense) of the growing peptide at the end of each coupling step. The microheterogeneity is now minimised by current synthesis protocols with improvements in the yields of the coupling step approaching 100%.

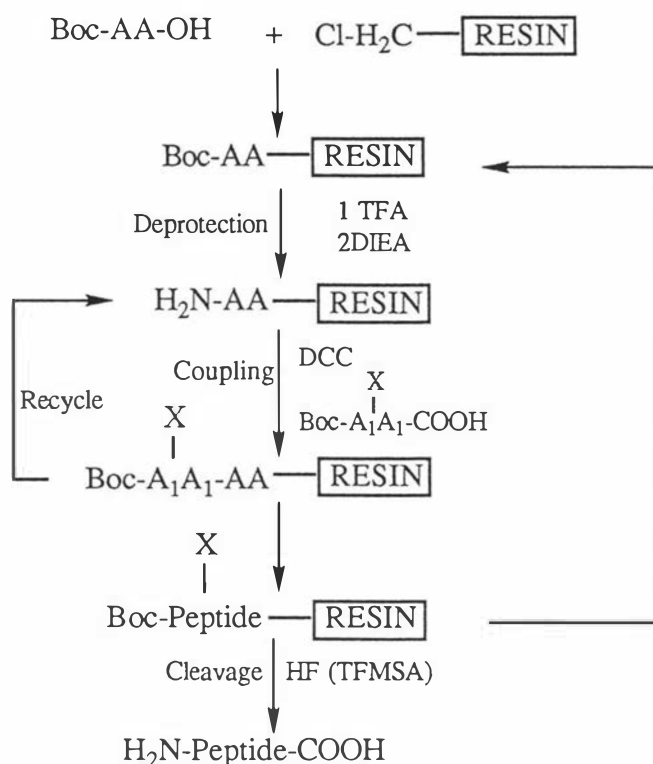


Figure 2.1.2: Merrifield solid phase peptide synthesis for Boc chemistry showing the coupling of an amino acid to a growing peptide chain on the resin. X=Side chain protecting group. AA= amino acid.

In Merrifield solid phase peptide synthesis, the first N-α-*tert* -butyloxycarbonyl (Boc) amino acid is covalently attached through the COOH group to a linker on the resin support. The Boc group is removed by trifluoroacetic acid (TFA) and the amino terminus neutralised by triethylamine (TEA) or more commonly by diisopropylethylamine (DIEA). The incoming amino acid is coupled to the free amino terminus with DCC, with the coupling step often repeated in order to maximise the yield. Generally coupling is achieved by the highly reactive symmetrical anhydride species generated from reacting DCC and the Boc amino acid in a 1:2 ratio (see fig 2.1.1). Any unreacted amino terminus is generally prevented from reacting further by treatment with acetic anhydride (Ac_2O)to acylate the free amino group. Dichloromethane (DCM) and dimethylformamide (DMF) are the preferred solvents for the deprotection, coupling and washing steps of the synthesis. Cleavage of the protected peptide from the resin is carried out using strong acids such as hydrogen fluoride (HF) or trifluoromethanesulphonic acid (TFMSA)[80].

The main type of resin used is a polystyrene polymer cross linked with 1% m-divinylbenzene to give beads which have a diameter of 20-80μm when dry. In solvents such as DCM and DMF the volume of the resin can swell five or six fold, acting like a well solvated gel, enabling reagents to react readily with the peptide terminus[81]. However the properties of the beads change significantly as the peptide grows in length with the peptide and the resin acting to enhance each others solvation properties[82]. Poor coupling yields have been related to the poor swelling of the peptide resin beads,[83] which has been shown to be sequence dependent and attributed to formation of H bonded peptide aggregates. Therefore a major objective is to swell the peptide resin to its maximum volume which has been attempted by using several procedures such as, coupling in aprotic solvent (DMF), adding trifluoroethanol to DCM in the coupling step[84], forming symmetrical anhydrides in N-methylpyrrolidinone at elevated temperature[85], adding dimethylsulfoxide (DMSO) to the coupling step in DCM[86], or adding base at the end of coupling step[84,86].

Several different linkers or 'handles' are available for attaching the first amino acid to the to the resin[72]. These handles are used as bifunctional spacers or extenders. For Boc amino acid chemistry the popular linker is the phenyl-acetamidomethyl (PAM) group (fig 2.1.3)[87]. The PAM group provides greater acid stability (100 times for TFA) compared to the chloromethyl resins developed by Merrifield[88]. The PAM handle reduces peptide loss from the resin caused by cleavage of the peptide from the resin by repeated treatments of TFA. The PAM linker also leads to a higher purity of peptide after cleavage from the resin with HF, particularly for longer peptides[89] .

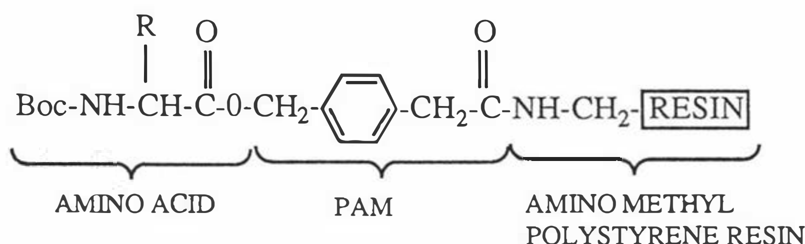


Figure 2.1.3: PAM handle attached to the resin

The most successful coupling reagent used in peptide synthesis is DCC particularly in aprotic solvents. Others such as N-ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline (EEDQ)[90] and diisopropylcarbodiimide[91] have been used in peptide

synthesis but have not found general favour. The activation and coupling processes can be carried out *in situ* with the resin-peptide which requires subsequent washing with methanol and DCM to remove the precipitated urea (DCU) that forms as a result of the coupling process (see Figure 2.1.1). An alternative approach is to preform the symmetrical anhydride (PSA) in DCM, filter away the insoluble DCU, then solvent exchange to DMF to permit coupling in 20-30 minutes[92]. Merrifield has shown that *in-situ* activation, ie in the presence of the resin, proceeds with less deletion and insertion peptides compared with the PSA procedure for small peptides[93]. The major draw back of *in-situ* activation is the most effective removal of the DCU is carried out under conditions that cause the resin to shrink (eg: alcohol containing washes) instead of keeping the resin as swollen as possible.

The amino acids asparagine and glutamine undergo dehydration of their carboxyamide side chains to give nitriles during coupling as symmetrical anhydrides[94], but this can be avoided by using active esters of the amino acids. The preferred active ester is 1-hydroxybenzotriazole (HOBt)[95] which in addition suppresses racemisation and also catalyses the slow coupling reaction of arginine. Other active esters also include *p* nitrophenyl esters[96] where coupling reaction is slow taking up to 18h and benzotriazolyl-N-oxytrismethylamino-phosphoniumhexafluorophosphate (BOP)[97] which is being used successfully for FMOC peptide synthesis[91].(see section 2.1.3)

Amino acids with reactive side chains need to have their reactive sites blocked with protecting groups that are stable to repeated treatments of TFA. The protecting groups commonly used for Boc amino acids today are presented in Table 2.1.1. Of concern to peptide chemists are the effects on the peptide of the repeated treatments with strong acids such as TFA, and the final deprotection by HF or TFMSA[98]. Other α-amino protecting groups requiring less harsh chemical treatments have been investigated such as the BPOC group which requires dilute TFA for removal (Table 2.1.2). Yet other groups such as 3-nitro-2-pyridinesulfenyl (NPyS) and diathiasuccinoyl (Dts) where deprotection is achieved by thiolysis have also been examined. These protecting groups in turn are also associated with handles that require milder conditions for cleavage of the peptide from the resin. The 3-nitro-4-bromomethylbenzoylaminomethyl resin handle is cleaved by 350nm light[106] for example.

Boc AMINO ACID	PROTECTING GROUP
Arginine	*p*-Toluenesulphonyl. (Tosyl)
Aspartic acid	O-Benzyl
Glutamic acid	O-Benzyl
Histidine	2,4-Dinitrophenyl (DNP) or Tosyl
Lysine	Chlorobenzyloxycarbonyl (ClZ)
Methionine	Sulphoxide (O)
Serine	Benzyl
Threonine	Benzyl
Tyrosine	2-Bromobenzyloxycarbonyl (2-Br Z)
Tryptophan	Formyl (CHO)
Cysteine	3-Methylbenzyl, 3-Methoxybenzyl
Alanine, Asparagine Glycine, Glutamine Isoleucine, Leucine Phenylalanine, Proline Valine	Unprotected side chains

Table 2.1.1: Common side chain protecting groups used for Boc amino acid chemistry.

Amino Acid Protecting Group.	Deprotection Method	Ref #
2-*p*-Biphenylpropyl-(2)-oxycarbonyl (BPOC)	Dilute TFA	98
αδ-Dimethyl-3,5-dimethyloxy carbonyl (Ddz)	Dilute TFA	99
2-Phenylpropyl-2-oxycarbonyl (POC)	Dilute TFA	100
α,2,4,5-Tetramethylbenzyloxycarbonyl (TMZ)	Dilute TFA	101
4-Methoxybenzyloxycarbonyl (Moz)	Dilute TFA	102
3-Nitro-2-pyridinesulfenyl (NPyS)	Thiolysis	103
Diathiasuccinoyl (Dts)	Thiolysis	104
Triphenylmethyl (Trt)	Dilute Acetic Acid	105

Table 2.1.2: Alternative α-amino groups cleaved by milder chemical conditions than those used normally for Boc amino acids.

2.1.1.3 FMOC PEPTIDE SYNTHESIS

The most successful peptide synthesic method employing milder chemistry utilizes the N^{α}-9-fluoromethyloxycarbonyl (FMOC) group.[107] This group was adapted to peptide synthesis by Atherton *et.al* in 1978[108]. Refinements to the original protocols have resulted in the general chemical scheme outlined in Figure 2.1.4.

The first FMOC amino acid is linked to an insoluble support via a base stable linker such as the 4-hydroxymethylphenoxyacetic acid (HMAP)[109]. The linked FMOC amino acid is deprotected by base, usually piperidine. The second FMOC amino acid is coupled as the symmetrical anhydride or as the HOBt active ester. Once the peptide has been synthesised, the removal of the base stable side chain protecting groups and peptide cleavage is carried out usually by treatment with TFA. Different solid supports (from 1% cross linked polystyrene) have been developed that take advantage of the milder chemistry[109]. The rationale in this case is to bring the solid support and the peptide chain to comparable polarities. The supports range from copolymerised polyacrylamides[110] with optimum solvation properties in aprotic solvents, to low

density highly permeable matrices like kieselguhr/polyamide[111] or polyHIPE[112]. The kieselguhr and and polyHIPE matrices are able to withstand pressures that enable FMOC peptide synthesis to be carried out in a continuous flow mode through a column packed with the matrix.

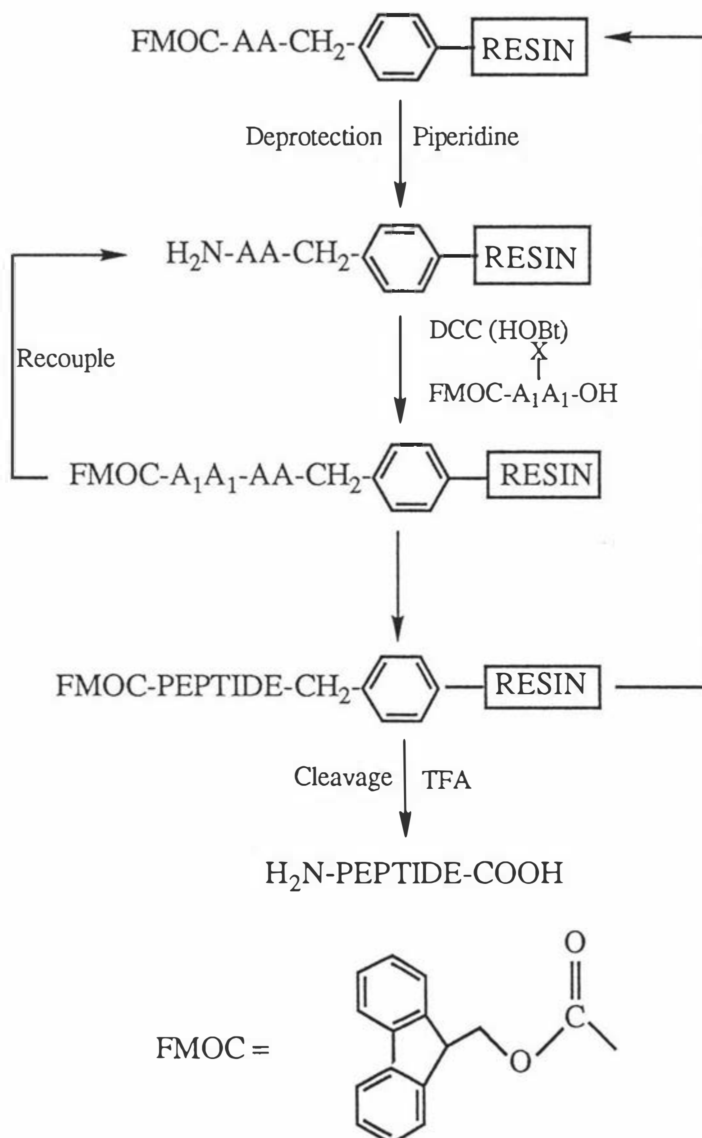


Figure 2.1.4: FMOC peptide synthesis showing the coupling of an FMOC protected amino acid with the peptide chain. X- side chain protecting group. AA= amino acid.

2.1.1.4 SUMMARY of PEPTIDE SYNTHESIS.

Both methods for Boc and FMOC chemistry have been automated for some years. The Applied Biosystems 430A Solid Phase Peptide Synthesiser[113],which can

operate both Boc and FMOC chemistries, was one of the first modern instruments developed. This type of instrument has allowed the rapid development of new synthetic protocols using new solvents such as the NMP-HOBt protocol, where N-methylpyrrolidinone (NMP) and dimethylsulphoxide (DMSO) are used during coupling, replacing DMF. NMP and DMSO swell the resin to the maximum volume providing greater exposure of the peptide chains to the surrounding coupling solution[114].

Rapid synthesis protocols for small amounts of resin have also been developed where the resin is washed only with DMF[92] enabling the resin to remain in the maximum swollen state throughout all cycles. Neat reagents are also used to enhance rapid deprotection and shorten cycle times. The multipeptide T-bag system developed by Houghten *et.al* [115] is a rapid synthesis method for simultaneous peptide synthesis of 40 or more peptides at a time. The peptides synthesised in this manner are used principally for rapid screening in biological assays. Another method used in this manner is peptide synthesis on polypropylene pins allowing direct probing of antibodies with the peptides still attached to the pins[116].

2.1.1.5 ANALYSIS AND PURIFICATION OF SYNTHETIC PEPTIDES.

Errors in the coupling of amino acids can occur at every step in the coupling process. In the past SPPS was criticised as cumulative errors appeared to greatly restrict the overall integrity of the synthesised peptide. An average coupling efficiency of 95% per amino acid corresponds to 36% of a correctly formed 20mer (20 amino acid peptide), 21.5% of a correct 30mer and 8% of a correct 50mer. However today coupling yields routinely average greater than 99.5% where the corresponding yields for peptides are 90.5, 86, and 77.8% respectively. Obvious sources of cumulative errors arise from impure amino acid derivatives such as N-α-*sec*-butoxycarbonyl amino acids[117] and impure solvents from which contaminants such as dimethylamine in DMF could give aspartic acid and glutamic acid side chain amides or can react with an activated carboxyl during coupling.

An incomplete deprotection or coupling step can result in deletion peptides lacking the intervening residue, if chain growth resumes. Irreversible termination of the peptide chain can occur either intentionally or unintentionally. In the case of capping, the result is an intended terminated peptide. Kinetically slow, sequence dependent , difficult couplings that are reproducible can occur, preventing the correct synthesis of a peptide, as in the case of a peptide from Acyl Carrier Protein (ACP) ACP 64-75[118]. An example of this type of peptide is described in this chapter with peptide 91-115. The

deletion and termination peptides can often be discriminated from the target peptide by chromatography on reverse phase or ion exchange chromatography.

Several quantitative and qualitative tests are used to monitor the success of the coupling step. The most widely used is the ninhydrin test[119] where a positive colorimetric result indicates incomplete coupling due to the presence of free Nα amino groups. An accurate quantitative test using the ninhydrin assay has been developed that enables the detection of 1nmole of free Nα-amino groups per gram of resin[120]. Non destructive tests such as the picric acid titration [121] have been used but find little favour in comparison. The drawback in all these methods is the background interference due to the peptide chain or by non specific binding as the peptide chain grows. In FMOC synthesis it is possible to monitor the decrease in absorbance at 300nm when FMOC amino acids are being coupled to the resin. However there is an inherent sensitivity loss by this method of determination as the concentration of chromophoric FMOC amino acid in solution decreases rapidly in the solution as coupling takes place.

Peptide resin hydrolysis is a powerful tool for monitoring a solid phase peptide synthesis, as it allows the accurate determination of the amino acids up to the selected step[122]. Peptide resin sequencing[123] is another tool for monitoring the completeness of a coupling step through preview analysis (see Figure 2.1.5), where an incomplete coupling step,is detected by a peak for the amino acid, at cycle 4 (termed the preview peak) instead of in cycle 5. The ratios of the amino acid peak between the preview peak

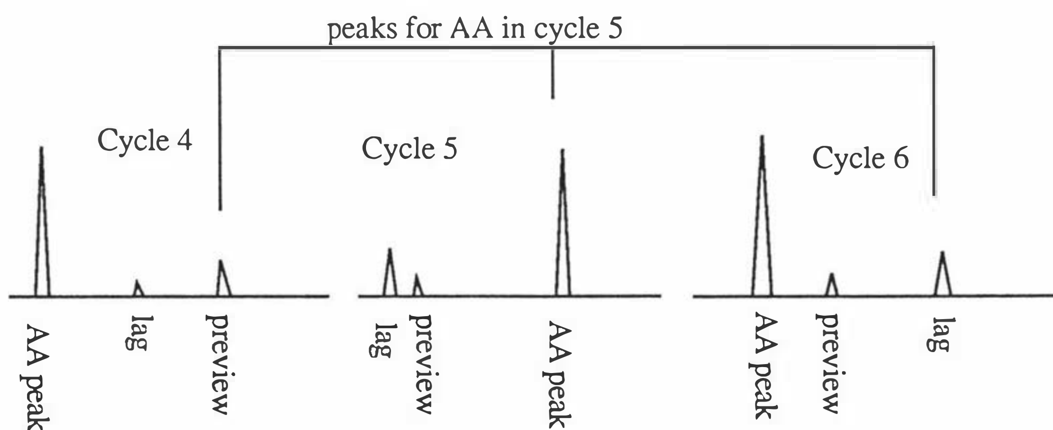


Figure 2.1.5: Demonstration of preview analysis in sequencing peptide resin samples. The height of the preview peak in cycle 4 divided by the height of the coressponding amino acid peak in cycle 5 indicates the % of failed or incomplete peptide chains.

in cycle 4 and the amino acid peak in cycle 5 indicates the % of terminated or deletion peptides up to that step. This method has been refined with the advent of High Performance Liquid Chromatography (HPLC) systems and automatic protein sequencers[124].

The cleavage of the peptide from the resin can be a step where problems occur affecting the purity of the peptide. For Boc chemistry the removal of protecting groups and cleavage from the resin is usually done in a single step process using liquid HF or TFMSA. The cleavage is performed in the presence of nucleophilic scavengers added to protect the peptide from carbocations produced by cleavage of the side chain protecting groups. In some cases the protecting groups are removed prior to cleavage of the peptide from the resin, using HF or TFMSA in a process known as the Lo-Hi or the S_n2 ,S_n1 procedure.[125,126]. This is advantageous in cases where there are many protecting groups or where there are protecting groups that are not removed by the standard single step S_n1 (Hi) cleavage method eg the formyl group on tryptophan. One of the disadvantages of cleavages using HF is the requirement for a specialised all teflon apparatus for safe handling, whereas TFMSA can be used with ordinary laboratory glassware. For FMOC peptides the peptide is cleaved using TFA and scavengers in laboratory glassware.

Following cleavage of the peptide from the resin the peptide must be purified to homogeneity on one or more chromatographic systems. As a first step, gel filtration can be invaluable, as the peptide can easily be separated from the nucleophilic scavengers and desalted from the acid salts. Gel filtration also gives some measure of size separation between complete and incomplete termination peptides. High Performance Liquid Chromatography (HPLC), in the reverse phase mode, is a powerful technique due to its high discrimination between a peptide and its impurities[92]. HPLC is used regularly to purify peptides in a single step from the crude peptide to greater than 95% purity. Other techniques used to purify peptides include Ion Exchange, Affinity Chromatography and Counter Current distribution.

Amino acid analysis (AAA) is routinely used to check the amino acid composition and the ratios of the amino acids either on or off the resin. The on resin hydrolysis performed at 130°C for 2h in propionic acid and 12M HCl (50:50) compares very favourably with normal off-resin hydrolysis with constant boiling 6N HCl[122]. The most difficult hydrolysis of peptide bonds are those of the carboxyl groups of isoleucine and valine with the most difficult being Ile-Ile and Val-Val. For example a hydrolysis of 192h is required to release all the Ile and Val in a peptide Tm-2 derived from papain where Ile-Ile and Val-Val occur once and twice respectively, the same number as in the 18 kDa protein[127]. Time course hydrolysis is often used to enable accurate estimates of

the amino acids[127]. The results from the resin hydrolysis will however give a better determination of Ile and Val.

The spectroscopic properties of peptides with formyl tryptophan and 2,4-DNP histidine are checked to ensure complete removal of these protecting groups. Mass Spectrometry is now becoming a powerful tool for the synthetic peptide chemist with the development of large mass visualisation and time of flight instruments.

2.1.2 IMMUNOLOGY: ANTIBODY AND T-CELL SITE RECOGNITION.[128]

2.1.2.1 INTRODUCTION.

The peptides synthesised in this chapter were used to screen regions of the 18 kDa protein to try and discover the immunologically active sites such as antibody binding regions and humoral immunity (T-cell) stimulation sites. These T-cell stimulation sites are important as they activate the memory cells which are involved in fighting infection. They also participate in suppressing an immune response as well as destroying infected cells. It is known that *M. leprae* can resist being killed by macrophages (killing cells) and proliferate within the cells which act to destroy antigenic material, thus evading the effector arm of the immune response. The large numbers of antibodies present in leprosy sufferers suggest that the B-cell (antibody producing cell) response is effective in recognising antigenic material from mycobacteria but for some reason the mycobacteria are not destroyed. One of the possible reasons for this is that the T-cells that regulate the level of the immune response by helping the B-cells to make antibodies and by stimulating the microbicidal and cytotoxic activity of other immune effector cells including the macrophages, are being prevented from functioning correctly due to either lack of communication between cells or non recognition of the potential antigens.

Each lymphocyte, either B-or T-cell, carries a surface receptor which is capable of recognising a particular antigen. The antigen receptor in B-cells is a membrane bound antibody which is ultimately secreted by the cell. B-cells are capable of recognising an unmodified antigen molecule either in solution or on the surface of the invading cells.

The T-cell antigen receptor is generated from a different set of genes which encode the cell surface receptor only. There are three main types of T-cells which are distinguished according to their function and surface proteins. These are the helper T_H,

suppressor T_S and cytotoxic T_C T-cells. T-cells only recognise antigen when it is presented to them in association with molecules encoded by the Major Histocompatibility Complex (MHC). Three types of molecule are associated with the MHC: Class I, Class II and complement components (Class III). Class I molecules are present on all mono nucleated cells and are involved in recognition of cytotoxic T-cells. Class II molecules present antigen to T_H cells. The presentation of antigen to the T-cells for recognition is performed by a group of cells called the Antigen Presenting Cells (APC). The cells which have receptors capable of recognising antigen, proliferate, usually accompanied by the release of antibodies, interleukins, and lymphokines. Macrophages release interleukin 1 (Il1) which acts on T-cells to produce interleukin 2 (Il2) and the expression of Il2 receptors. This causes the proliferation of T-cells, which have been previously activated by an antigen/MHC signal. T-cells also release lymphokines which act on B-cells (B-cell stimulating factors).

There are three main steps in an immune response. 1. The uptake processing and presentation of antigen to T and B-cells which is effected by the APCs. 2. Interaction between T-cells and the effector cells including macrophages (killing cells), B-cells and cytotoxic T-cells. These interactions are mediated by soluble factors or by direct cellular action. 3. The level and type of response is regulated by the T_S cells as well as APCs, T_H cells and antibody. The immune system then interacts with other cells such as the granulocytes, mast cells, fibroblasts and endothelial cells that produce the inflammation reactions aimed at eliminating the source of the antigen. Elimination of the antigen following the resolution of an infection limits further stimulation of the immune system, therefore the immune system as a whole is directed toward the elimination of antigen, producing a feedback inhibition of further activation of the antigen specific cells. The production of antibody feeds back on the antigen specific B-cells to limit further proliferation.

In addition to T_H cells there is a second set of T_H cells, idiotype specific which are responsible for selectively expanding populations of B-cells carrying particular idiotypes which are single antigenic determinants on an antibody variable region. Therefore T_H cells act at two levels: 1 antigen specific T-cells expand the population of antigen specific B-cells and 2 they selectively expand particular populations of B-cells (idiotype specific).

Some individuals have a high level of immune response while others have a low response, hence they are called high and low responders. The overall level of the immune response is determined by a large number of genes. The MHC Class II molecules are particularly important in this respect as they are not only involved in

interactions between the APC and T_H cells but also in the co-operation between T-cells and B-cells in the production of antibody responses to T-cell dependent antigens. In man it is found that T_H cells carry the surface molecule CD4 (T4) which is involved in the recognition of MHC Class II molecules while the CD8 (T8) molecule is related to the recognition of Class I molecules.

2.1.2.2 DEVELOPMENT OF ANTIBODY RESPONSES.

Antibodies are separated into subclasses IgM, IgG, IgA, IgD and IgE. Surface IgM is the first antibody to be expressed on the surface of B-cells and secreted IgM is the major component in the primary response to an invading antigen. The relatively low affinity of this antibody is offset by the multivalency so that the overall avidity of IgM antibodies is quite high especially when binding to antigens with multiple antigenic determinants. This sort of polyvalent antigen includes polymeric carbohydrates and peptidoglycans on bactericidal cell walls which are capable of inducing a B-cell response but not a T_H cell response.

The increase in antibody affinity is associated with a switch to IgG antibody production. IgG is the first line of defence to microorganisms and the most effective class of antibody for sensitising targets for antibody dependent cytotoxicity, where IgG is capable of inhibiting pathogens by virtue of its agglutinating ability. IgG is also the second line adaptor molecule which permits many other immune effector systems to be brought into play. IgG is the major component of the secondary immune response to T-cell dependent antigens. The fixed chain (Fc) fragment of the antibody's bind to an antigens on a bacteria and attaches to the Fc receptors of phagocytic cells, thereby facilitating the destruction of the bacteria. There are also receptors for IgG on B-cells and T-cells which are thought to involve immunoregulation.

2.1.2.3 ANTIGENICITY OF PROTEINS.

Antigenicity refers to the ability of a particular structure to be recognised by B or T-cells whereas immunogenicity refers to the ability of a molecule to induce an immune response and is dependent on a variety of regulatory influences such as the balance of the various types of T-cells, MHC, APC,and antigens. These factors relate to immunodominance, where reactions to certain potential antigenic regions of a protein predominate following immunisation with the native molecule.

There are two main types of epitopes, conformational where the folding pattern of the native protein is required and sequential which relies on the linear sequence of the

amino acids. Most of the accessible surface of a globular protein is potentially immunogenic and these proteins appear to consist of a continuium of multiple overlapping epitopes. The repertoire of antibodies binding to the surface of a protein is diverse with most amino acids in a protein having the potential to be detected immunologically. Protein antigenic sites recognised by antibodies consist of three dimensional arrays of amino acid side chains, presenting particular electron cloud shapes capable of forming ionic interactions with the antibody's binding site(paratope), which requires the native protein conformation for their integrity.

The ability of an antipeptide antibody to bind native protein appears to correlate better with the mobility of the target sites on the peptide chain than with the surface exposure of these sites. Antigenicity is related to the high local mobility of the amino acid chain, with many antigenic sites involving N terminal or C terminal residues, regions that frequently have high mobility. Many antibodies directed against antigenic peptides fail to bind native protein, those that do so bind to regions of local flexibility. With the mobility of amino acid side chains associated to antigenicity, the old idea of antibody binding to antigen through induced fit represented by a conformational change in the antibody might in fact be partly true but with conformational changes also occurring in the antigen. Major rearrangements of protein structure do not occur but small local conformational changes in the antigen and or antibody may take place in some instances.

Hence the recognition of antigen by antibodies is viewed as a multistep process rather than a simple lock and key complimentarity. Electrostatic forces and perhaps the breaking of salt bridges orientate the antibody to the epitope. The interaction with the side chains of amino acids may be followed by induced fit, involving small local changes in both epitope and paratope.

2.1.2.4 T-CELL ANTIGENS.

Antigenic determinants of T-cells are often different from those recognising antibodies. Early experiments using polyclonal T-cell populations indicated that native, denatured and often fragmented forms of the antigen could be recognised by many of the T-cells. This reflects the requirement of many T-cells that the antigen is processed into a form that can interact appropriately with MHC proteins and subsequently with the T-cell receptors. No simple ligand binding assay exists due to the three way interaction between antigen Class II molecule and T-cell receptors, therefore the biological responses of T-cells, such as proliferation or Il2 secretion are used as a measure of

antigen recognition. It will be some time before molecular information akin to antigen antibody interactions will be obtained for T-cell Class II antigen interactions.

With the requirement for antigen /Class II interactions in the stimulation of T_H cells it is most likely that T-cells recognise a conformational structure but whether or not the antigen is seen in a native, fragmented or processed form depends on which form is most capable of association with class II MHC molecules to provide a suitable determinant. In the case of cytotoxic T-cells conformational determinants on Class I molecules are recognised.

The identification of amino acids in the region of the epitope is important, allowing the determination of critical amino acids required for contact with the T-cell receptors, along with the appropriate region of the MHC protein. This is considered as the true epitope region. This then allows the determination of which amino acids are required for interaction with the Class II molecule and whether any other amino acids are important in stabilising the epitope in the best conformation, for instance an α-helix, in order to make contact with the Class II molecule and T-cell receptor. Data on T-cell epitopes for myoglobin and cytochrome c suggest that α-helical conformations are correlated with increased antigenicity. It has been suggested that amphipathic helixes with hydrophilic amino acid side chains on one face of the helix and hydrophobic amino acids on the other face (or amphipathic β-turns) are structures common in T-cell epitopes[129]. However some known T-cell epitopes do not fit this model and not all amphipathic regions of a protein provide T-cell epitopes. T-cell epitopes appear to be dependent on linear peptide sequences in which the native tertiary structure also is important, e.g. porcine insulin.

Different regions of a protein are recognised by the T-cells of different species and even between individuals within a species. Studies on lysozyme, β-glactosidase, myelin basic protein and others indicate that epitopes recognised by T_H and T_S cells also differ in many cases. The epitopic specificities of predominant T-cells can also significantly influence the specificity of B-cells stimulated to produce antibody during the immune response. There are more constraints on regions of a protein capable of stimulating T-cells and this could be the reason why there are fewer T-cell epitopes than B-cell epitopes. Certain regions of the protein constitute processing determinants that somehow influence the way a protein is processed and presented to T-cells with enzyme cleavage sites being obvious candidates for this sort of region. Other regions of the native protein could influence which of the potential T-cell epitopes are recognised. It must be emphasised that, as with the antibody response, host regulatory mechanisms

influence T-cell responses in such a way that the responses depend on whether the antigenic determinant stimulates T_S or T_H cells.[128]

2.1.2.5 SUMMARY.

In this chapter the details of the synthesis of twenty three peptides are discussed. These peptides were used to investigate the binding site for the L5 monoclonal antibody, which is used to specifically detect the presence of the 18kDa protein in isolates of *Mycobacterium leprae*.[54]. It was also possible to examine a potential major T-cell stimulation site of the recombinant 18 kDa protein[55]. IgG responses to the peptides have also been examined leading to the identification of four regions on the protein that are capable of raising IgG antibodies. For these regions to give IgG responses there must also be corresponding T helper cell epitopes that allow the production of the IgG. The identification of these T-cell epitopes has led to the hypothesis that there are two types of T-cell interactions based on the T-cell sub populations T_H1 and T_H2[130,131] as the responses between two strains of mice used in the study vary in their IgG responses.

2.2 EXPERIMENTAL METHODS.

2.2.1 General Experimental Methods.

2.2.1.1 PEPTIDE SYNTHESIS.

All peptides were synthesised on an Applied Biosystems 430A Peptide Synthesiser using either the standard protocols supplied by Applied Biosystems (Foster City, CA. disk part #400809, Version 1.4) or modified cycles supplied by S.B.H. Kent.

The following side chain protection for the amino acids was used, N^G-tosyl for arginine, benzoyl for aspartic acid and glutamic acid, N^{im}-dinitrophenyl for histidine, N -*p* -chlorocarbobenzoxy for lysine, benzyl for serine and threonine, N^{in}-formyl for tryptophan, and 2-bromocarbobenzoxy for tyrosine.

t-Boc amino acids were supplied by Bachem Fine Chemicals (Torrance CA.), Sigma (St Louis), Vega Biochemicals (Arizona), Peninsula Laboratories or the Protein Research Foundation (minoh-Shi, Osaka, Japan). Trifluoroacetic acid (TFA), (Halocarbon, USA) was glass distilled before use. Diisopropylethylamine (DIEA) (Aldrich,Germany) was distilled from ninhydrin then redistilled immediately before use. Dichloromethane (DCM), (BDH Chemicals, NZ) was distilled from P_2O_5 then redistilled prior to use. Dimethylformamide (DMF) was degassed over calcium hydride overnight then distilled under vacuum at 40-42°C immediately before use. The solutions of dicyclohexylcarbodiimide (DCC), (Protein Research Foundation or Peninsula Laboratories) (0.5mM) in DCM and hydroxybenzyltriazole (HOBt) (Protein Research Foundation or Peninsula Laboratories) (0.5mM) in DMF were prepared immediately prior to use.

The Boc amino acid PAM resins were purchased from ABI. The loading of the resins was between 0.4 and 0.78 mMol/gram of amino acid and for each synthesis typically 0.2-0.5 mMol of the Boc amino acid PAM resin was used. All peptides were synthesised using double couple cycles to minimise the chances of a failed coupling step. For peptides synthesised using Applied Biosystems protocols the Boc protecting group was removed with one 2min wash with TFA (33% in DCM) and one 19 min treatment with TFA (50% in DCM). Neutralization to give the free amino group was carried out with two 1 min treatments of DIEA (10% in DCM). Meanwhile the Boc protected amino acid was activated by DCC to give either the symmetrical anhydride or as in the case of

arginine, asparagine and glutamine the HOBt active esters. The amino acids were coupled in two mMol quantities. Treatment with DIEA (10% in DCM), to neutralise any remaining protonated amino groups[132b] and disrupt any peptide peptide hydrogen bonds was carried out before the recoupling step with a further 2mmol of amino acid. After the coupling of an amino acid, the unreacted amino groups were blocked by reaction with acetic anhydride (10% in DCM). A second synthesis of peptide 91-115 was carried out using ABI's NMP/HOBt protocols (Disk part 400674 Version 1.4), using the method described above with NMP replacing DMF, and dimethylsulphoxide (DMSO, 20% vol : vol) added to the coupling solution for the last 10 min of the coupling step to increase the swollen state of the resin. In this case all amino acids were coupled as the HOBt active esters and at the end of every coupling reaction the unreacted N-terminal was capped by reaction with acetic anhydride.

For the peptides synthesised under Steve Kent's protocols, the synthesiser was modified by removing the inline filters to the top and the bottom of the reaction vessel to prevent non-reproducible deliveries due to clogging of the filters (NB. This is not recommended by ABI). A modified 20 ml reaction vessel designed and made at the California Institute of Technology was used with the associated rapid synthesis protocols [113,132]. The chemistry involved in these rapid chemistry cycles closely parallels the Applied Biosystems protocols. A single solvent, DMF, was used in the reaction vessel throughout the synthesis in combination with flow washes to maintain the high purity of the solutions within the reaction vessel. Reactions were carried out using the highest concentration of reactants and neat TFA was used for the removal of the Boc group by first flowing the TFA through the reaction vessel to wash out the first 80% or so of the butyloxy radicals that are generated. All peptides synthesised using these protocols were routinely double coupled to ensure complete coupling. Capping was not used after the second coupling as the conditions available involved shrinking the resin from its swollen state in DMF.

2.2.1.2 NINHYDRIN ANALYSIS.

The coupling efficiency for each amino acid was monitored using the quantitative ninhydrin assay developed by Sarin *et.al* [120]. For the ninhydrin assay the chemicals were either purchased from Applied Biosystems or made up one day prior to use. Ninhydrin (2.5g) was dissolved in absolute ethanol (50ml). Phenol (40g) was mixed with absolute ethanol (10ml) and warmed until dissolved. The solution was then stirred with MB-3 Amberlite mixed bed resin (4g, ethanol washed) for 45min then filtered. Potassium cyanide (65 mg) was dissolved in H_2O (100 ml) and stirred with MB-3 amberlite for 1hr. After filtering off the resin, the cyanide solution (2 ml) was added to

pyridine (100ml) which had been distilled twice from ninhydrin immediately before use. All solutions were stored in dark bottles. Resin samples of approximately 4-8mg were removed automatically from the reaction vessel during the synthesis and stored under DIEA/DMF to reduce the possibility of acidolysis of some protecting groups, particularly the DNP group on histidine. The resin was filtered, washed 3-4 times with DCM/MeOH (50:50 vol : vol), dried by air suction and weighed. Phenol solution (75 µl), 0.2M KCN in pyridine (100 µl) and ninhydrin solution (75 µl) were addedto the resin. This mixture was heated at 100°C for 5 min before 5 mls of 60% EtOH was added to prevent further reaction. After cooling and vigorous mixing of the solution, the absorbance was measured against a blank (ie reagents only) at 510 nm on a PYE, PU 8610 Unicam UV/VIS spectrophotometer

2.2.1.3 CLEAVAGE AND PEPTIDE PURIFICATION.

After the removal of the protecting groups on histidine (DNP) and tryptophan (CHO) which are not removed by cleavage with HF, the Boc N-terminal group was removed from all peptides by treating the peptide resin with 65% TFA in DCM (1x 1min, 1 x 15 min). The peptide resin was then neutralised with a solution of 10% DIEA in DCM before HF cleavage. The HF cleavage was carried out using an all Teflon apparatus (Peptide Research Institute Japan). A sample of the peptide resin (usually 500 mg) was placed in the reaction vessel, with the appropriate nucleophilic scavengers. The scavenger mixture was dependent on the amino acid composition of the peptide. The peptide resin was chilled in liquid air under vacuum then HF(approx 10 ml) was distilled into the reaction vessel. The reaction vessel was warmed to 0° C and stirred for one hour. The HF was evaporated *in-vacuo* for 20 min. Then the peptide and resin were triturated with anhydrous diethyl ether (2x 20 ml) to remove any residual HF and the nucleophilic scavengers. The peptide was dissolved in the minimum possible concentration of aqueous acetic acid solution (usually 20-30%) and lyophilised. The purity of the lyophilised product was checked by analytical HPLC and purification was carried out by loading the crude peptide (approximately 50 mg) onto the preparative HPLC column. Fractions were taken across the solvent peak and across any other major peptide peak eluted off the column as detected at 214 nm. Each fraction was analysed on the analytical HPLC before combining the purest fractions and lyophilising to give the purified peptide. The lyophilised pooled fractions were checked for purity by analytical HPLC, amino acid analysis and if necessary sequencing of the peptide.

2.2.1.4 REMOVAL of the DNP GROUP from HISTIDINE.

2.2.1.4.1: Removal of the DNP group Prior to HF Cleavage.

The peptide resin (500 mg) was first swollen in DMF then treated with a solution of DIEA (5%) and β-mercaptoethanol (20%) in DMF(vol: vol) for 2X 30 min.[132] If the yellow colour persisted in the second treatment, the peptide resin was treated for a third time. An alternative method was used later where the β-mercaptoethanol was replaced by thiophenol for a single 1h treatment[134]. The peptide resin was washed with DMF, DMF:DCM (50:50) then DCM. The peptide resin was treated with TFA (65% in DCM) for 3 min then the peptide resin was filtered and treated for a further 15 min with the TFA solution. The peptide resin was reneutralised with DIEA (10% in DCM) for 3 min washed in DCM and dried before being cleaved. If the cleavage had to be done the following day the resin was dried under vacuum in a desiccator without neutralisation.

2.2.1.4.2: Removal of Residual DNP group after HF Cleavage.

The peptide was dissolved in 6M guanidine hydrochloride (guan.HCl) 25mM tris.HCl at pH 8 at a concentration of 10-20 mg/ml. β-Mercaptoethanol was added to 20% (vol : vol) and the solution was heated at 37° for 2hr or 50° for 1hr. The solution was loaded directly onto a gel filtration column (G-10) or directly onto the HPLC column to remove the β-mercaptoethanol-DNP complex from the peptide.

2.2.1.5. HIGH PERFORMANCE LIQUID CHROMATOGRAPHY.

Analytical High Performance Liquid Chromatography (HPLC) was carried out on a Waters HPLC system using Waters Model 510 pumps with an automated gradient controller, model 680. The column was maintained at a constant 35^oC in a Waters Temperature Controlled oven. Detection was achieved with a Model 441 detector with Extended Wavelength Module 440 for detection at 214 nm.

Samples were automatically injected using a Waters WISP Model 710. Data handling was performed by a Waters Data Module 730. The analytical column was a reverse phase Vydac C-4 column (4mm id x 250mm). All peptide analytical samples were chromatographed using linear gradients from initial eluent A: (2% acetonitrile (CH_3CN), 0.1% TFA, in H_2O) to eluent B: (0.1% TFA in CH_3CN) using flow rates of either 0.7ml per min or 1ml per min unless otherwise stated.

Preparative chromatography was carried out using Waters Model 510 pumps with Gradient Controller Model 660. Detection was by a Waters Variable Wavelength Detector Model 450. The fraction collector was a Pharmacia Frac100 model. The column was a C-18 Syncroprep column (10mm id x 250mm) packed in house. The conditions of elution of the peptides are summarised in Table 2.2.1. These conditions varied for each peptide, according to the peptides behaviour in the analytical HPLC of the crude peptide, after HF cleavage. The crude peptide solutions were filtered (0.2μm) and injected onto the column at a flow rate of 0.5 or 1ml/min. The eluents were the same as described above for the analytical column.

2.2.1.6 PEPTIDE AND PROTEIN SEQUENCING.

Sequencing of peptides and proteins was carried out on an Applied Biosystems 470A Protein Sequencer with an online Applied Biosystems 120A PTH Amino Acid Analyser. Lyophilised peptide was loaded onto the filter in 50% formic acid for sequencing.

2.2.1.7 AMINO ACID ANALYSIS.

Amino Acid Analysis (AAA) was carried out on a Beckman 119BL Amino Acid Analyser or an LKB Alpha Plus AAA with a Spectra Physics Integrator. Hydrolysis of peptides and proteins was carried out using 6M hydrochloric acid (HCl) containing 1% phenol in an evacuated Teflon capped tube at 110°C for 24 hrs[133]. AAA on peptide resin samples was done adding propanoic acid, degassing the resin and acid on a water vacuum line then adding concentrated HCl to make a 50:50 (vol/vol) solution. The tube was then flushed with nitrogen sealed and heated at 130°C for 2h[122,135]. For both types of hydrolyses the resulting sample was dried down *in-vacuo*. The residue was dissolved in H_2O or 2M citric acid solution to give a final concentration of approximately 25 nmole per 100 μl. Mole ratios of the amino acids were calculated relative to alanine, except in cases where alanine was attached directly to the resin, then calculations were relative to glycine.

Detection of tryptophan in AAA was carried out using 3M mercaptoethane sulphonic acid in an evacuated Teflon capped glass tube at 110°C for 24 hr [136]. The acid was neutralised with 3M sodium hydroxide and made up to the appropriate volume to give the 25 nmole per 100 μl final concentration.

2.2.2 PEPTIDES SLP-1 TO SLP-16.

The synthesis of the peptides was carried out using Kent's protocols as this would provide information on potentially difficult coupling steps that could occur in the total synthesis of the 18 kDa protein (Kent's cycles were to be used in the total synthesis of the 148 amino acid protein, as discussed in Chapter 3). The specific conditions for the synthesis, cleavage and purification of the peptides are summarised in Table 2.2.1. All of the peptides were handled as discussed in the general experimental conditions outlined above, with some exceptions that are noted below. The first three peptides were synthesised and purified while the the ABI 430A peptide synthesiser was being checked for its operation under Kent's protocols. Consequently the scale of the syntheses was kept small to minimise the costs of any failures. In later syntheses the scale was increased to 0.5 mMol to provide the maximum amount of peptide possible in one synthesis and to perhaps prevent their resynthesis in the advent of excessive losses during purification as with peptide SLP-8 and SLP-13. However it was discovered that as the peptide grew to about 15 amino acids in length, the peptide resin had increased in volume inhibiting the correct vortexing of the resin in the coupling solvent. Subsequently the scale was decreased to 0.4 mMol to allow the peptide resin to vortex in the reagents.

Two syntheses of peptide SLP-16 were carried out because it was discovered in the first synthesis that there was a significant amount of a prematurely terminated peptide. As the synthesis had failed using Kent's protocols it was decided to repeat the synthesis of this peptide with ABI's NMP/HOBt protocols but this gave a slightly worse result than the first synthesis. This added strength to the decision to use Kent's protocols for the total synthesis of the 18 kDa protein.

There were also changes in the HF cleavage of the peptides when a smaller 50 ml cleavage vessel was purchased. This provided cleaner crude peptide material than the 250 ml vessel which was more suitable to larger scale cleavages. The size of the cleavages was kept around the 500 mg size as the smaller quantities of peptide resin gave less pure crude peptides as judged by analytical HPLC. For the peptides containing tryptophan, the HF cleavage was carried out with 1,4-butanedithiol present as a scavenger to enable the removal of

SLP-9 61-80	Boc Ala 0.780mMol/g	0.4mMol 0.513g	498	9	*p* -Cresol(0.5ml) *p* -thiocresol(0.5ml)	299	90	0-40 2h 1.5ml/min	50.9
SLP-10 76-95	Boc-Glu 0.720mMol/g	0.4mMol 0.550g	501	9	*p* -Cresol(0.8ml) *p* -thiocresol(0.2ml)	277	47.5	10-40 2h 1.5ml/min	22.4
SLP-11 91-110	Boc Val 0.780mMol/g	0.4mMol 0.512g	498	9	Anisole (1ml)	239	98.7	10-40 2h 1.5ml/min	27.0
SLP-12 106-125	Boc-Lys(ClZ) 0.64 mMol/g	0.4mMol 0.629g	498	9	Anisole (1ml)	214	63	0-40 2h 1.5ml/min	33.0
SLP-13 121-140	Boc-Lys(ClZ) 0.64 mMol/g	0.4mMol 0.625g	498	9	Anisole (1ml)	293	60	0-40 2h 1.5ml/min	14.3
SLP-14 136-148	Boc Ala 0.780mMol/g	0.5mMol 0.662	504	9	Anisole (1ml)	214	51.1	0-40 2h 1.5ml/min	10.1
SLP-15 101-125	Boc-Lys(ClZ) 0.64 mMol/g	0.5mMol 0.781g	302	11	Anisole (1ml)	131.6	49.7	0-40 2h 1.5ml/min	19.6
SLP-16 #1 91-115	Boc-Ile 0.73 mMol/g	0.5mMol 0.730g	502	11	Anisole (1ml)	230.8	50.7	0-50 2h 1.5ml/min	22 and 16
SLP-16 #2 91-115	Boc-Ile 0.73 mMol/g	0.5mMol 0.685g	500	11	Anisole (1ml)	178.8	46.6	0-40 2h 1.5ml/min	11 and 18

Table 2.2.1. Experimental detail for peptides SLP-1 to SLP-16 giving details of the scale of the synthesis, the method of HF cleavage, yields of peptide, the preparative HPLC conditions and the final yield of purified peptide

the formyl group on tryptophan in a single step cleavage instead of applying the two step Lo-Hi HF cleavage method[137]. This was done to reduce the handling time associated with the cleavage of the large number of the small peptides.

The data presented for peptides SLP-4 and SLP-5 are from the initial cleavages of these peptides. In the subsequent cleavages of peptide SLP-4, the size of the vessel was changed to 50 ml and the amount of HF decreased to 9 ml, with the quantity of peptide resin remaining the same. The scavengers were changed to those used in the cleavage of peptide SLP-10, which significantly improved the yield and the HPLC profile of the crude peptide. For peptide SLP-5 the size of the HF vessel was also changed to the smaller size with half the quantity of scavengers.

Problems were encountered with the pretreatment of peptide SLP-13 to remove the DNP group from histidine. It was noticed that the β–mercaptoethanol solution used to remove the DNP group went a dark red colour which remained throughout the first treatment instead of the normal yellow colour. There was residual DNP group present in the crude peptide as judged by the analytical HPLC of the crude peptide. Therefore the crude peptide was treated in solution to remove the residual DNP then gel filtered. This also failed to effectively remove the DNP group as did treatments of the peptide resin with thiophenol.

The yields for most of the crude peptides were as expected (calculated from the substitution level of the peptide on the resin) except in the case of peptide SLP-8, and peptide SLP-13. For peptide SLP-8 the yield was decreased due to not being able to dissolve the peptide from the resin even though the amount of acetic acid was 70%. In the case of peptide SLP-13 the yield was much higher than expected probably due to the presence of DNP-β-mercaptoethanol complex still associated with the crude peptide.

The yield of the purified peptides from the preparative HPLC column was reasonably consistent given the same loading conditions. Crude peptides were dissolved in buffer A where possible otherwise the peptide was dissolved by the addition of buffer B to give a 5% to 10% buffer B solution. If the crude product still would not dissolve, the peptide was dissolved by the addition of 10% to 20% acetic acid. A maximum of 50% acetic acid was used in the case of peptide SLP-16. Peptide SLP-8 however was very difficult to dissolve even in high percentages of acetic acid or 6M guanidine HCl. There was also difficulty in eluting this peptide off the column as shown by the requirement for up to three blank gradient runs on the HPLC to completely remove the peptide from the column. In the case of peptide SLP-13 the low yield is a result of

having to cut the peptide peak at the very leading edge to avoid any DNP associated peptide contaminating the pure peptide.

2.2.3 Immunology Assay Experimental

This section of work and section 3.2.3.11 was carried out in Professor J.D. Watson's laboratory in the Molecular Medicine Department at the Auckland Medical School under the supervision of Dr R.L.Prestidge and has been published.[54, 55, 130, 131] At the conception of this work it was intended that there would be some participation by the writer in this area particulaly with the antibody assay s and the western blots (section 3.2.3.11), but this did not eventuate.

2.2.3.1 ANIMALS.

Strains of mice were purchased from the Jackson Laboratory, Bar Harbor, ME. BALB/cJ (H-2^d), C57BL/10J (H-2^b), B10.M (H-2^f), B10.BR (H-2^k) and BALB.B (H-2^b) were maintained in breeding facilities of the department of Molecular Medicine , Auckland University. For each experiment, 6-10 week old mice were used. F1 hybrid mice were bred at Auckland University.

2.2.3.2 ENZYME LINKED IMMUNOSORBANCE ASSAY.

Serum from mice was obtained by bleeding from the tail at three and six days after each injection of antigen, clotting the blood at 4°C and removing particulate material by centrifugation. Serum was divided into portions and either used immediately or frozen at -70°C.

Antibody-binding assays were carried out using an Enzyme Linked Immunosorbance Assay (ELISA). The peptides were bound to 96-well plastic immunoassay plates (either Nunc, Roskilde, Denmark or Linbro, Herts, UK). Peptide binding was found to be optimal over the range of concentrations from 12.5-25 μg per well, while 18 kDa protein binding was optimal over the range 100 to 500 ng per well. A colorimetric assay[138] allowed the precise amount of antigen bound to immunoassay plates to be assessed. Binding was performed using a bicarbonate coating buffer (15 mM sodium carbonate, 35 mM sodium hydrogen carbonate, pH 9.8) and incubating at 37°C overnight. Immunoassay plates were then washed once with phosphate buffered saline (PBS) plus 0.05% Tween 20 (250 μl per well), before being blocked by the addition of 250 μl of 1% bovine serum albumin per well and incubation for 1 hr at room

temperature. The plate was then washed three times with PBS plus 0.05% Tween 20 for three minutes per wash. The primary antibody (usually diluted serum prepared as above in 100 μl of PBS / Tween 20) was then allowed to bind for one hour at room temperature, before being removed, and the plate washed three times as before. The secondary antibody (biotinylated anti-mouse IgM or IgG, Becton Dickinson, CA diluted 1/2000 in 100 μl of PBS/Tween per well), was allowed to bind for 1 hr at room temperature then removed and the plate washed three times with 250 μl of PBS/Tween 20 per well as before. Streptavidin biotin horse -radish peroxidase conjugate (Becton Dickinson, CA diluted 1/1000 in 100 μl of PBS) was allowed to react for one hour at $37^{o}C$, then removed and the wells washed three times for three minutes each time with 250 μl of PBS per well. The colour was developed using 100 μl per well of 1% 1,2-phenylenediamine dihydrochloride in 25mM citric acid, 50 mM disodium hydrogen phosphate, 0.00012% hydrogen peroxide for one to five minutes in the dark. The reaction was stopped by the addition of HCl to a final concentration of 3.3% and the absorbance immediately read in a Sci-Med Kontron SLT 20 automated ELISA reader at a wavelength of 486 nm.

2.2.3.3 DATA ANALYSIS.

For each experiment , groups of 3-5 mice were immunised with identical amounts of antigen and bled separately. Each individual mouse serum was separated, analysed and the mean titre expressed together with standard deviations. The titre of IgG expressed in each figure is the lowest dilution of sera at which absorbance measurements were similar to that of sera of non-immunised mice.

2.2.3.4 PURIFICATION of the 18 kDa PROTEIN ANTIGEN.

The recombinant plasmid pML3, which contains the 18kDa gene of *M.Leprae*, was used to transform the *E.coli* strain DH5α. Cells were grown to late log phase, harvested by centrifugation, and the 18 kDa protein extracted and purified as detailed[52]. The purified 18 kDa protein was checked for purity by polyacrylamide gel electrophoresis and was consistently of greater than 95% purity.

2.2.3.5 IMMUNISATION.

Antigens were prepared by dissolving lyophilised material in PBS, (150 mM NaCl, 2 mM NaH_2PO_4, 2 mM Na_2HPO_4,pH7), or in 1% acetic acid. For primary immunisation, the standard protocol was to inject antigen at the base of the tail in 100 μl of an emulsion comprised of 50% Freunds Incomplete Adjuvant (FIA) and 50% PBS.

Subsequent challenges with antigen were delivered intraeritoneally in 100 μl of saline at ten days after the initial immunisation and at nine day intervals thereafter.

2.2.3.6 T CELL PROLIFERATION ASSAY.

Mice were immunised in the base of the tail with 50 μl or 100μl of a mixture containing either 10 mg of peptide or 10 mg of 18 kDa protein in saline, or at the time intervals detailed in the text, emulsified in an equal volume of FIA. Seven days after immunisation mice were sacrificed and para-aortic plus inguinal lymph node cells (LNC) were prepared in medium, washed twice and diluted to a final concentration of 2×10^6 cells/ml. For microcultures 200 μl of this suspension was cultured with 0-50 μg/ml of 18 kDa protein in 96 well flat-bottomed microtitre plates (Nunc, Roskilde, Denmark) for 4 days at 37°C in an atmosphere of 10% CO_2. Microcultures were then radiolabelled for 8 hr with 0.25μCi of ^{3}H-thymidine (^{3}H-Tdr) (New England Nuclear, Boston, US) and subsequently harvested onto glass microfibre filters (Whatman, Singapore). Radioactive incorporation was measured using a liquid scintillation counter. The results were expressed as the mean of triplicate cultures and in each case the standard error was less than 10% of the mean.

2.3 RESULTS and DISCUSSION

2.3.1 PEPTIDE SYNTHESIS.

2.3.1.1 PEPTIDES SYNTHESISED PRIOR TO THE COMMENCEMENT OF THIS WORK.

Some peptides discussed in this work were synthesised in the initial phase of the project and used as a starting point for the development of the immunological assays to detect immunologically reactive areas of the protein. Peptides RLP-23,24,25,29,30, and 31 were synthesised on a Schwartz Mann peptide synthesiser that was controlled by a punched-paper tape computer. The peptides were synthesised with capping protocols and were cleaved in the normal manner as described in the general experimental methods section. The peptides were however dissolved in neat TFA after the cleavage, evaporated *in-vacuo* , diluted with H_2O or H_2O/ acetic acid and freeze dried. The peptides were then dissolved in 50% acetic acid and gel filtered on a G-25 Sephadex column with the front half of the first peak, detected at 280 nm, being collected. AAA on the isolated peptides were used to confirm the amino acid compositions. The peptides were generally not purified beyond this stage. Analytical HPLC run on some of the peptides stored for over a year at -18°C showed that other peaks were also present with the peptide peak. HPLC purification was done by R.L. Prestidge in Auckland on the peptides that did not exhibit one main peak in an analytical HPLC of the peptide.

Peptide RLP-39 was synthesised and purified by Gibson [139] prior to the start of this work. The peptide was synthesised using the ABI synthesiser with the standard program (disk 400809, version 1.4) altered to allow capping. The purification of the peptide using a FPLC (Pharmacia) equipped with a RP Pep (HR 5/10) column, showed a series of multiple peaks. Though the size of the peptide is fairly large, 40 amino acids, for synthesis under ABI protocols (ABI recommends the instrument has a capacity for synthesis of a length of 50 amino acids), a comparison of the crude HPLC profiles for this peptide compared unfavourably with the 50 mers as described in Chapter 3. This provided further impetus for the choice of using Kent's synthesis protocols for the synthesis of the total 18kDa protein and the larger fragments (Chapter 3).

Amino Acid	Expect Ratio	Purified Peptide Relative to Ala
Ser	2	2.09
Glx	2	2.16
Gly	1	1.12
Ala	1	1.13
Val	1	1.00
Ile	2	1.95
Leu	3	3.10
Tyr	1	1.08
Lys	1	1.13
Arg	1	0.99

Table 2.3.1: AAA peptide SLP-1 (101-115).

Amino Acid	Expect Ratio	Purified Peptide Relative to Ala
Ser	1	0.97
Glx	1	0.96
Pro	2	1.87
Ala	2	2.00
Val	1	0.97
Ile	1	0.93
Leu	2	1.94
Lys	3	2.84
Arg	2	2.06

Table 2.3.2: AAA peptide SLP-2 (110-125)

Amino Acid	Expect Ratio	Purified Peptide Relative to Ala
Asp	3	2.92
Ser	1	0.96
Pro	1	1.05
Gly	2	1.90
Ala	1	1.0
Val	1	0.93
Ile	1	0.81
Lys	2	1.76
Arg	2	2.08

Table 2.3.3: AAA of peptide SLP-3 (115-134)

2.3.1.2 PEPTDES SYNTHESISED

2.3.1.2.1 Peptides SLP 1-4

Sequences:
SLP-1: (101-115): Arg Ile Leu Ala Ser Tyr Gln Glu Gly Val Leu Lys Leu Ser Ile

SLP-2: (111-125): Leu Lys Leu Ser Ile Pro Val Ala Glu Arg Ala Lys Pro Arg Lys

SLP-3: (125-134): Ala Lys Pro Arg Lys Ile Ser Val Asp Arg Gly Asn Asn Gly

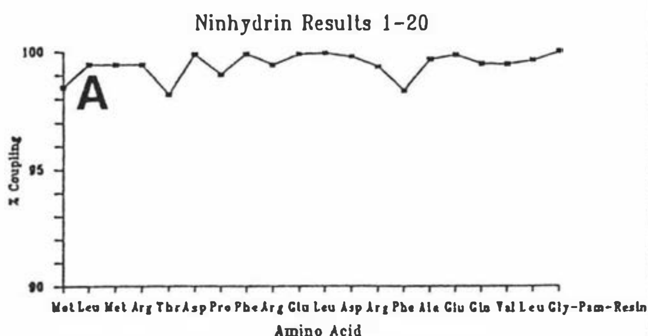
SLP 4: (109-125): Gly Val Leu Lys Leu Ser Ile Pro Val Ala Glu Arg Ala Lys Pro Arg Lys

These peptides were synthesised at the very early stages of adapting the instrument to Kent's protocols where the flow rates and mixing times were adjusted to account for the pressure differences between instruments, preventing amino acids like asparagine from foaming over the top of the activator and concentrator vessels. Purification protocols were also being tested to decide which HPLC columns and gradients would allow the purification of the majority of the peptides. Therefore the yields were not maximised for these earlier peptides. In addition the auto sampler was not removing consistent size resin samples for ninhydrin analysis at this stage. This caused inaccurate readings due to the samples having a concentration of detectable amino groups below the threshold for the assay. Average coupling percentages determined on the samples that were removed were 99.65, 98.94, 99.31 and 99.62% respectively for four the peptides. The yields of the peptides are given in Table 2.2.1. The AAA figures (Table 2.3.1-3) for the purified peptides were in agreement with the expected ratios for the peptides. For peptide SLP-4 no AAA was done but the peptide was sequenced to confirm amino acid integrity of the peptide.

2.3.1.2.2 SLP-5 (1-20)

Sequence:
Met Leu Met Arg Thr Asp Pro Phe Arg Glu Leu Asp Arg Phe Ala Glu Gln Val Leu Gly

After completion of the synthesis the dry weight of the resin was 2.154 g (a weight increase of 1.5 g). The ninhydrin analysis indicated that the average coupling was 99.43% with an overall yield of the peptide of 90.4% (Figure 2.3.1). AAA of the



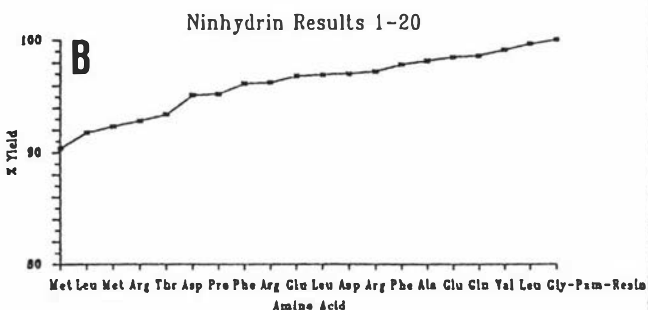


Figure 2.3.1 a) % Coupling for peptide 1-20 b) Cumulative predicted yield for the peptide on the resin

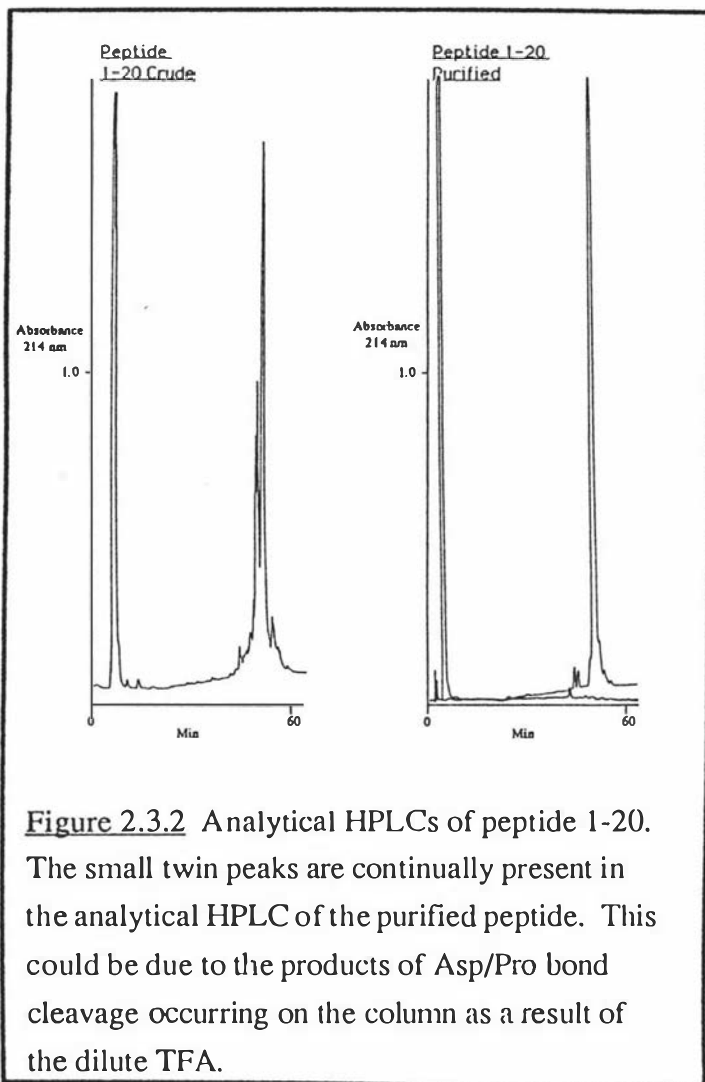


Figure 2.3.2 Analytical HPLCs of peptide 1-20. The small twin peaks are continually present in the analytical HPLC of the purified peptide. This could be due to the products of Asp/Pro bond cleavage occurring on the column as a result of the dilute TFA.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide Relative to Ala
Asx	2	1.88	1.93
Thr	1	0.84	0.93
Glx	3	3.00	2.94
Pro	1	0.88	0.93
Gly	1	1.05	1.07
Ala	1	1.0	1.0
Val	1	1.12	0.76
Met	2	0.87	1.76
Leu	3	2.34	2.77
Phe	2	1.97	1.95
Arg	3	2.95	3.21

Table 2.3.4: AAA peptide SLP-5 (1-20)

peptide- resin gave the correct ratio for the amino acid composition of the peptide (Table 2.3.4). Crude peptide 213mg (50.7%) of was obtained after cleavage of 598 mg of peptide resin. Preparative HPLC on 50 mg of crude peptide gave 15.1mg (30.2% based on the crude peptide). Analytical HPLC (Figure 2.3.2) and AAA (Table 2.3.4) showed the peptide to be greater than 95% pure.

In cleaving other peptides containing methionine a slightly different protocol was used which significantly improved the purity of the peptides containing methionine as judged by analytical HPLC. Subsequent cleavages of this peptide SLP-5 were carried out using the method for the cleavage of peptide SLP-10. This significantly improved the HPLC profile of the crude peptide giving increased yields of the purified peptide.

2.3.1.2.3 SLP-6 (16-35)

Sequence:
Glu Gln Val Leu Gly Thr Ser Ala Arg Pro Ala Val Met Pro Met Asp Ala Trp Arg Glu .

After completion of the synthesis the dry weight of the peptide resin was 2.039 (1.35 g weight increase). The ninhydrin analysis indicated an average coupling yield of 99.68% with predicted yield of 94.22% for the complete peptide (Figure 2.3.3). AAA of the peptide-resin gave the expected amino acid ratios (Table 2.3.5). Crude lyophilised peptide 286 mg (72.2% based on the dried peptide resin) was isolated after cleavage of 600 mg of peptide resin. Preparative HPLC on 50 mg of crude peptide gave 13.5 mg (27.2%) of purified peptide. Analytical HPLC (Figure 2.3.4) and AAA (Table 2.3.5) indicated a purity of greater than 95%. A check of the UV spectrum (Figure 2.3.5) of the purified peptide showed the peak height at 300nm being less than the peak height at 280nm which indicated the complete removal of the formyl group from tryptophan[140]by 1,4-butanedithiol[141].

2.3.1.2.4 SLP-7 (31-50)

Sequence:
Asp Ala Trp Arg Glu Gly Glu Glu Phe Val Val Glu Phe Asp Leu Pro Gly Ile Lys Ala

After completion of the synthesis the dried peptide-resin weighed 1.61 g (a 1.08 g weight increase). The ninhydrin analysis gave an average coupling yield of 99.77% with a cumulative coupling yield of 95.58% (Figure 2.3.6). AAA of the peptide resin

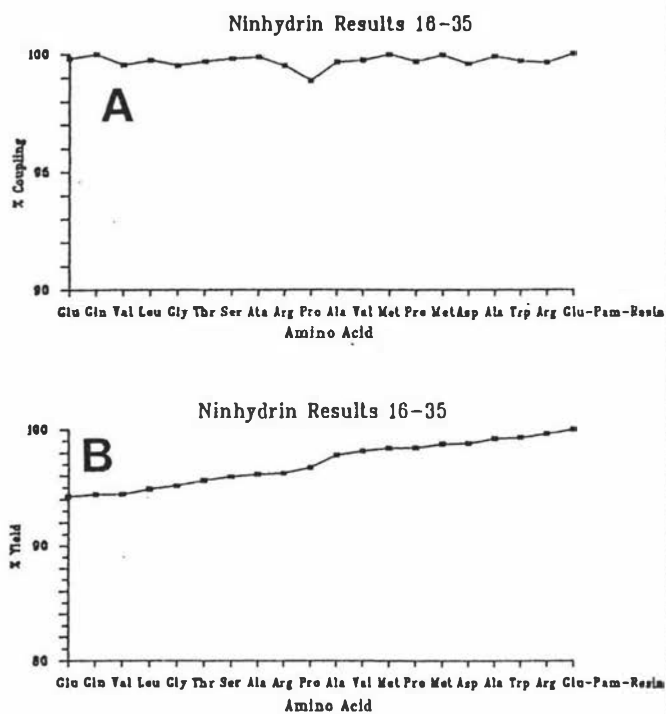


Figure 2.3.3 a) % Coupling for peptide 16-35 b) Cumulative predicted yield for the peptide on the resin

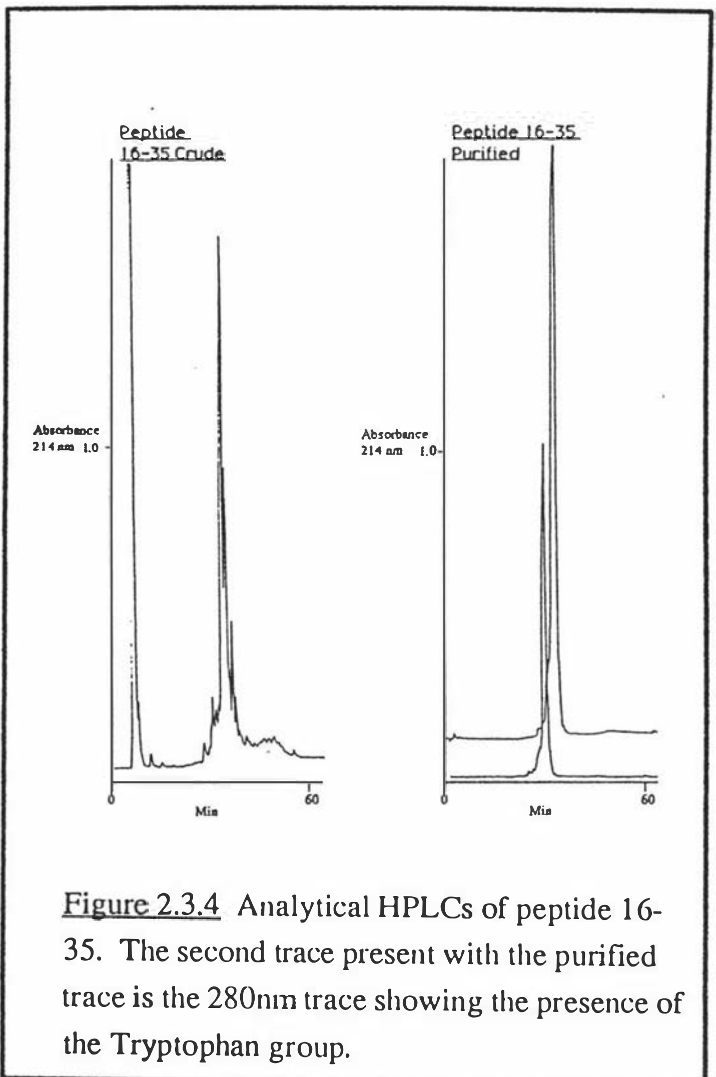
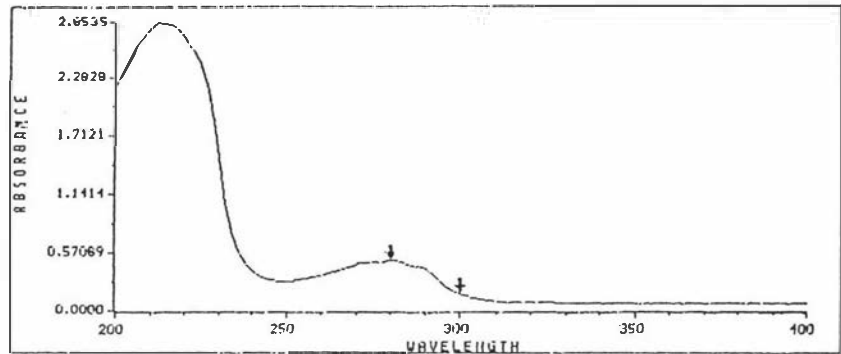


Figure 2.3.4 Analytical HPLCs of peptide 16-35. The second trace present with the purified trace is the 280nm trace showing the presence of the Tryptophan group.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide Relative to Ala
Asx	1	1.08	0.98
Thr	1	0.99	0.93
Ser	1	0.55	0.96
Glx	3	2.90	2.73
Pro	2	1.90	1.86
Gly	1	0.95	1.02
Ala	3	3.0	3.0
Val	2	1.99	1.34
Met	2	1.26	1.79
Leu	1	0.95	0.63
Trp	1	ND	0.85
Arg	2	2.14	2.01

Table 2.3.5 AAA peptide SLP-6 (16-35). Detection of Trp was by hydrolysis with 3N Methanesulphonic acid.



Annotated Wavelengths:
1 : Wavelength = 280 Result = 0.486
2 : Wavelength = 300 Result = 0.157

Figure 2.3.5 UV spectra of the purified peptide 16-35 showing that there is no sign of the formyl tryptophan group which gives a large peak at 300nm.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide Relative to Ala
Asx	2	2.08	2.01
Glx	4	4.09	3.97
Pro	1	1.10	0.93
Gly	2	2.11	2,05
Ala	2	2.0	2.00
Val	2	1.81	0.66
Ile	1	0.96	0.66
Leu	1	1.07	1.01
Phe	2	2.05	1.90
Trp	1	ND	ND
Lys	1	1.03	0.79
Arg	1	1.09	0.86

Table 2.3.6: AAA peptide SLP-7 (31-50). Low Val content as a result of Val-Val bond as discussed in section 2.1.

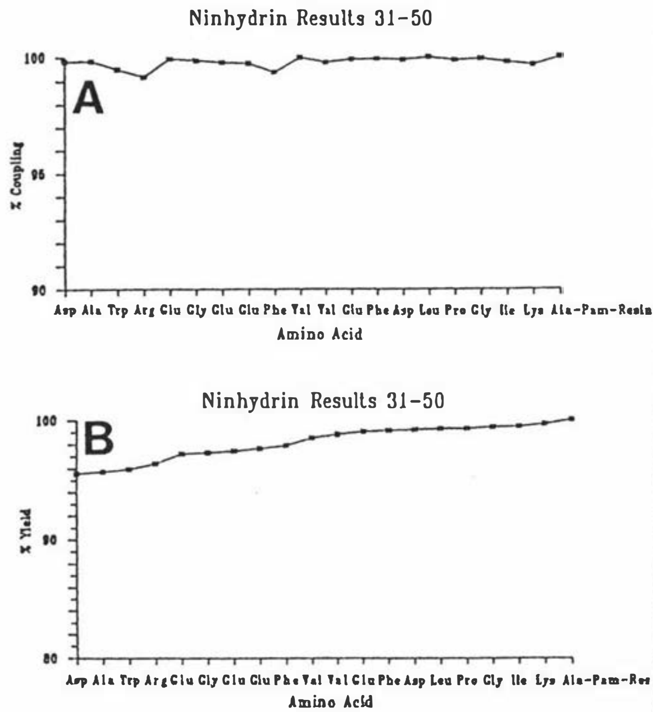


Figure 2.3.6 a) % Coupling for peptide 31-50 b) Cumulative predicted yield for the peptide on the resin

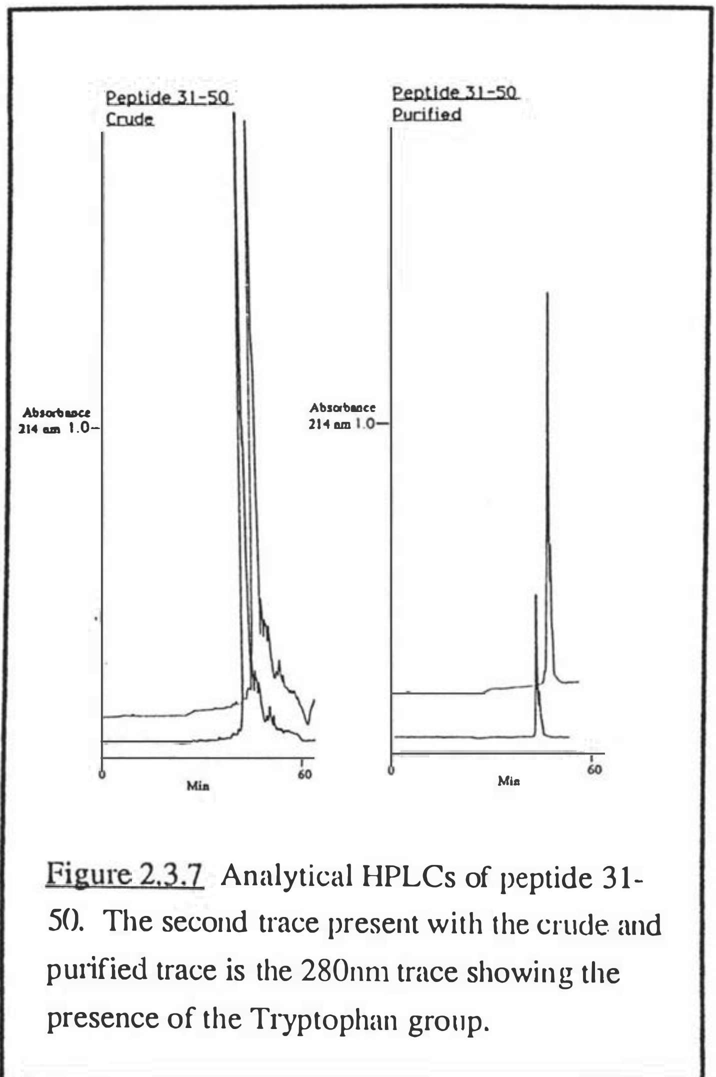


Figure 2.3.7 Analytical HPLCs of peptide 31-50. The second trace present with the crude and purified trace is the 280nm trace showing the presence of the Tryptophan group.

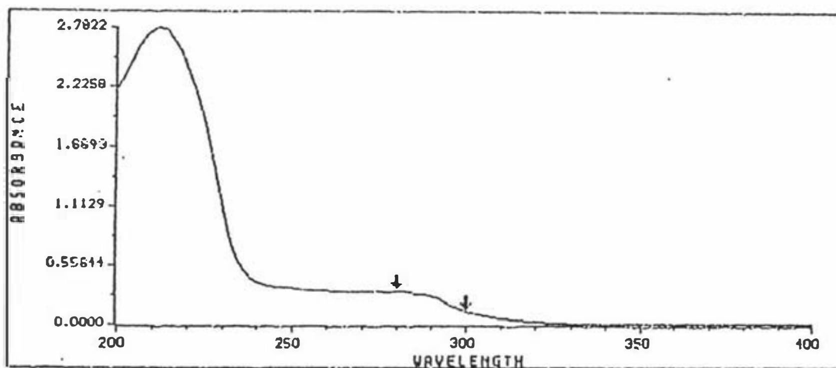


Figure 2.3.8 UV spectra of the purified peptide 31-50 showing that there is no sign of the Formyl Tryptophan group which would give a large peak at 300nm.

gave the expected ratios of amino acids (Table 2.3.6). The HF cleavage of 500 mg of peptide resin gave 250 mg of crude peptide (74.3% yield). Preparative HPLC on 65 mg of crude peptide gave 46.3 mg (69% yield) of purified peptide. AAA (Table 2.3.6) and analytical HPLC (Figure 2.3.7) showed that the peptide was greater than 95% pure. Examination of the UV spectrum (Figure 2.3.8) of the purified peptide showed that the ratio of the absorbances at 280 and 300nm indicated the complete removal of the formyl group on tryptophan.

The initial synthesis of this peptide failed to couple the phenylalanines at position 8 and 12 as detected by ninhydrin analysis (% couplings were 70 and 40% respectively) because the amino acid failed to dissolve in the delivery cartridge. It was discovered that DMF (1 ml) needed to be added to the cartridge to dissolve the phenylalanine. This problem occurred with one source of Boc-phenylalanine (Vega) and no such problems were encountered with other sources.

2.3.1.2.5 SLP-8 (46-65)

Sequence:
Pro Gly Ile Lys Ala Asp Ser Leu Asp Ile Asp Ile Glu Arg Asn Val Val Thr Val Arg

At the completion of the synthesis the dry weight of peptide-resin was 1.74 g (an increase of 0.933 g). The average coupling yield by ninhydrin analysis was 99.46% with an overall peptide yield of 92.7% (Figure 2.3.9). The AAA of the peptide-resin showed the amino acids were in the correct ratios for the peptide (Table 2.3.7). From the HF cleavage, only 170 mg (62.7%) of crude peptide was isolated because the peptide was very difficult to dissolve in dilute acetic acid. Although the acetic acid concentration went up to over 70% there was still some peptide left behind as shown by the foaming in water as the funnel was washed out. Purification of this peptide by preparative HPLC presented some difficulties. To dissolve the crude peptide greater than 50% acetic acid was required or a combined mixture of 6M guanidine.HCL and formic acid. Purification of 90 mg of the crude peptide gave 17.7 mg (19%) of pure peptide. AAA (table2.3.7) and analytical HPLC (Figure 2.3.10) showed the peptide to be about 85% pure. Due to the losses incurred in the purification to this stage no attempt was made to purify the peptide further.

Though the same amount of peptide (150 µl of a 2mg/ml solution) was loaded on the analytical column, the peak height of this peptide was considerably reduced in comparison with other peptides. Despite increasing the amount of peptide loaded the size of the peaks remained reduced in absorbance height. Up to 50% of the peptide was

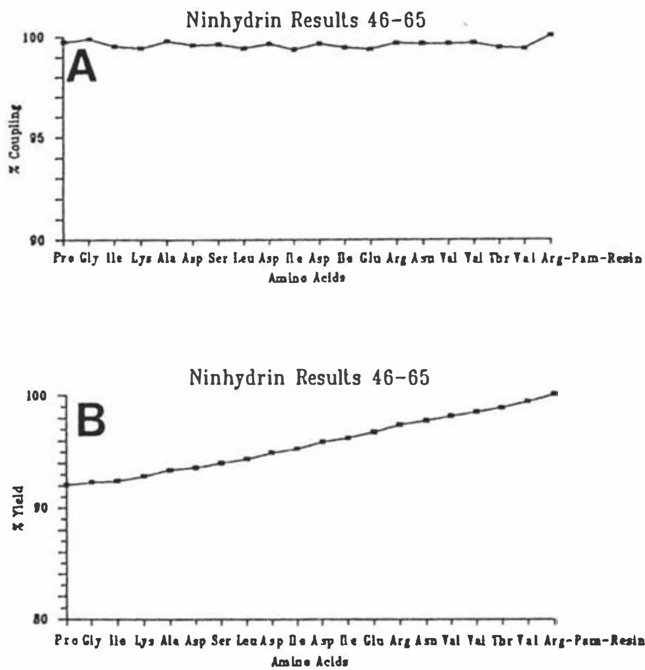


Figure 2.3.9 a) % Coupling for peptide 46-65 b) Cumulative predicted yield for the peptide on the resin

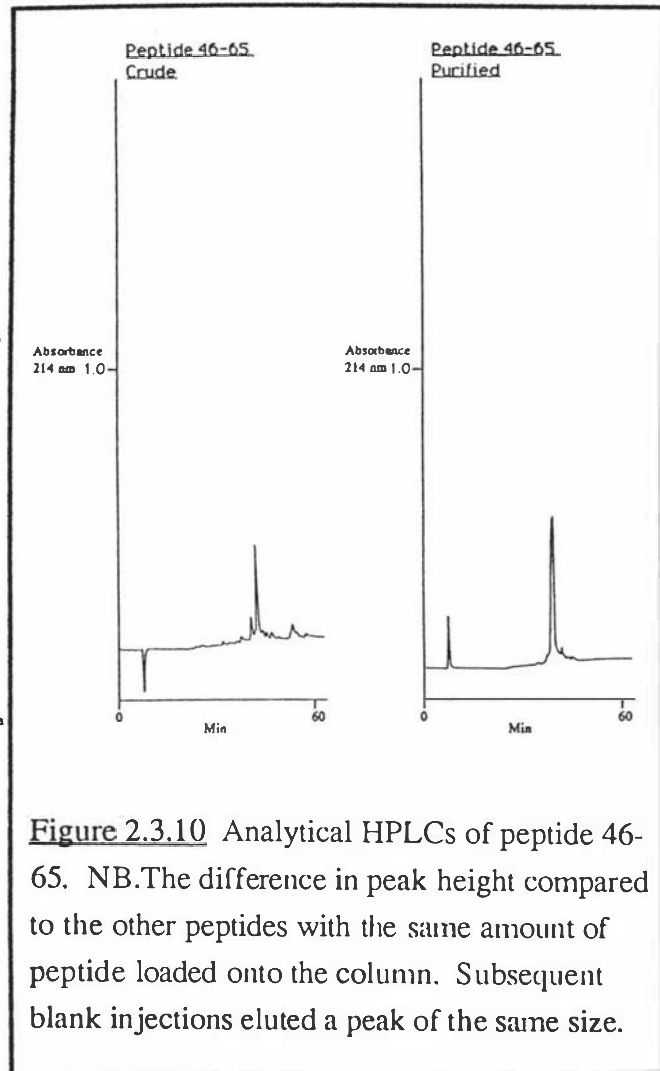


Figure 2.3.10 Analytical HPLCs of peptide 46-65. NB.The difference in peak height compared to the other peptides with the same amount of peptide loaded onto the column. Subsequent blank injections eluted a peak of the same size.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide Relative to Ala
Asx	4	3.79	3.71
Thr	1	0.93	0.79
Ser	1	0.48	0.89
Glx	1	1.01	0.84
Pro	1	0.89	0.91
Gly	1	0.89	0.98
Ala	1	1.00	1.00
Val	3	2.63	1.37
Ile	3	2.37	2.21
Leu	1	0.98	0.98
Lys	1	0.93	0.64
Arg	2	1.93	1.93

Table 3.2.7: AAA peptide SLP-8 (46-65). Val-Val bond causing low Val content as discussed in section 2.1.

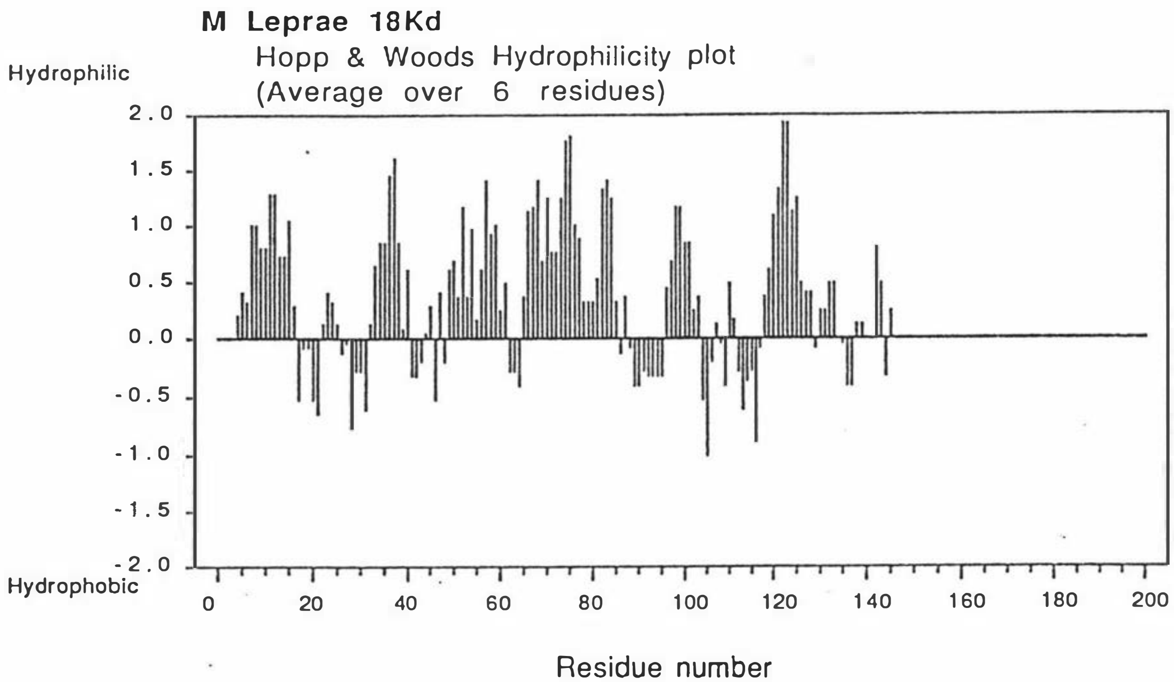


Figure2.3.11 Hopp and Woods Hydrophilicity plot for the 18kDa protein from the Mac Pro Mass program [142] The region 46-65 shows a predicted hydrophilic nature.

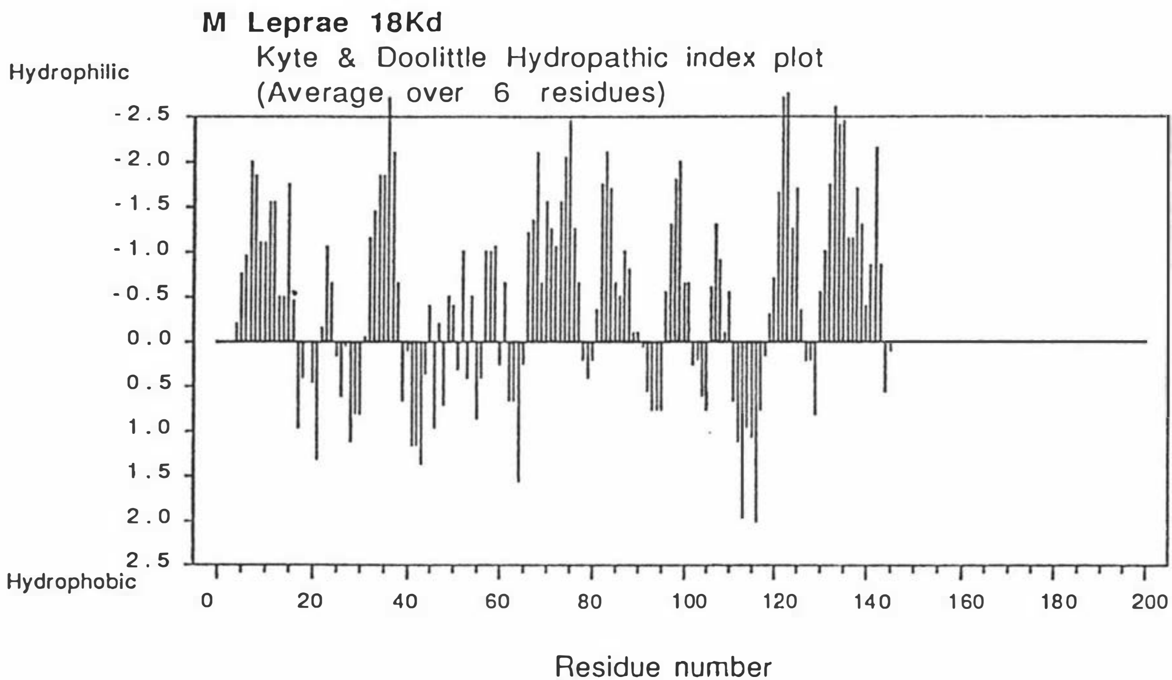


Figure 2.3.12 Kyte and Doolittle Hydropathic Index plot from the Mac Pro Mass program [143]. The region 46-65 shows a greater hydrophobic nature compared to the Hopp and Woods plot.

retained on the HPLC columns as shown by the elution of more peptide material in subsequent (up to 3) blank injections on the HPLC columns. By examining the hydrophilicity plot (Hopp and Woods[142]) (Figure 2.3.11) of the protein, the peptide is predicted to be slightly hydrophilic which does not explain the almost total hydrophobic nature of this peptide. The Kyte and Doolittle[143] plot (Figure 2.3.12) does predict some degree of hydrophobicity. However,the secondary structure prediction plot of Garnler-Osguthorpe and Robson[144] predicts that in the region of this peptide (residues 53-60) an alpha helix would be formed. If the peptide does form an α-helix in this region then it is possible that the residues Leu^{53},Ile^{57} and Val^{61} being 4 residues apart would form an extremely hydrophobic face on one side of the helix which would cause the peptide to absorb strongly to the HPLC column. This has been observed previously by Knighton who designed alpha helical peptides with a hydrophobic face on one side, which caused peptides to behave similarly to peptide SLP-8 on an HPLC column.[145]. The more commonly used Chou-Fasman[146] plot for the protein did not predict formation of a helix in this region. Another factor that needs to be considered also is the prediction of both secondary structure programs that the tail end of this peptide will form a β-sheet. Aggregation of the peptide brought about by this structural motif could also play a part in the peptides behaviour on the HPLC column.

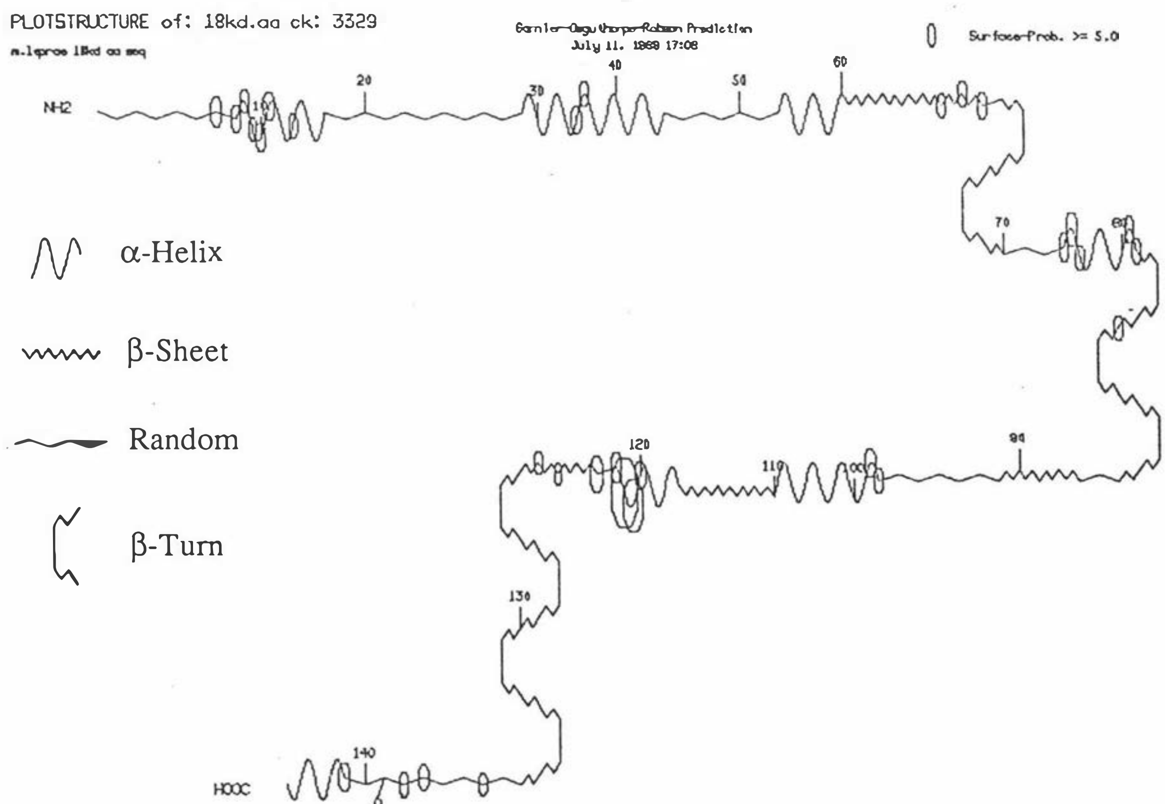


Figure2.3 13 Predicted secondary stucture plot using Garnier-Osguthorpe-Robinson prediction program. The region 46-65 indicates that residues 53-60 form an α-Helix where there is the possibility of hydrophobic interactions.

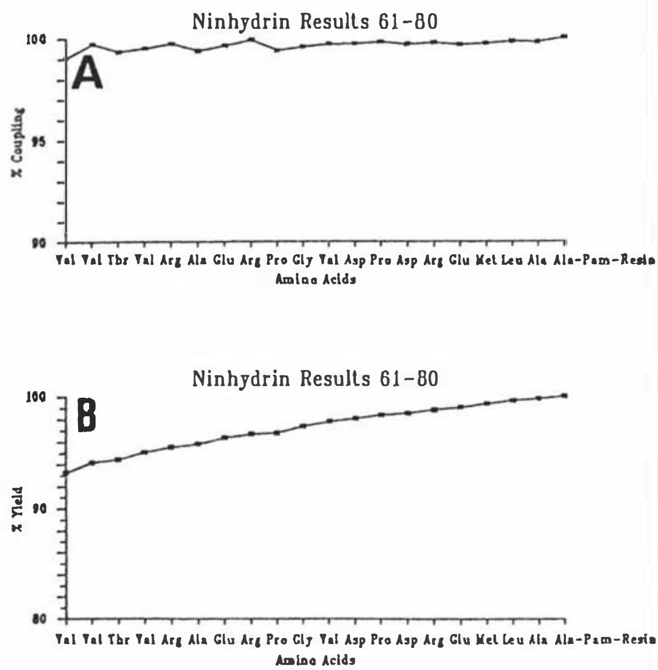


Figure 2.3.14 a) % Coupling for peptide 61-80 b) Cumulative predicted yield for the peptide on the resin

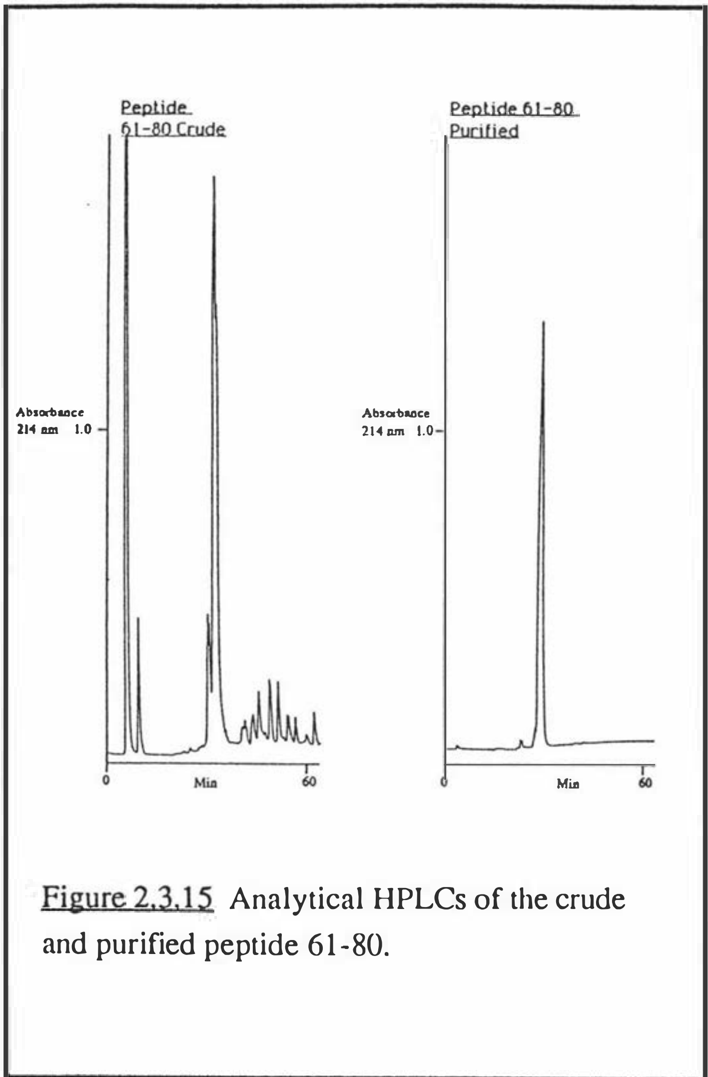


Figure 2.3.15 Analytical HPLCs of the crude and purified peptide 61-80.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide Relative to Ala
Asx	2	2.09	2.08
Thr	1	0.54	0.92
Glx	2	1.95	2.12
Pro	2	1.98	1.96
Gly	1	0.97	1.18
Ala	3	3.0	3.0
Val	4	3.63	2.48
Met	1	0.04	0.99
Leu	1	1.04	1.09
Arg	3	3.07	3.29

Table 2.3.8: AAA of peptide SLP-9 (61-80). Low val content due to Val-Val ,as discussed in section 2.1

2.3.1.2.6 SLP-9 (61-80)

Sequence:
Val Val Thr Val Arg Ala Glu Arg Pro Gly Val Asp Pro Asp Arg Glu Met Leu Ala Ala

At the completion of the synthesis the dry weight of the peptide-resin was 1.54g an increase of (1.0 g). The average coupling yield per amino acid was 99.63% giving an expected yield for the peptide of 93.28% from the ninhydrin analysis (Figure 2.3.14). AAA of the peptide-resin was consistent for the amino acid composition of the peptide (Table 2.3.8). After HF cleavage the weight of the crude lyophilised peptide was 299 mg (89.7%). Preparatve HPLC on 90 mg of the peptide gave 50 mg (55%) of peptide of greater than 95% purity as judged by analytical HPLC (Figure 2.3.15) and AAA (Table 2.3.8).

2.3.1.2.7 SLP-10 (76-95).

Sequence:
Glu Met Leu Ala Ala Glu Arg Pro Arg Gly Val Phe Asn Arg Gln Leu Val Leu Gly Glu

At the completion of the synthesis the dry weight of the peptide-resin was 1.40 g a weight increase of 0.8 g slightly under the expected amount of 0.9 g. The ninhydrin assay indicated an average coupling of 99.74% with an overall yield for the peptide of 95.11% (Figure 2.3.16). AAA of the peptide on the resin showed the amino acid composition for the peptide was consistent (Table 2.3.9). After HF cleavage 277 mg (91% Yield) of crude lyophilised peptide was isolated. Preparative HPLC on 47.5 mg of crude peptide gave 22.4 mg (47%) of peptide with greater than 98% purity as judged by AAA (Table 2.3.9) and analytical HPLC (Figure 2.3.17).

2.3.1.2.8 SLP-11 (91-110)

Sequence:
Leu Val Leu Gly Glu Asn Leu Asp Thr Glu Arg Ile Leu Ala Ser Tyr Gln Glu Gly Val

At the completion of the synthesis the dry weight of the peptide-resin was 1.36 g an increase of 0.84 g. The average coupling yield per amino acid was 98.4% with an expected yield of peptide of 72.28% (Figure 2.3.18). The decreased yield was due to two incomplete couplings at Leu^{93} and Leu^{103} of 95 and 81% respectively. After examination of the AAA of the peptide-resin, the amino acid ratios showed that a

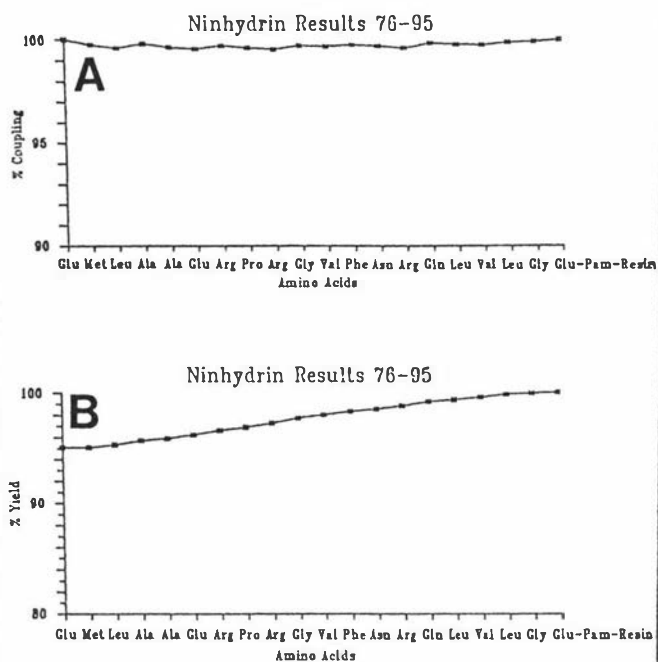


Figure 2.3.16 a) % Coupling for peptide 76-95 b) Cumulative predicted yield for the peptide on the resin

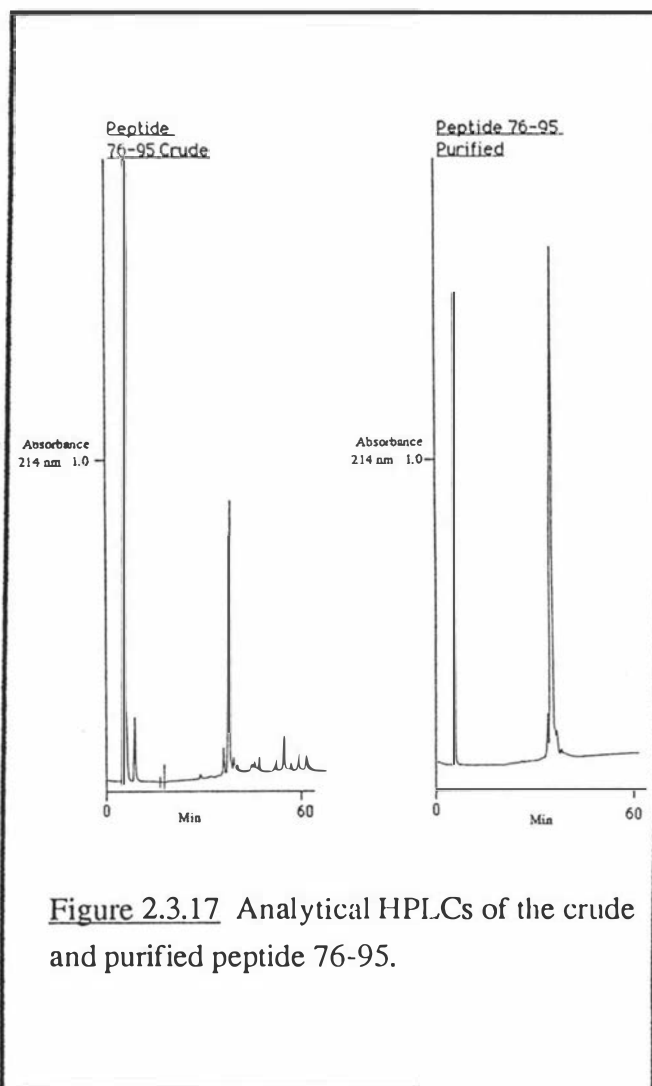


Figure 2.3.17 Analytical HPLCs of the crude and purified peptide 76-95.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide Relative to Ala
Asx	1	1.07	1.07
Glx	4	4.27	4.08
Pro	1	1.09	1.00
Gly	2	2.17	2.09
Ala	2	2.0	2.00
Val	2	2.11	1.99
Met	1	0.39	1.01
Leu	3	3.30	3.20
Phe	1	1.03	0.99
Arg	3	3.32	3.42

Table 2.3.9: AAA of peptide SLP-10 (76-95).

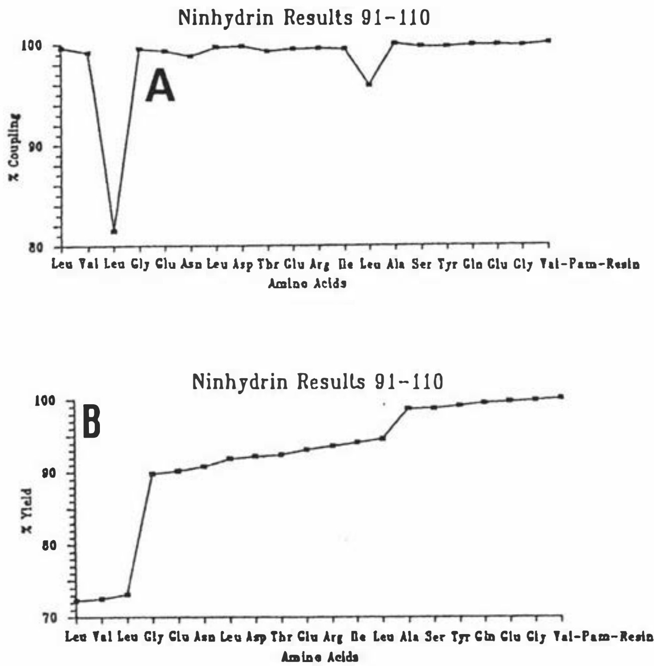


Figure 2.3.18 a) % Coupling yield for peptide 91-110 with a failed coupling at position 18.
b) Predicted cumulative yield for the peptide on the resin.

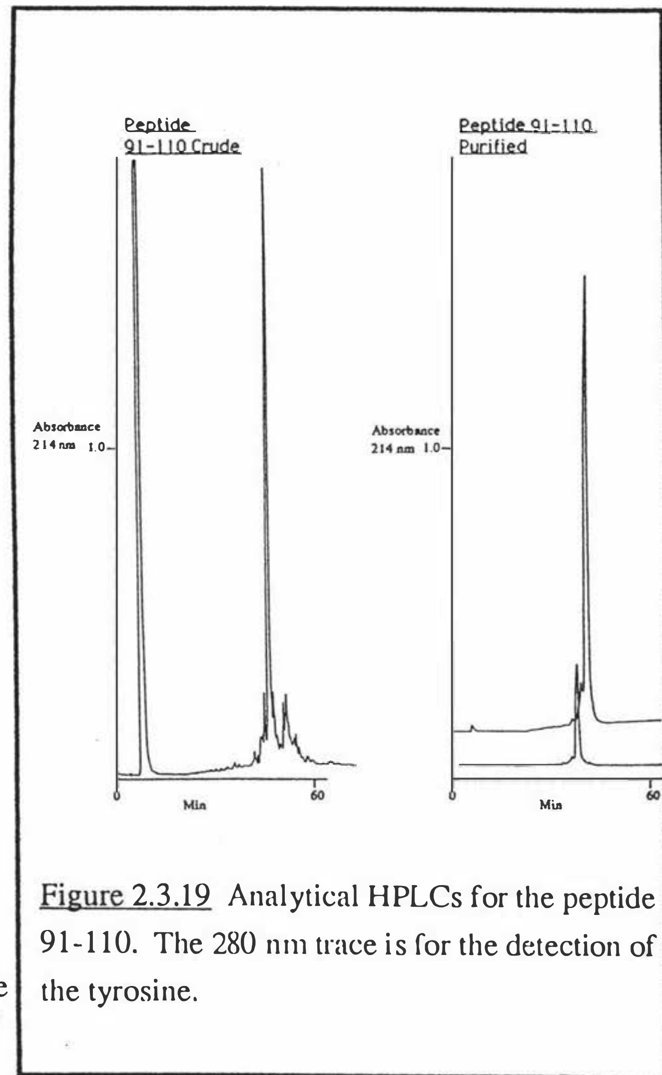


Figure 2.3.19 Analytical HPLCs for the peptide 91-110. The 280 nm trace is for the detection of the tyrosine.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide relative to Ala
Asx	2	1.93	1.97
Thr	1	0.94	0.91
Ser	1	0.54	0.97
Glx	4	3.99	3.97
Gly	2	1.97	1.87
Ala	1	1.00	1.00
Val	2	1.93	1.73
Ile	1	0.89	0.64
Leu	4	3.77	3.31
Tyr	1	0.59	0.99
Arg	1	1.02	0.89

Table 2.3.10: AAA of peptide SLP-11 (91-110). Val, Ile and Leu all low due to Leu-Val-Leu and Ile-Leu. If referenced to Gly ratios increase.

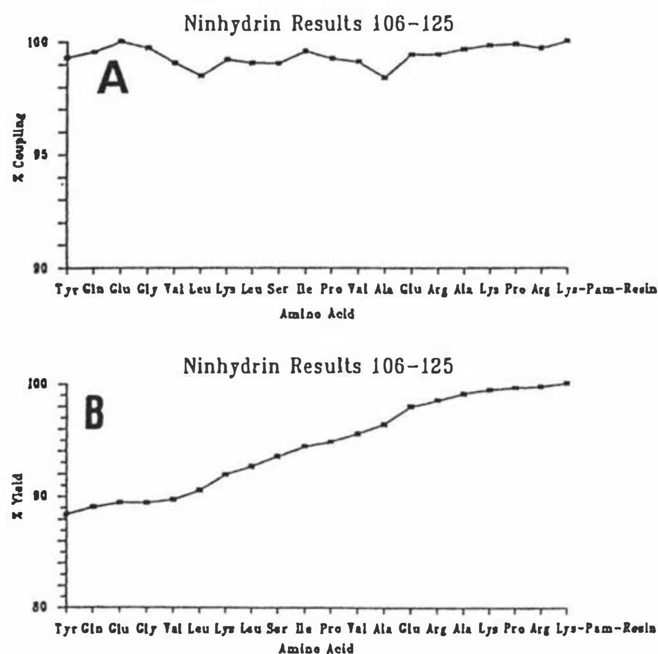


Figure 2.3 20 a) % Coupling yield for peptide 106-125.
b) Predicted cumulative yield for the peptide on the resin.

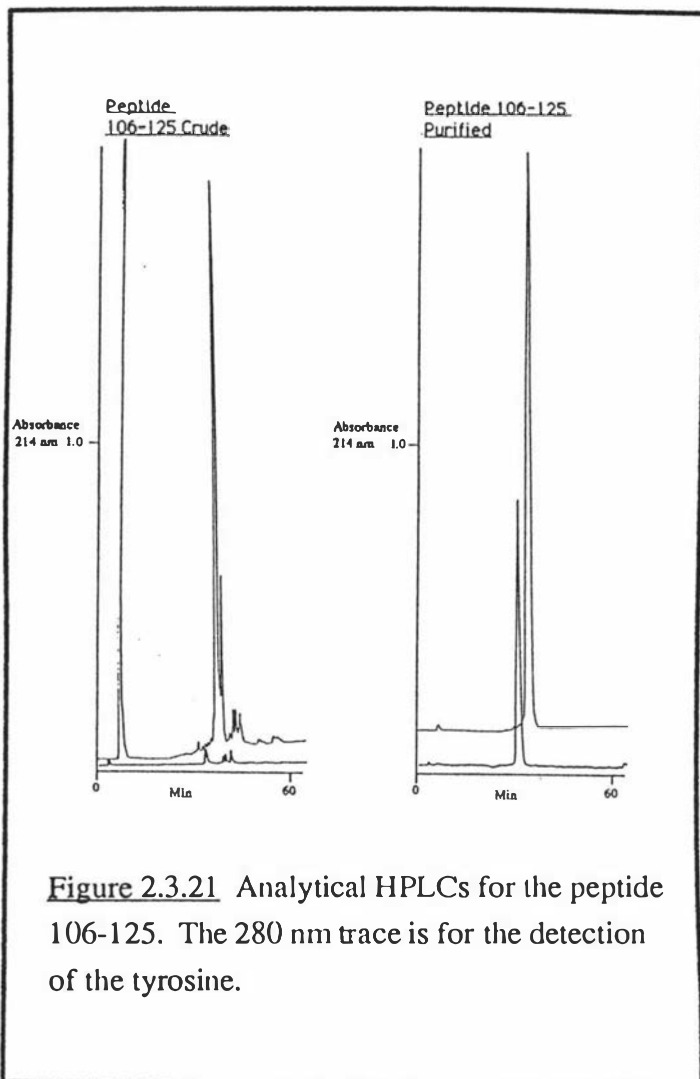


Figure 2.3.21 Analytical HPLCs for the peptide 106-125. The 280 nm trace is for the detection of the tyrosine.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide Relative to Ala
Ser	1	0.85	0.97
Glx	3	2.95	3.04
Pro	2	1.62	1.95
Gly	1	1.13	1.10
Ala	2	2.00	2.00
Val	2	1.97	1.91
Ile	1	0.93	1.00
Leu	2	2.14	2.11
Tyr	1	0.95	1.03
Lys	3	3.43	3.02
Arg	2	2.62	2.24

Table 2.3.11: AAA of peptide SLP-12 (106-125).

reasonable amount of the completed peptide was on the resin and a resynthesis was not warranted (Table 2.3.10). After HF cleavage 239 mg (77.0%) of crude peptide was isolated. Preparative HPLC on 98.7 mg of the crude peptide gave 27 mg of pure peptide with a purity greater than 90% as judged by AAA (Table 2.3.10) and analytical HPLC (Figure 2.3.19).

2.3.1.2.9 SLP-12 (106-125)

Sequence:
Tyr Gln Glu Gly Val Leu Lys Leu Ser Ile Pro Val Ala Glu Arg Ala Lys Pro Arg Lys

After completion of the synthesis the dried peptide-resin weight was 1.76 g a weight increase of 1.13 g. The ninhydrin assay indicated an average coupling yield of 99.36%, with an overall yield for the peptide of 88.45% (Figure 2.3.20). AAA of the peptide resin was in accord with the amino acid composition of the peptide (Table 2.3.11). After HF cleavage of the peptide 214 mg (66% yield) of crude peptide was isolated. Preparatve HPLC on 63 mg of the crude peptide gave 33mg (52% yield) of lyophilised pure peptide. AAA (Table 2.3.11) and analytical HPLC (Figure 2.3.21) showed that the peptide was greater the 97% pure. This peptide was also sequenced as a final confirmation of the peptide structure, because the antibody assays were not as predicted (see section 2.3.3).

2.3.1.2.10 SLP-13 (121-140)

Sequence:
Ala Lys Pro Arg Lys Ile Ser Val Asp Arg Gly Asn Asn Gly His Gln Thr Ile Asn Lys

After completion of the synthesis the dried peptide resin weighed 1.45 g, a weight increase of 0.825 g. The average coupling yield per amino acid was 99.24% with an expected yield for the peptide of 86.49% as judged by ninhydrin analysis (Figure 2.3.22). The slightly decreased yield could be due to the colour associated with the DNP group on the histidine interfering with the readings during the analysis. AAA of the peptide-resin showed the amino acids to be in the correct ratios (Table 2.3.12), except for histidine as the DNP protecting group on the histidine was not removed by acid hydrolysis. During the removal of the DNP group it was noted that the solution went a dark red colour for a period of time instead of the normal pale yellow colour. After removal of the Boc group and HF cleavage 293 mg (greater than 100% yield) of crude peptide was isolated, which had a bright yellow colour associated with it.

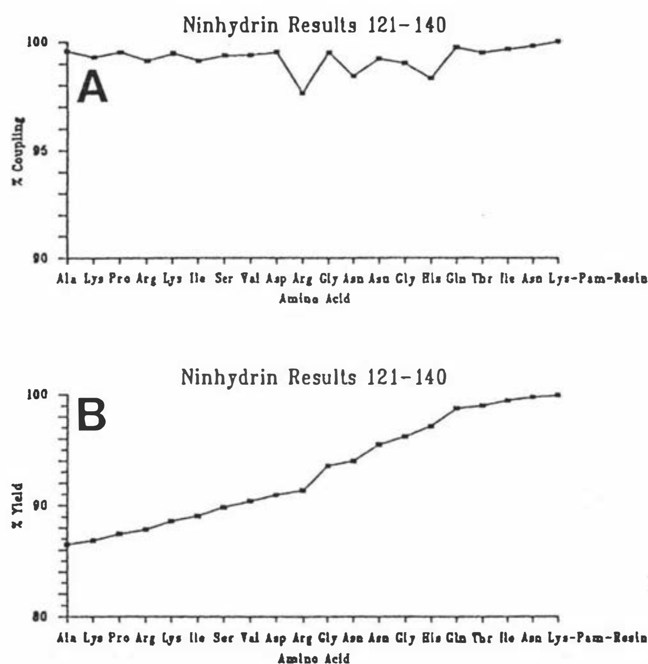


Figure 2.3 22 a) %coupling yield for peptide 121-140

b) Predicted cumulative yield for the peptide on the resin.

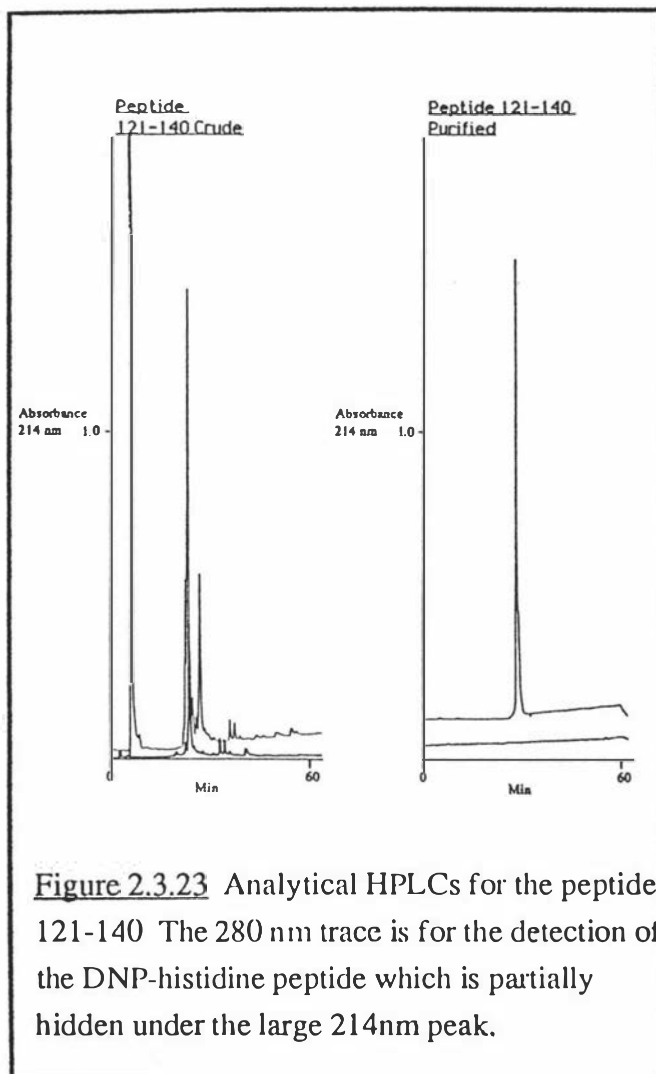


Figure 2.3.23 Analytical HPLCs for the peptide 121-140 The 280 nm trace is for the detection of the DNP-histidine peptide which is partially hidden under the large 214nm peak.

Amino Acid	Expect Ratio	Peptide Resin Relative to Gly*	Purified Peptide Relative to Gly*
Asx	4	3.07	3.64
Thr	1	1.06	0.97
Ser	1	0.66	0.99
Glx	1	1.07	1.09
Pro	1	0.48	0.74
Gly	2	2.0	2.0
Ala	1	0.72	0.75
Val	1	0.98	1.03
Ile	2	1.69	1.86
His	1	ND	1.04
Lys	3	2.79	2.44
Arg	2	1.78	1.67

Table 2.3.12: AAA of peptide SLP-13 (121-140). In the resin hydrolysis His exists as His(DNP) *NB low value for Ala, Lys, Arg and Pro (see text) Therefore the molar ratios are referenced to Gly.

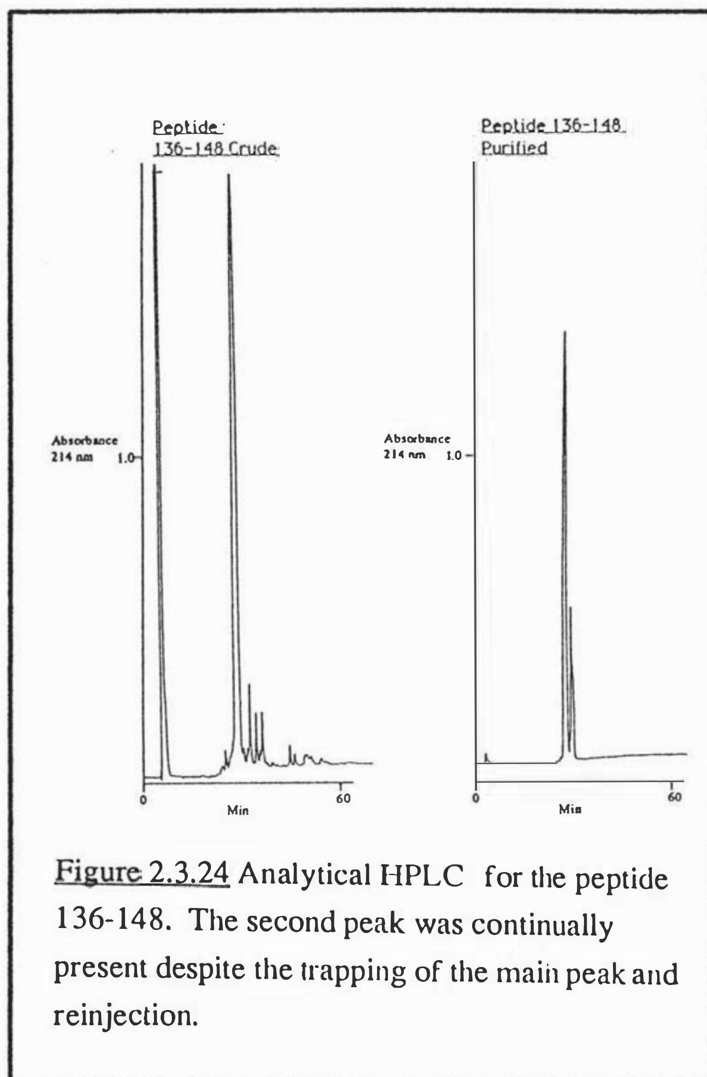


Figure 2.3.24 Analytical HPLC for the peptide 136-148. The second peak was continually present despite the trapping of the main peak and reinjection.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide Relative to Ala
Asx	2	1.89	1.91
Thr	2	1.56	1.80
Glx	2	1.88	2.00
Ala	2	2.0	2.00
Ile	3	2.20	2.04
His	1	ND	0.99
Lys	1	1.09	0.98

Table 2.3.13: AAA of peptide SLP-14 (136-148). Ile-Ile bond not hyrolysed.For the resin hydrolysis His exits as His(DNP)

Initially preparative HPLC on 60 mg of crude peptide initially gave a very small amount of peptide (14 mg, 23%) which still had a faint yellow colour associated with it. In comparison, the bulk of the material isolated had a bright yellow DNP colour. Attempts to separate the DNP complex from the peptide in solution failed. Recleavage of the peptide using thiophenol to remove the DNP group prior to the HF gave a similar result to that previously obtained with β-mercaptoethanol. Finally by dissolving the crude peptide (50 mg) in 2M guanidine .HCl and loading onto the preparative HPLC, 14.3 mg (28.6%) of purified peptide was isolated as judged by AAA (Table 2.3.12) and analytical HPLC (Figure 2.3.23). The purified peptide was able to be isolated by taking very small fractions from the leading edge of the peptide peak which shows the complete absence of any DNP associated material.

The AAA did however show that the N-terminal end of the peptide was consistently 0.3 mole ratio lower than expected over the last four amino acids over all fractions isolated across the peptide peak. There was no sign, from the ninhydrin results of any bad coupling in this area.

2.3.1.2.11 SLP-14 (136-148)

Sequence: Gln Thr Ile Asn Lys Thr Ala His Glu Ile Ile Asp Ala .

At the completion of the synthesis the dried peptide-resin weight was 1.62 g an increase of 0.96 g. AAA of the peptide-resin showed that the amino acids had the correct ratios except for the His value (as above) . The average coupling yield for the amino acids determined by ninhydrin analysis was 97.92%, with five missing samples due to a resin sampler malfunction. This along with the interference of the DNP group on the histidine could contribute to the low coupling. AAA (Table 2.3.13) of the peptide-resin was in accord with the amino acid composition of the peptide. No problems were encountered in the removal of the DNP group, with this peptide, as with the previous peptide. After HF cleavage 214 mg (71.4%) of the lyophilised crude peptide was isolated. Preparative HPLC on 51.1 mg of crude peptide gave 10.1 mg (20.1%) of pure peptide. The analytical HPLC showed a peak that could not be separated on the preparative column (Figure 2.3.24). AAA of the purified material showed it to be greater than 95% (Table 2.3.13).

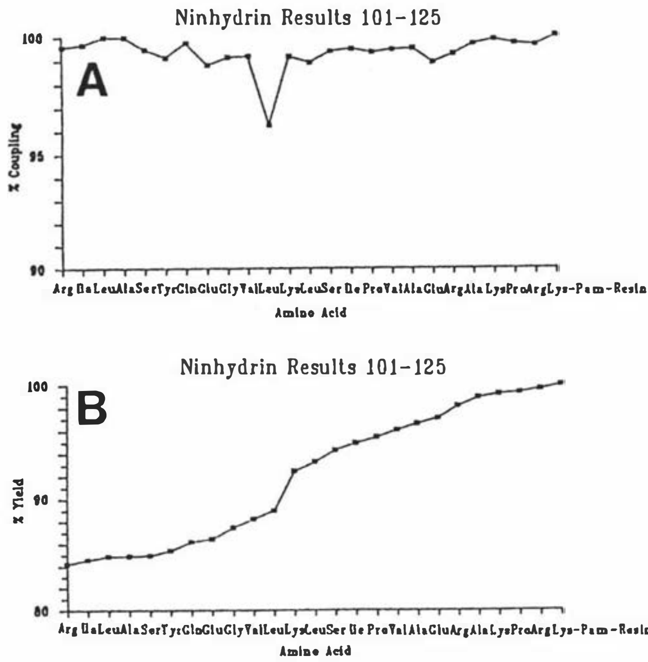


Figure 2.3 25 a) %coupling yield for peptide 101 -125. NB the failed coupling at position 15.
b) Predicted cumulative yield for the peptide on the resin.

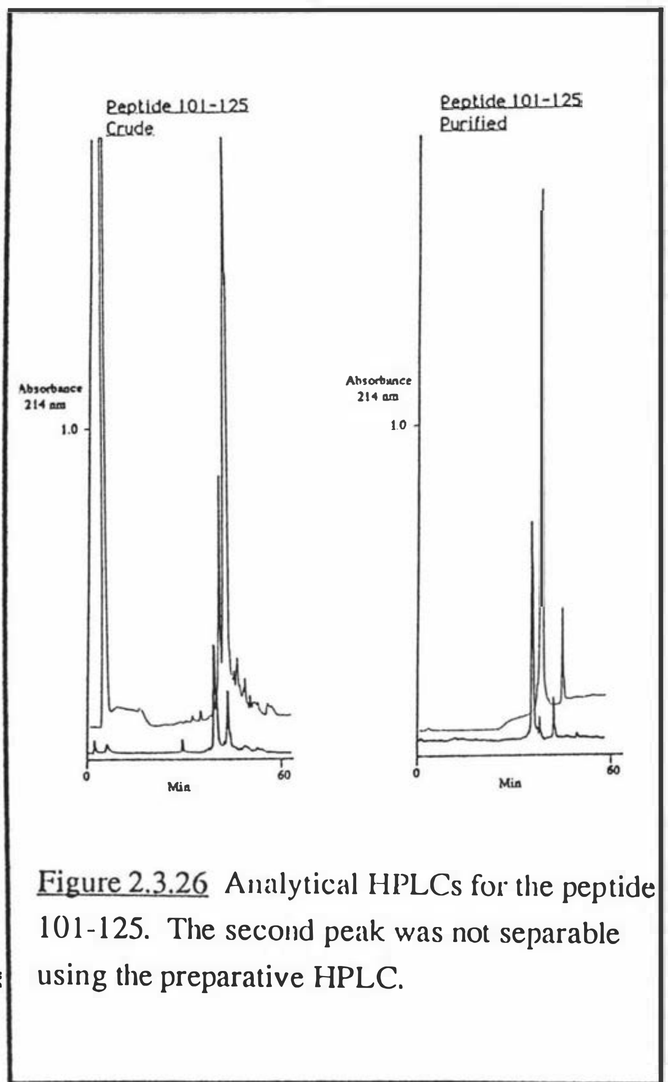


Figure 2.3.26 Analytical HPLCs for the peptide 101-125. The second peak was not separable using the preparative HPLC.

Amino Acid	Expect Ratio	Peptide Resin Relative to Ala	Purified Peptide Relative to Ala
Ser	2	1.18	1.88
Glx	3	2.79	3.07
Pro	2	2.26	1.93
Gly	1	0.83	1.11
Ala	3	3.00	3.00
Val	2	1.91	1.94
Ile	2	1.69	1.79
Leu	3	2.43	2.89
Tyr	1	0.63	0.94
Lys	3	3.11	3.06
Arg	3	3.10	3.00

Table 2.3.14: AAA of peptide SLP-15 (101-125).

2.3.1.2.12 SLP-15 (101-125)

Sequence
Arg Ile Leu Ala Ser Tyr Gln Glu Gly Val Leu Lys Leu Ser Ile Pro Val Ala Glu Arg Ala Lys Pro Arg Lys

After completion of the synthesis the weight of the peptide resin was 2.17 g, a weight increase of 1.39 g. The average coupling percentage determined by ninhydrin analysis was 99.30% with the overall peptide yield of 84.2% (Figure 2.3.25). The one coupling failure of 96% could be attributed to the weight of the resin sample being below the threshold level for detection of NH_2. AAA of the peptide resin showed the amino acid ratios to be in accord with the amino acid composition of the peptide (Table 2.3.14). After HF cleavage of 302 mg of peptide resin, 131.6 mg (67%) of crude peptide was isolated. Preparative HPLC on 47.5 mg of material gave 19.6 mg (41%) of purified peptide as determined by AAA (Table 2.3.14) and analytical HPLC (Figure 2.3.26) which showed the purified peptide to be greater than 94 % pure.

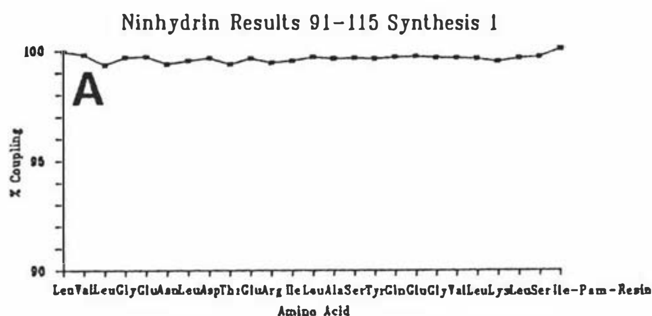
2.3.1.2.13 SLP-16 (91-115)

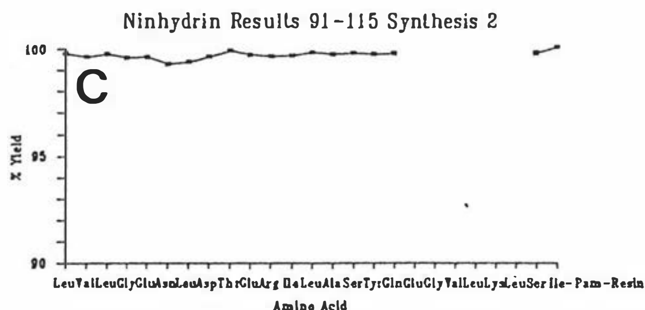
Sequence:
Leu Val Leu Gly Glu Asn Leu Asp Thr Glu Arg Ile Leu Ala Ser Tyr Gln Glu Gly Val Leu Lys Leu Ser Ile

This peptide warranted two syntheses due to the problems encountered in the initial synthesis with Kent's protocols. The problem of an incomplete coupling step meant that the second synthesis using ABI's NMP/HOBt would provide an interesting comparison between the two synthesis protocols.

Synthesis #1:

This synthesis was performed, as with previous peptides using Kent's protocols. The weight of the peptide-resin after completion of the synthesis was 1.84 g an increase of 1.15 g, slightly down on the expected of 1.38 g. The ninhydrin analysis, apart from missing two of the first four samples, gave an average coupling for each amino acid of 99.60% with a predicted yield of the peptide of 90.79% (Figure 2.3.27). There was no sign of any failed coupling. AAA on the peptide-resin showed the presence of large amount of prematurely terminated peptide (Table 2.3.15). After HF cleavage of 502 mg of peptide-resin, 230.8 mg (67.6%) of crude peptide was isolated. Analytical HPLC (Figure 2.3.28) of this crude material showed the presence of two





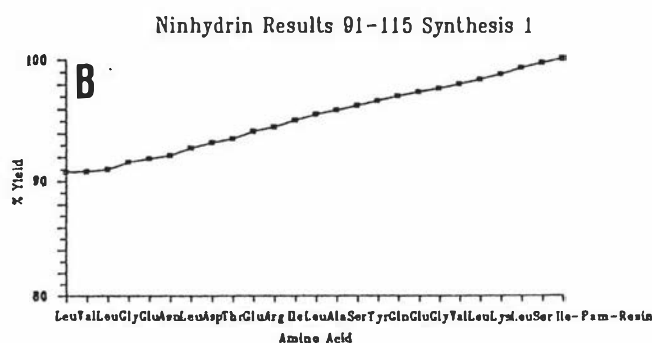


Figure 2.3.27 a) % Coupling yield for the first synthesis of peptide 91-115. There is no sign of the failed coupling at the Arg Ile coupling. b) Predicted cumulative yield for the first synthesis for the peptide on the resin. c) % Coupling yield for the second synthesis using NMP/HOBt chemistry with the eight missing values. No sign of a failed coupling at the Arg Ile.

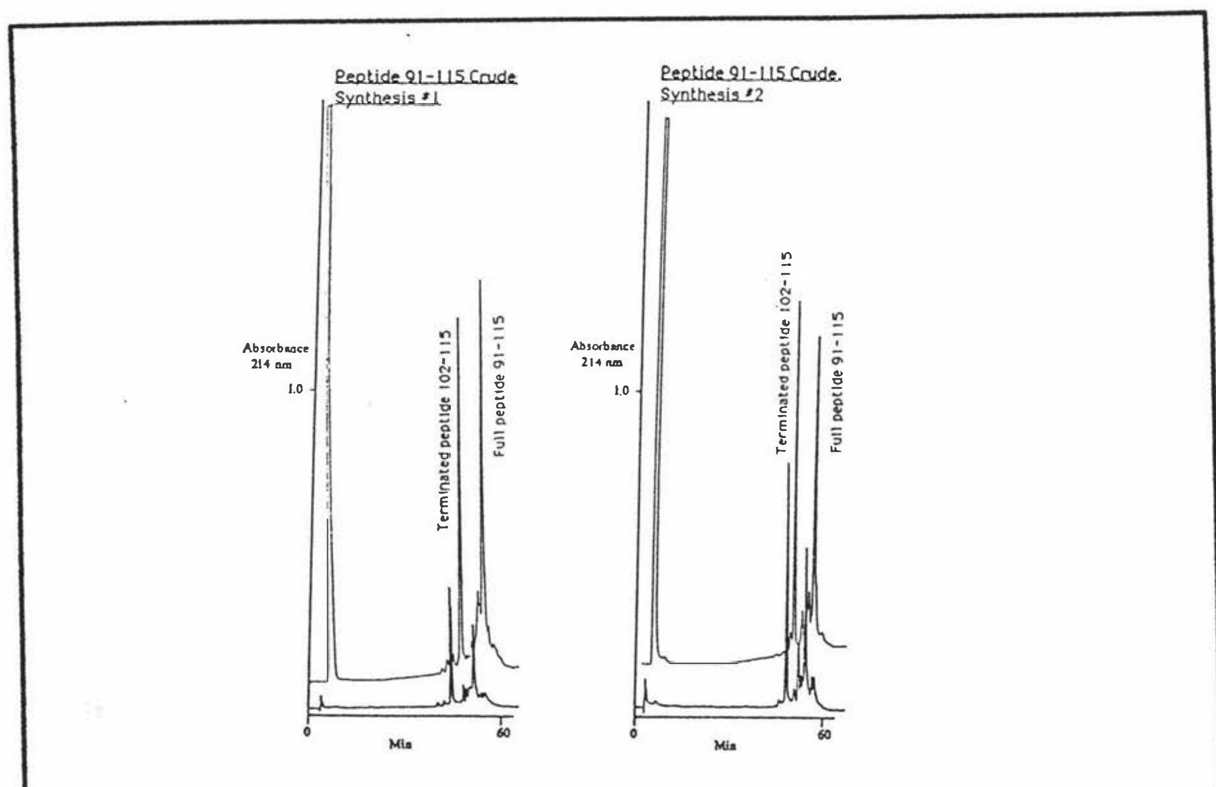


Figure 2.3.28 Analytical HPLC traces for peptide 91-115. a) Crude trace for synthesis 1 with Kents protocols. b) Crude peptide for synthesis 2 with NMP/HOBt chemistry. NB the difference in the ratios of the peak heights between the terminated peptide and the full length peptide for the two traces.

major peaks, one which had a retention time of 46.6 min comprising 46.6% of the total integrated area and the other with a retention time of 53.1 min comprising 53% of the total integrated area. After preparative HPLC on 50.7 mg and isolation of the two peaks, AAA were carried out which shows that the earlier peak consisted of a 14 amino acid long peptide from 102-115

102 115
Ile Leu Ala Ser Tyr Gln Glu Gly Val Leu Lys Leu Ser Ile

The second peak was the full length peptide 91-115. For some reason the Arg failed to couple fully to Ile, terminating in a way that no exposed NH_2 groups could be detected by ninhydrin analysis.

		Expected AA ratio		Synthesis 1		Synthesis 2	
Amino Acid	Peptide on Resin	Frac 1	Frac 2	Fraction 1	Fraction 2	Fraction 1	Fraction 2
Asx	1.46	0	2	0.21	2.01	0.23	1.79
Thr	1.01	0	1	Nd*	0.90	0.11	0.83
Ser	1.57	2	2	1.66	2.08	1.52	1.96
Glx	4.08	2	4	2.0	3.96	2.08	3.96
Gly	2.00	1	2	1.0	2.0	2.0	2.0
Ala	1.21	1	1	0.94	1.15	0.93	1.20
Val	2.00	1	2	1.01	1.56	1.02	1.44
Ile	2.41	2	2	1.72	1.58	1.72	2.08
Leu	6.12	3	6	3.02	4.84	3.02	5.84
Tyr	0.87	1	1	0.85	1.31	0.85	1.00
Lys	1.35	1	1	0.89	1.02	0.90	1.12
Arg	0.74	0	1	Nd*	0.76	0.15	1.0

Table 2.3.15: AAA of peptide SLP-16 (91-115).
Synthesis 1 Synthesised under Kent's protocols.Synthesis 2 Synthesised using NMP/HOBt chemistry
Fraction 1 is for the Terminated peptide Fraction 2 is for the full length peptide.
* Nd Not detected.

Synthesis #2:

For this synthesis the standard ABI NMP/HOBt synthesis protocol was used which included capping cycles that should block any unreacted NH_2. After completion of the synthesis the peptide-resin weighed 1.58g, a weight increase of 902 mg. Ninhydrin analysis minus six of the first eight samples (due to fraction collector failure) showed no sign of failed couplings in either the first coupling or the second (recoupling) step (Figure 2.3.27). HF cleavage on 500 mg of peptide-resin gave 195.9 mg of crude peptide. Analytical HPLC (Figure 2.3.28) of the crude peptide showed two major peaks with similar retention times as those in the first synthesis with the ratio between the two peaks reversed to 52% for the terminated peptide and 47% for the full length peptide. Isolation of the two separate peaks again confirmed the composition of the terminated peptide and the full length peptide. Sequencing of the crude peptide mixture confirmed the presence of the the two peptides with ratios of approximately 43% termination peptide. There was no sign that the N-terminal was blocked (ie acylated), or that a rearrangement of the Ile amino acid had prevented coupling, in the terminated peptide.

This peptide could possibly be a peptide that exhibits a specific sequence dependent coupling, since none of the peptides synthesised previously in this region of the protein indicated a problem with this coupling step. It is quite possible that coupling Arg at this position (15 th in the peptide chain) to a sterically hindered amino acid such as Ile caused the coupling failure. At this distance from the peptide resin it could be hypothesised that the peptide chain is just beginning to form some type of secondary structure, which could cause the amino terminus to become trapped thereby preventing coupling. This does not however explain the totally negative result from the ninhydrin analysis which failed to detect the presence of NH_2 under conditions that should have exposed any trapped NH_2.

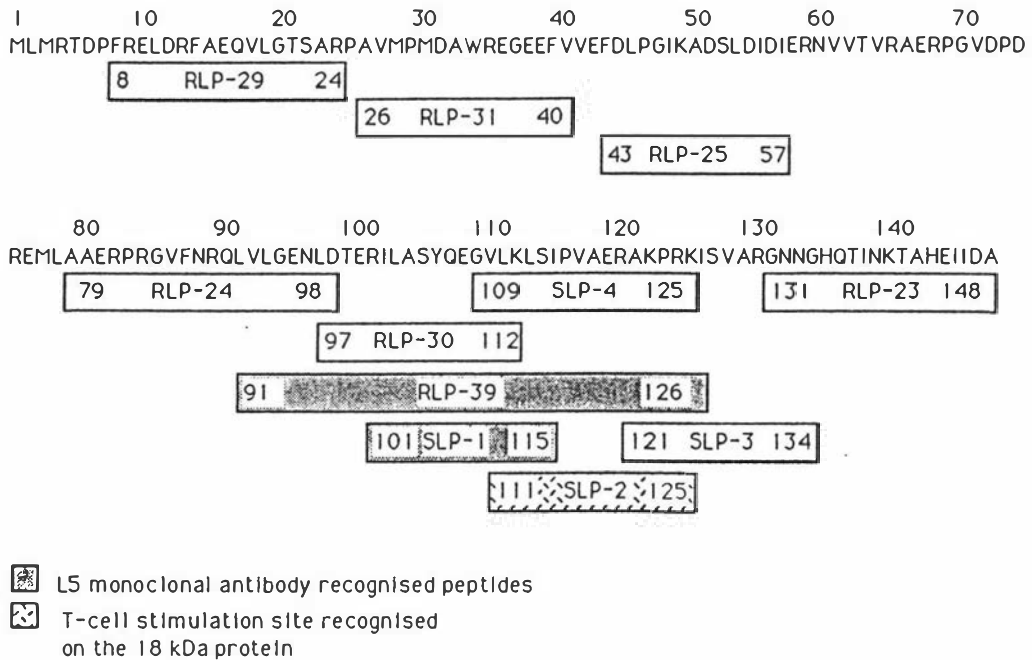


Figure 2.3.29. Protein sequence showing the peptides synthesised in the first phase of this work (RLP peptides) and the earlier SLP peptides. The peptides RLP-39 and SLP-1 are recognised by the L5 Monoclonal antibody, while only peptide SLP-2 is recognised as a T-cell stimulation epitope from the 18 kDa protein.

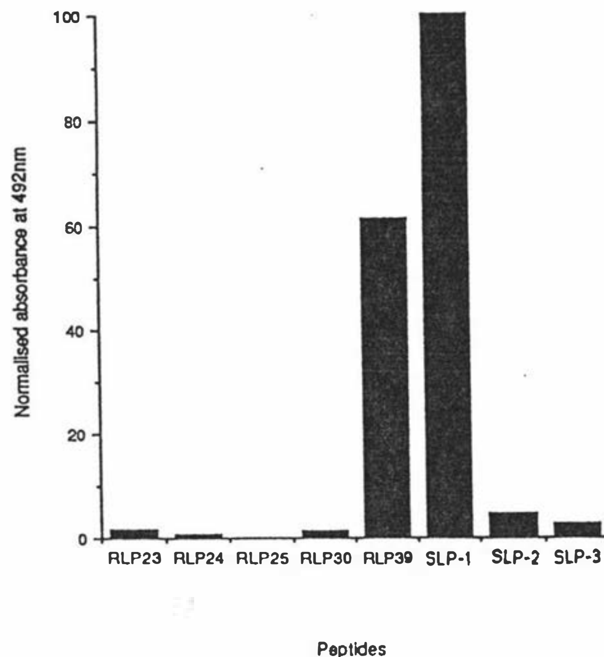


Figure 2.3.30. A comparison of L5 Binding to synthetic peptides in ELISA assays.

2.3.2 IMMUNOLOGICAL ASSAYS.

2.3.2.1 RLP PEPTIDES AND SLP-1 to SLP-4.

Figure 2.3.29 shows the first set of peptides synthesised for this project. Peptides RLP-29,31,24,30,and 23 were predicted by a the computer program of Margalit[147] to have a propensity to form amphipathic helices, a peptide structure that is reported to be good T-cell epitope stimulation regions[129]. RLP-25 was synthesised as a control as it showed no propensity in the computer program to form an amphipathic helix. From the initial screening of these peptides for possible T-cell stimulation it was found that none of these peptides stimulated T-cells.

By using a combination of recombinant DNA technology and synthetic peptides the epitope on the 18 kDa protein that is recognised by the murine monoclonal antibody L-5, which is used specifically to detect the 18kDa protein, was isolated. The peptides in Figure 2.3.29 were used as a final confirmation of the binding site, as the exonuclease III deletion clones had localised the binding region between residues 109-115. The peptides of interest were RLP-39 which duplicated the sequence of one of the clones and RLP-30 and SLP-1 to 3 which cover the regions 97-112, 100-115, 111-125, and 121-135 respectively. Only RLP-39 and SLP-1 which contain the full region 109-115 bound a significant quantity of L-5 monoclonal antibody (Figure 2.3.30)[54].

It was found later, using the series of overlapping 20 mer peptides, that amino acid residues outside this area were also required for recognition of the antibody. Peptide 106-125, which from the results discussed here would be predicted to bind the antibody failed to recognise the Mab L-5. It is now thought that residues in the region between 101-115 (ie beyond 109 as first thought), are also required for antibody recognition.

Using the series of peptides RLP-30 to SLP-3, further screening for T-cell proliferation epitopes was carried out. Mice were immunised *in vivo* with the intact 18kDa protein and the lymph node cells were subsequently challenged with the synthetic peptides. The data of Figure 2.3.31 show that the peptides RLP-23, 24,25, 29, 30, 31, and SLP-1 and SLP-3 elicited no significant response in microculture assay. Peptides RLP-39 (91-126) and SLP-2 (111-125) consistently stimulated proliferative responses in lymph node cells immunised with the 18kDa antigen. When the mice were first immunised with the peptide and the lymph node cells challenged with the 18 kDa protein, the peptide SLP-2 consistently gave the maximum response,

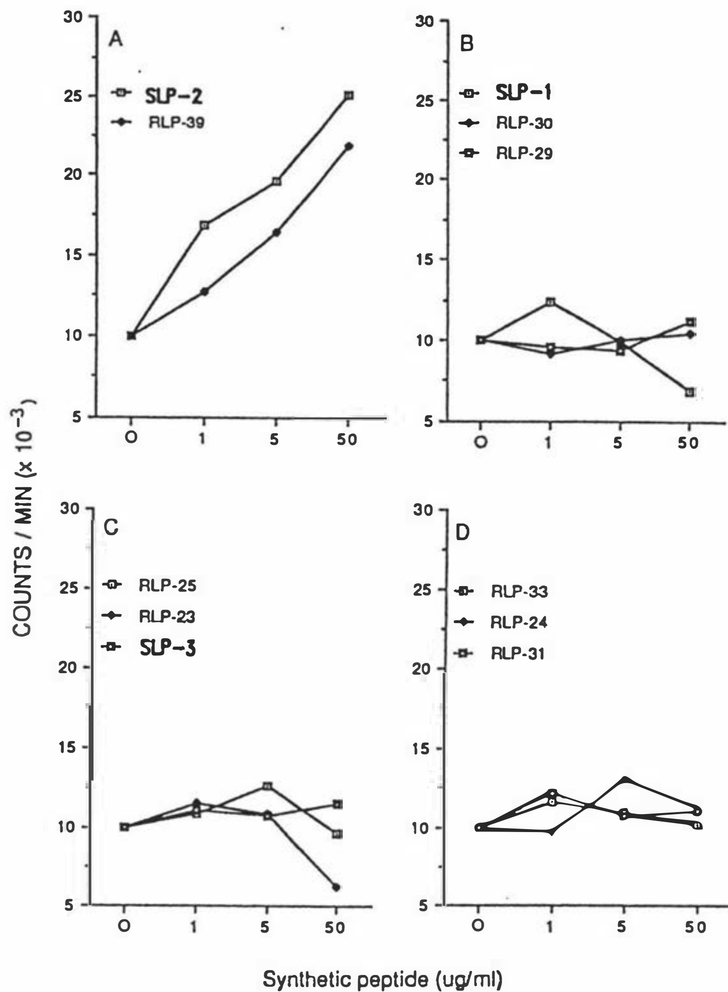


Figure 2.3.31. Proliferative response lymph node cells from mice immunised with 18 kDa protein and challenged in culture with synthetic peptides. As a control peptide RLP-33 which included the amino acids 198-219 of the 65 kDa protein from *M.leprae*. Significant stimulation (>50000 cpm) was always observed when lymph node cells were challenged in culture with 18 kDa protein.

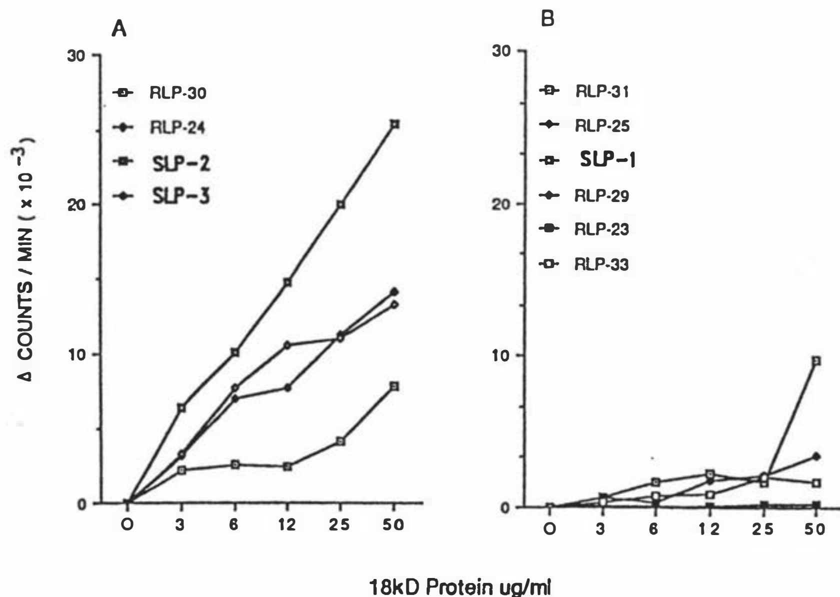


Figure 2.3.32. Proliferative response of lymph node cells from mice immunised with synthetic peptide then challenged in culture with 18 kDa protein.

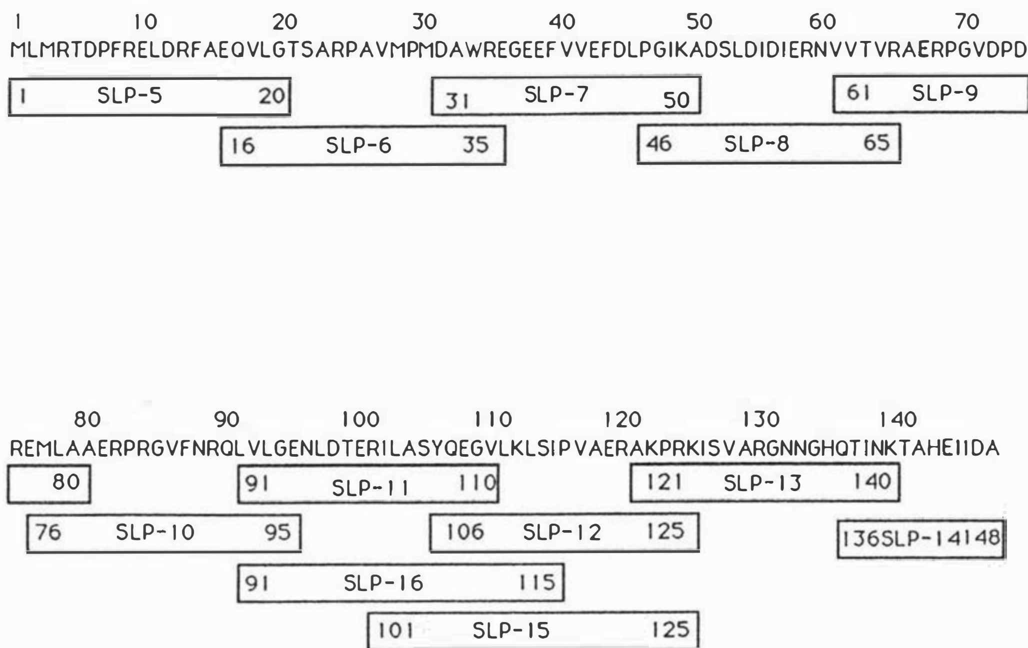


Figure 2.3.33. Protein sequence showing the overlapping 20 amino acid peptides and the two 25 amino acid peptides.

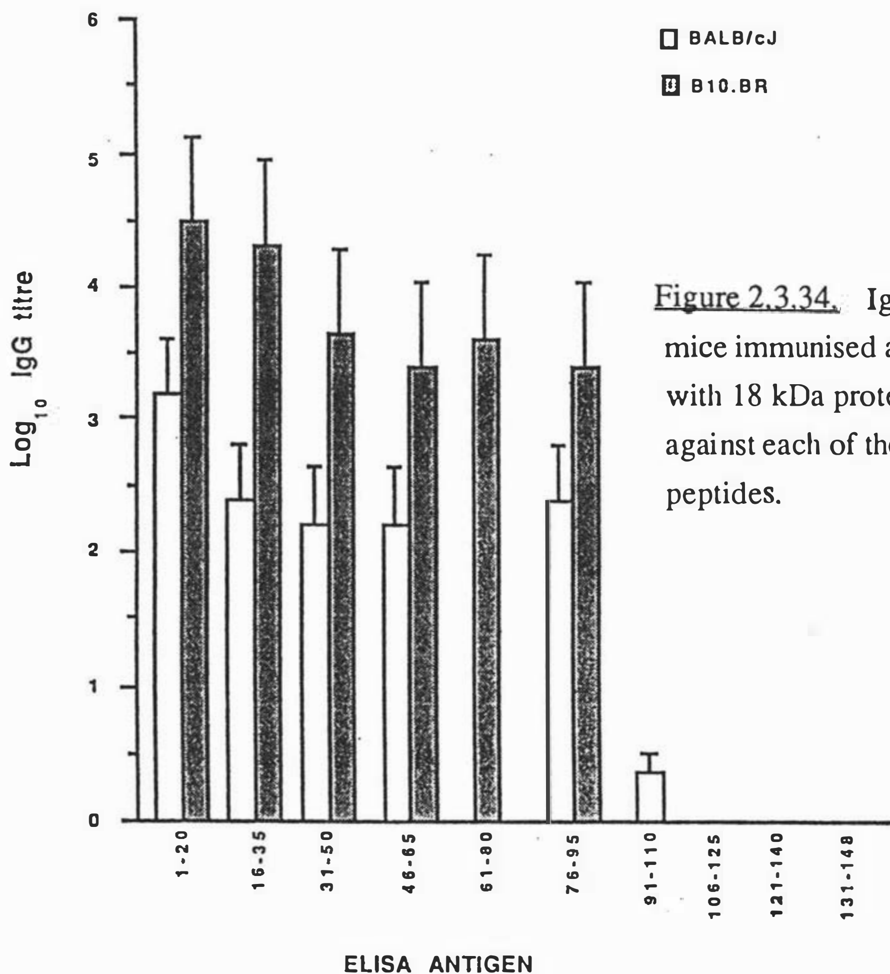


Figure 2.3.34. IgG from BALB/cJ and B10.BR mice immunised and subsequently challenged with 18 kDa protein. Sera were assayed for IgG against each of the overlapping synthetic peptides.

but the peptides RLP-30,RLP-24 and SLP-3 gave lower responses (Figure 2.3.32), adding further evidence that the peptide SLP-2 (111-125) contained an important T-cell epitope.

It is notable that SLP-2 does not contain part of any of the predicted motifs from the computer predictions, however from the surface probability predictions (Figure 2.3.13) the region 120-125 has a high probability of being on the outer surface of the protein and therefore more likely to be recognised by immunological recognition systems. It is also probable that additional recognition areas would be required for optimal binding to Major Histocompatability Complex encoded products such as antigen presenting cells, or for T-cell recognition and binding.

2.3.2.2 PEPTIDES SLP-5 to SLP-16

In order to have complete coverage of the protein the 20 mer overlapping peptides SLP-5 to SLP-14 (ie 1-20, 16-35, 31-50,...121-140, 136-148) were synthesised (Figure 2.3.33). By isolating the regions of the protein that contain IgG epitopes it should be possible to systematically identify peptides that contain B-cell epitopes and to study the number and location of the associated T_H cell epitopes which are required for the induction of IgG responses. Previously only one T-cell epitope (peptide region SLP-2) was identified on the recombinant 18 kDa protein However, T_H cell epitopes can exist either adjacent or very close to B-cell epitopes therefore rendering a peptide capable of inducing IgG responses on its own. In this manner the identification of these T-cell epitopes should allow the engineering of dominant epitopes as synthetic polypeptides with a view to developing a vaccine.

By immunising different strains of mice with the 18kDa protein it was possible to identify the types of responses as low, medium or high. It was clear from these results that the genes encoded in the H-2 complex influenced the IgG responses and that the gene responsible for the high responses was inherited in a dominant manner[130]. Groups of the high responding breeds of mice B10.BR and BALB/cJ were immunised then challenged with the 18kDa protein. The sera were isolated six days after the third challenge and assayed for binding of each 20 mer peptide (Figure 2.3.34) High IgG responses were found for peptides 1-20, 16-35, 31-50, 46-65, 61-80, and 76-95 in BALB/cJ and B10.BR mice. Differences occurred in IgG responses to peptide 61-80 which gave a high response in B10.BR mice but no response in BALB/cJ mice while for peptide 91-110 a low response was found in BALB/cJ mice but no response in B10.BR mice. It was also observed that for peptides 1-20,16-35, 31-50 and 76-95

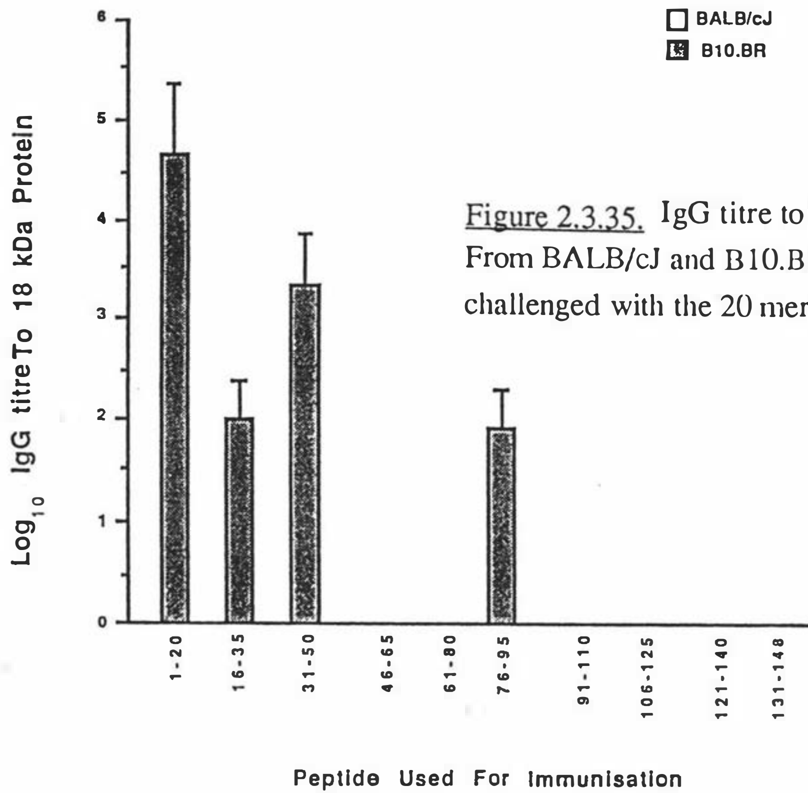


Figure 2.3.35. IgG titre to the 18kDa protein From BALB/cJ and B10.BR mice immunised and challenged with the 20 mer synthetic peptides.

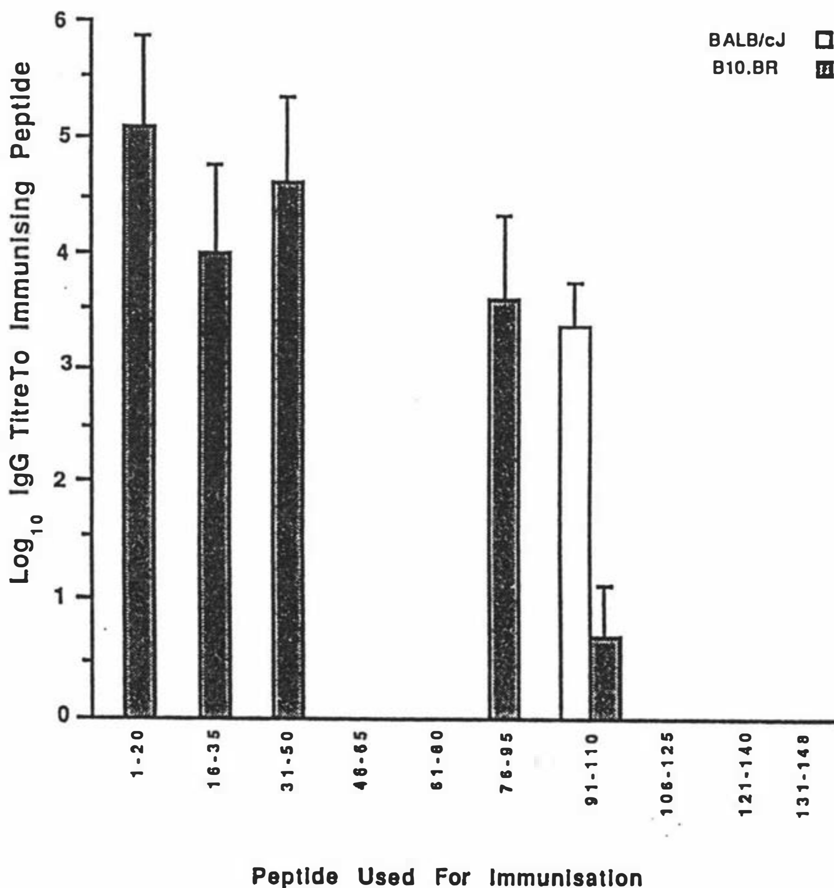


Figure 2.3.36. IgG titre to the immunising peptide from BALB/cJ and B10.BR mice immunised and with the synthetic peptides. Sera were assayed for IgG that bound the immunising synthetic peptide.

there were significant, but varied IgG titre, responses in low, intermediate and high responding breeds of mice.

If peptides contain both B-cell epitopes and adjacent T_H epitopes then, using the peptides as antigens, the peptides themselves should be able to produce IgG responses. Immunisation of groups of B10.BR and BALB/cJ mice with the peptides (three times with 10 µg of each 20-mer) and analysing their sera, by ELISA for IgG that bound either the 18kDa protein (Figure 2.3.35), or the immunising peptide (Figure 2.3.36), showed that there were significant differences between the B10.Br mice and BALB/cJ mice. The B10.BR mice had high titres to peptides 1-20, 16-35, 31-50 and 76-95. Only 91-110 elicited IgG response to the peptide and not the intact 18kDa protein in both B10.Br and BALB/cJ mice. From this it can be concluded that there are B-cell and T_H cell epitopes seen by B10.BR mice and not the BALB/cJ mice for peptides 1-20, 16-35, 31-50, 76-95 and 91-110. The fact that peptide 91-110 elicits an IgG response to the peptide and not the protein indicates that the T_H epitope on this peptide is recognised by both strains of mice but the B-cell epitope differs from that recognised by the 18kDa protein.

As both strains of mice responded to immunisation with the 18kDa protein by producing IgG that binds peptides 1-20, 16-35, 31-50 and 76-95, both strains clearly recognise the B-cell epitopes of these peptides. The differences in IgG recognition when the peptides were used as immunising antigens suggest that the T_H epitopes seen by the different breeds of mice were different. This clearly does not imply that the T_H epitopes are the same as those revealed by *in vivo* processing of the 18kDa protein. There are two possible reasons for the different responses. When the protein is used to immunise H-2^d mice, T_H epitopes located elsewhere on the protein could be processed to enable activation of the B-cells with specificities to these peptides.

The other alternative is that different subpopulations of T_H cells either T_H1 or T_H2, are being activated[148,149]. The four peptides 1-20, 16-35, 31-50 and 76-95 may activate T_H1 cells in BALB/cJ mice but not in B10.BR mice. There could also be a difference in the binding affinities of the peptides where there is much stronger binding in H-2^d than to H-2^k Class II MHC proteins. The results tend to suggest that the induced responses are caused by the activation of different subpopulations of T-helper cells, due to the recognition of different T-cell epitopes.

Another point to note is the lack of IgG response for peptides 91-110,106-125, 121-140 and 136-148, while immunisation with 18kDa protein resulted in IgG that bound peptide 101-148 (see Chapter 3)[130]. This implies that the B-cell epitopes

detected by binding to the 50 mer were of a conformational nature and lost in the smaller 20-mer peptides. This is supported by the fact that the 25-mers 91-115 and 101-125 (ie SLP-15 and SLP-16), which should be able to form secondary strucure both bound to the sera derived from immunisation of the B10.BR mice (data not shown)[130].

In an effort to correlate the T-cell responses, T-cell recognition assays were carried out using the high responder strains of mice BALB/cJ and B10.BR, by immunising with the overlapping 20-mer peptides as antigen *in vivo* then challenging the lymph node cells *in vitro* with the 18kDa protein. Groups of mice were immunised separately with 10μg of synthetic peptide in FIA in the base of the tail and challenged 10 and 19 days later with 10μg of the same peptide in saline. Ten days after the final immunisation lymph nodes were excised, cell suspensions were prepared and challenged in vitro with 2, 10 and 50 μg/ml of 18kDa protein.

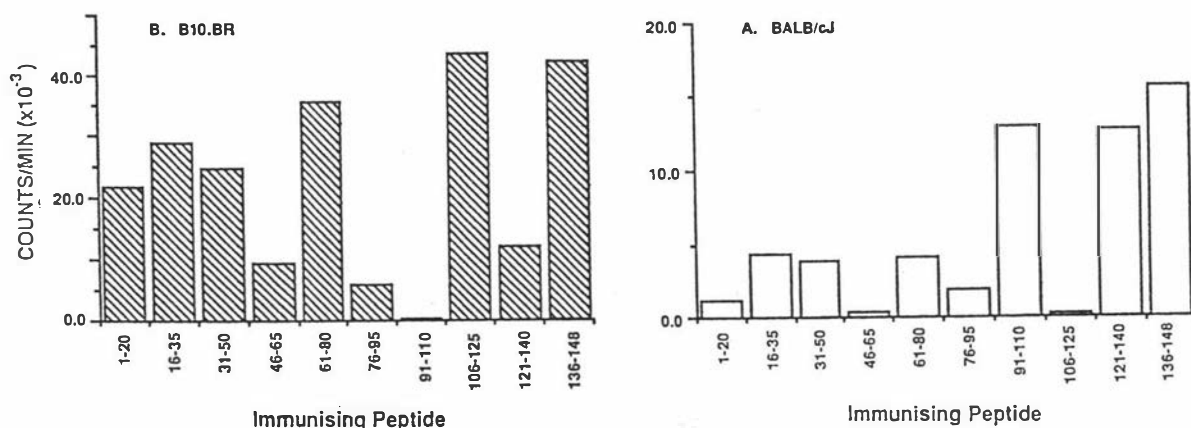


Figure 2.3.37. Proliferative T-cell responses to lymph node cells from mice immunised with 20 mer peptides and challenged in vitro with 18 kDa protein.

B10.BR mice lymph node cells responded well to peptides 1-20,16-35, 31-50, 61-80, 106-125 and 136-148 with smaller responses to 46-65, 76-95, 121-140, and no response to peptide 91-110 (Figure 2.3.37). In contrast BALB/cJ lymph node cells responded well to peptides 91-110, 121-140 and 136-148 with smaller lymph node responses observed to peptides 1-20, 16-35, 31-50, 61-80 and 76-95.

As discussed previously when H-2^d mice were immunised with 18kDa protein *in vivo* , their lymph node cells when challenged *in vitro* with synthetic peptides responded only to peptide SLP-2 (111-125). This therefore implies that the epitopes resulting from the *in vivo* processing of the 18kDa protein could be different, yielding only one epitope capable of being recognised by CD4$^+$ T cells. A comparison between the T-cell proliferative (Figure 2.3.37) and IgG responses to peptides 1-20, 16-35, 31-

50 and 76-95 (Figure 2.3.35) show that there were peptide specific responses in B10.BR mice and not in BALB/cJ mice. As the proliferation assay measures the response of $CD4^+$ only, it would be expected that these cells were of the T_H cell type, therefore the difference in the responses must be due to the difference of T_H1 and T_H2 subpopulations[148,149]. As these four peptides each elicit IgG responses, this assay would measure the functional activity of the T_H2 cells. The failure of the development of IgG in BALB/cJ mice could be due to activation of a subpopulation of T-cells that do not induce IgG responses in particular the T_H1 subpopulation.

The peptides that do not elicit T-cell proliferative responses 46-65 and 106-125 in BALB/cJ mice and peptide 91-110 in B10.Br mice could lack class II MHC-binding epitopes for the haplotypes concerned, or there may be an absence of responder T-cell for them. For the two other peptides 121-140 and 136-148, in which there are strong T-cell responses, but no IgG responses, the responding T-cells could be the T_H1 subpopulation for both strains of mice.

Protective immunity to any given pathogen must be viewed in the context of the antigenic epitopes they contain and the processes by which T_H effector phenotype is either selected or activated. A greater understanding of the types of response particularly the protective immunity responses is necessary before there can be development of a new vaccine. Despite the knowledge that the 18kDa protein stimulated nearly half of the *M.Leprae* specific T-cell clones, the type of activation of these T-cell clones could be the factor governing the immunological responses in humans, thereby explaining the relative immunological unresponsiveness to the mycobacterium in most of the population infected by the disease, and why so many people are resistant to it.

2.3.3 SUMMARY OF CHAPTER 2.

A total of sixteen peptides along with seven others were synthesised for immunologically assaying at the Auckland School of Medicine. The majority of the syntheses proceeded smoothly with only three peptides presenting difficulty, SLP-8, SLP-13 and SLP-16.

In subsequent (after completion of this work) syntheses of peptide 121-140 (SLP-13) under ABI protocols with capping was carried out [139]. In this synthesis the DNP group was replaced with the Boc histidine side chain protected with the Bom group which is cleaved in HF thus overcoming the problem of removing the DNP

group. However this synthesis had a much reduced yield of peptide on the resin with two failed couplings and an average coupling yield of 98.23%. Peptide 1-20 (SLP-5) was also repeated under ABI protocols and this synthesis also did not achieve the same results as the synthesis under Kent's protocols with the average coupling yield considerably lower (92.7%). Using the method developed by Milton *et.al.*[150] ,peptide 46-65 (SLP-8) was predicted to be a very difficult peptide to synthesise. When this prediction method was applied to the two previous peptides SLP-5 and SLP-13 the difficult coupling steps were where the two failed couplings occur in the synthesis using ABI protocols.

The synthesis of the peptides discussed in this chapter offers an independent assessment on the effectiveness of Kent's protocols which do not appear to have gained wide acceptance. Certainly in the cases discussed above Kent's protocols appear to out perform the standard ABI protocols and in one case for peptide SLP-16, the newer ABI NMP/HOBt protocols, even though Kent's cycles did not completely succeed with this peptide. The lack of a capping procedure (now incorporated into Kent's cycles in his laboratory[151]) did not appear to interfere with the final purity of the peptides synthesised, though in cases where there is a failed coupling, as for peptide SLP-11 it would have been preferable to have a capping step.

The binding site for the monoclonal antibody L-5 has been isolated, involving the key residues 109-115. There also appears to be residues outside this region around 101-109 that are also involved in the recognition of the antibody. The peptides SLP-5, SLP-6, SLP-7 and SLP-10 appear to carry B-cell and T_H-cell epitopes. These are recognised differently by the murine immune system in two breeds of mice,giving the discrimination between the two types of T_H-cell responses. In the murine system, the dominant T-cell epitope recognised on the 18 kDa protein is within the residues 110-125. However for the 18kDa protein it has been reported in studies using human T-cell lines that the dominant epitope recognised consists of residues 38-50[56]. Whether the results discussed here in the murine system can be transposed to the human model is as yet unknown. Certainly there have been problems previously in transferring the results in murine models to human systems as with the case of the 65 kDa protein (discussed in Chapter 1).

The overall aim of the work discussed in this thesis is for the total synthesis of the 148 amino acid 18 kDa protein. The peptides synthesised in this chapter have offered some insight into the possible problems that could be faced in the total synthesis. In particular, the evaluation of Kent's protocols have proved valuable. Certainly it seemed that Kent's cycles out performed the standard ABI cycles and the

newer NMP/HOBt cycles. Therefore it was decided at this stage that Kent's cycles offered the best chance for the successful total synthesis of the 148 amino acid 18 kDa protein. It was postulated that the distance from the resin would alleviate any potential difficulty at any of the coupling steps past twenty five residues. Indeed most of the problems will occur in the C-terminal fragment the region closest to the resin due to the folding charateristics of the initial 15-20 residues[82]. The synthesis of peptide RLP-23 (131-148) was almost impossible to carry out on the Schwartz Mann synthesiser with extremely low yields of peptide. The two peptides synthesised in this region on the ABI synthesiser, apart from the low coupling percentages, which was probably due to the DNP histidine, have shown that there should not be any difficulty. The problem of the detection of low coupling yields could be solved by the use of the peptide resin sequencing method as discussed earlier. The peptides synthesised gave total coverage over the whole protein. There seemed at this stage to be no difficult coupling steps that might not be overcome in the synthesis of the longer peptides and the full length protein.

Chapter 3.

Protein Synthesis.

3.1 INTRODUCTION.

Soon after Merrifield reported the first automated solid phase synthesis, papers were published demonstrating how the technique could be used to synthesise small proteins from 50 to 190 amino acids (see table 3.1 and references therein). Many of the improvements in peptide chemistry came about as a result of attempts to synthesise proteins containing about 100 amino acids. As shown in Table 3.1.1 there has been a wide variety of proteins synthesised nearly all with Boc chemistry, with the major driving force being production of increased quantities of proteins, as proteins are often present in minute quantities in biological systems. The increased quantity allows investigation of the biochemical role of these proteins e.g. Apo CII[167] or investigation of the structural requirements for activity of a protein as in the case of interleukin 3[132c].

The original protein synthesised by Marglin and Merrifield[152], insulin, required synthesising two separate peptide chains, mixing them under reducing conditions and allowing them to oxidise together over a period of time to give a crude product with an activity of 8%. This compared favourably with native insulins which had been reduced then oxidised.

The synthesis of the protein that awakened interest in total chemical protein synthesis was the synthesis of ribonuclease A performed by Gutte and Merrifield[153]. This 124 amino acid protein was synthesised using the original protocols for peptide synthesis developed by Merrifield. The final yield of the purified product based on the original starting resin was 2.5%. However it was found that 83%, 1.4% per step, of the starting resin peptide chains were lost from the resin by cleavage of the anchoring benzyl ester bond. It was also discovered that some of the protecting groups were not stable to 50% TFA in DCM, for the entire length of the synthesis, leading to chain branching at some of the amino acid residues such as lysines. The final purified product had a specific activity of 78%. This synthesis provided further evidence that it was the primary sequence of amino acids that determined the tertiary and enzymatic activity of proteins.

Despite the drawbacks discovered during the synthesis of ribonuclease A there were many other syntheses attempted at this time with small variations in the synthetic protocols, or changes in the protecting groups for some of the amino acids. The attempt to synthesise human pituitary growth hormone by Li and Yamashiro[154] demonstrated

PROTEIN	AMINO ACID	YEAR	TYPE OF SYNTHESIS	REFERENCE
Insulin	55	1966	Original Merrifield Boc AA Chloromethyl resin	Marglin and Merrifield[152]
Ribonuclease A	124	1969-71	Original Merrifield Boc AA Chloromethyl resin	Merrifield and Gutte[153]
Human Pituity Growth Hormone	188	1970	Original Merrifield Boc AA Chloromethyl resin	Li and Yamashiro[154]
Bovine Pancreatic Trypsin Inhibitor (BPTI)	58	1971 1976	Original Merrifield Boc AA Chloromethyl resin. Boc Merrifield chemistry New protecting groups HF cleavage	Noda *et.al* [155] Tam and Kaiser[156]
Cobratoxin	62	1972	Boc Merrifield chemistry 4 disulphide bonds	Aoyagi *et.al* [157]
Lysozyme	129	1972-1973	4 syntheses using Boc Chemistry modified with DMF as coupling solvent	Barstow *etal* [158] Sharp *et.al* [159]
Acyl Carrier Protein (ACP)	74	1971	Manual then automatic Boc chemistry. Shrink and swell Chloromethyl resin 1. HF cleavage 2. HBr /TFA and H_2 Pd cleavage	Hancock *et.al* [160]
Cytochrome C	104	1972	Boc chemistry HF cleavage. Chloromethyl resin	Sano *et.al* [161]
Lipoprotein CI	57	1976	Low substitution chloromethyl resin Boc chemistry HF cleavage	Harding *et.al* [162]
Ovine β Lipotropin	91	1978	Brominated styrene resin Boc chemistry preformed symetrical anhydrides HOBt active esters HF cleavage	Yamashiro and Li [163]

Table 3.1. Proteins synthesised using solid phase chemistry developed by Merrifield. The chemistry used for the synthesis of the proteins is indicated.

PROTEIN	AMINO ACID	YEAR	TYPE OF SYNTHESIS	REFERENCE
Human Parathyroid Hormone	84	1983	Boc Chemistry Pam resin P-Nitro active esters HF cleavage	Fairwell *et.al* [164]
Interleukin 3	140	1985	Optimised Boc Chemistry in DMF Pam Resin ABI 430A Lo-Hi HF	Clarke Lewis and Kent [132c]
Murine Transforming Growth Factor-α (TGF-α)	53	1986	Optimised Boc Chemistry Pam resin.Lo-Hi HF	Heath and Merrifield [165]
Human Transforming Growth Factor-α (TGF-α)	50	1986	Optimised Boc Chemistry Pam resin.Lo-Hi HF	Tam *et.al* [166]
Apoliprotein CII	79	1987	Automated Beckman instrument Pam resin Boc Chemistry S_n1 HF	Fairwell *et.al* [167]
Porcine Cardiodilatin	88	1988	ABI 430A Pam resin Boc chemistry Lo-Hi HF	Nokihara and Semba [168]
Human Granulocyte-Macrophage Colony-Stimulating Factor	127	1988	ABI 430A Pam Resin Optimised cycles DMF Lo-Hi HF	Clarke Lewis *et.al* [169]
Human Immunodeficency virus Protease.	99	1988 1988	ABI 430A all DMF cycles Pam Resin Lo-Hi HF	Schneider and Kent [170]. Nutt *et.al* [171]
Ubiquitin	77	1989 1989	FMOC ABI 430A FMOC NPS 4000 synthesiser Bop/HOBt as coupling catalyst	Ramage *et.al* [172] Briand *et.al* [173]
Human Interleukin 1-β	153	1988	Boc chemistry Pam resin ABI 430A Lo-Hi HF	Lobl *et.al* [174]

Table 3.1 Continued

that proteins up to 188 amino acids were capable of being synthesised with some biological activity. Unfortunately the wrong sequence for the protein was synthesised[175].

The synthesis of Cobratoxin by Aoyagi *et.al.* [157] with its four disulphide bridges investigated the formation of disulphide bonds, as it was becoming clear that one of the major drawbacks in the synthesis of proteins with cysteine, was the correct formation of disulphide bonds to give correct tertiary structure. It has now been demonstrated by the synthesis of α–mTGF[165] and α–hTGF[166] proteins, which both have 3 disulpide bridges, that it is possible to correctly form the disulphide bond. The use of new protecting groups that had greater stability to acid treatments, in a synthesis of BPTI, produced a much cleaner product[156]. The new protecting groups 2-chlorobenzyloxycarbonyl (2ClZ) for tyrosine and chlorobenzyloxycarbnyl (Cl-Z) for lysine reduced the loss of the protecting groups thereby reducing the possibility of side chain branching. Asparagine and glutamine were coupled as the *p* -nitrophenol ester, for 18h in each case, to reduce dehydration of the side chain caused by using the symmetrical anhydrides of these amino acids. The synthesis was also followed using the ninhydrin assay, where a positive test (i.e. colour development) resulted in that amino acid being recoupled then capped if the test was still positive. After cleavage the protein was subject to purification using a trypsin column to which correctly folded BPTI would bind while the other material was eluted off the column. These protocols yielded a protein that was 80-90% as active the native inhibitor. Further purification by ion exchange chromatography gave a protein that behaved identically with the native protein inhibitor, as judged by CD spectra (indicating correct formation of secondary structure), tryptic mapping and inhibitor kinetics. This synthesis was more successful than two previous syntheses of BPTI, ie, a continuous solid phase synthesis by Noda *et.al* [155] and a fragment condensation synthesis on a solid support by Yajima *et.al* [176].

With the losses of considerable amounts of peptide chains from the resin, a more acid stable link between the peptide and the resin was required to reduce the losses brought about by the repeated TFA treatments. The synthesis of ovine-β-lipotropin utilized a brominated styrene resin which has greater than 20 times the acid stability of Merrifield's original chloromethyl polystyrene resin[163]. From 1980 onwards regular use in protein syntheses was made of the phenyl-acetamidomethyl (PAM) resin, developed by Merrifield's group. This resin with greater than 100 times the acid stability of the chloromethyl resin with respect to TFA has proven ideal for protein synthesis. The increased purity of products after cleavage soon became evident.[87] The use of the PAM resin has been proven in syntheses of proteins from murine-α-TGF[165] with its three disulphide bridges to the HIV protease enzyme[170].

The four syntheses of the 129 amino acid protein lysozyme by Barstow *et.al* [158] and Sharp *et.al* [159] demonstrated one of the fundamental problems in the synthesis of peptides and proteins. During the original synthesis, it was detected by ninhydrin analysis that coupling became more difficult as the chain length increased particularly past 40 amino acids. In subsequent syntheses of lysozyme the coupling yield was improved by altering the amino acid recoupling solvent from DCM to DMF, or by carrying out both the coupling and the recoupling step in DMF, or by using a mixture of DMF and urea. This observation relates to the solvation requirements of the peptide chain and the polystyrene matrix. As the peptide chain grows the solid support environment changes rapidly from a relatively apolar environment due to the polystyrene, to a polar environment due to the peptide chain and the amino acids[82]. The use of more polar solvent such as DMF would be required to provide greater solvation of the peptide chain and the amino acid to be coupled. To overcome the apparent sequence dependent coupling problems that occurred during the synthesis of lysozyme, capping with acetic anhydride after 2 or 3 recoupling steps was used, if there was still a positive result from the ninhydrin assay.

Hancock *et.al* [160] used a method of shrinking and swelling the polystyrene resin during the synthesis of ACP which supposedly exposed the previously buried functional groups, so that a recoupling step would lead to further reaction and also squeeze out any byproducts that were trapped within the resin beads[118]. The shrinking of the resin was carried out by washing the peptide-resin with butanol then reswelling the peptide-resin by washing with DCM. This method for supposedly increasing the yield of the peptide was disputed later when more sophisticated techniques for monitoring the internal matrix of the peptide-resin such as NMR became available[177]. It was discovered by Kent *et.al* [82] that the peptide and resin moieties mutually enhanced each other's solvation with no restriction on the degree of swelling of the peptide-resin, or the amount of peptide that was able to be synthesised on the resin. Hence the strategy used by Kent *et.al* [132c, 170] of keeping the peptide resin in its maximum swollen state by using DMF for all washes and couplings during the synthesis was successful in the synthesis of interleukin 3 and the HIV protease enzyme.

It is critical in solid phase peptide synthesis to maintain the highest possible purity for the chemicals used throughout protein synthesis, to minimise chances of possible side reactions such as chain termination. The chemistry used to make the resin can produce resin bound benzaldehyde groups that react with α-amino groups to give Schiff base imines. This is a reversible reaction which can produce aldehyde groups randomly over the whole protein reforming Schiff bases at each step on the growing

protein chain[177]. Another major source of terminated peptides has been impure Boc amino acids contaminated with TFA resistant secondary protected amino acids thereby causing chain termination (ie *sec* butyl amino acids).[117]

Along with the improved yield of pure protein on the resin, modified procedures have also been developed for HF cleavage of proteins[125]. The S_n2 deprotection method removes the side chain protecting groups using a low concentration of HF in dimethyl sulphide (DMS), which scavenges the harmful carbonium ions before they react with the peptide chain. The S_n2 process is then followed by the standard high concentration HF cleavage procedure, an S_n1 process, which cleaves the peptide anchor to give the free protein. This procedure has been used extensively in most protein synthesis cleavages since 1983 and has been very successful with cleavage of the α-TGF proteins which are particularly susceptible to damage in the HF cleavage[165,166].

In contrast however, Fairwell *et.al* [167] who synthesised apolipoprotein CII discovered that they achieved better yields of protein with the ordinary Hi 90%, HF cleavage, which gave less aggregated material than the Lo-Hi procedure. As aggregation of the protein occurred throughout the purification, it was important to reduce the amount of aggregated material resulting from the HF cleavage.

Reverse phase HPLC as a purification method was first used in synthetic protein purification of interleukin 3[132c]. However the most powerful purification procedure still appears to be gel filtration, for desalting, scavenger removal, size separation from terminated peptides and separation of aggregated products. As a final step, ion exchange chromatography can raise an 80-90% pure protein to close to 100% purity.

An interesting strategy has been used by Diebel and Lobl in the synthesis of human-interluekin-1β[174]. They used capping throughout the synthesis, then removed the N-terminal Boc protecting group and biotinylated the N-terminal amino acid of the protein . After the HF cleavage, a single step purification was used, where the biotinylated protein was bound to an avidin agarose column and the non-biotinylated material washed off. The biotinylated protein was then eluted off the column providing pure protein with a bioactive label at a specific site. This single step process could become a valuable tool in the future for the synthetic peptide chemist.

Most chemical protein syntheses have utilised the t-Boc protecting group, with the associated benzyl based side chain protecting groups, and HF cleavage. There have been some FMOC protecting group protein syntheses two of which were carried out on the protein ubiquitin, a 77 amino acid protein[172]. Another group[173] compared a new

coupling catalyst BOP with the standard DCC procedures used by Ramage *et.al* [172] and found that the new coupling catalyst gave a higher yield of protein. Ramage *et.al* [172] have since repeated their synthesis working on altering the protecting groups of the protein and looking at the differences in the purity of the product.

With the advent of the synthesis of longer peptides, it is now becoming possible to synthesise whole structural domains of proteins. By understanding the factors that determine protein structure and function, it could be possible to customise proteins as working enzymes. This has been demonstrated by Hahn *et.al* [178] where they synthesised a four helix bundle synthetic protein that exhibited enzymic activity similar to chymotrypsin. This was achieved by placing at the amino terminal of each helix, a critical residue that occurs at the serine proteases catalytic site, ie: serine, histidine or aspartic acid, in the same spatial conformation as in chymotrypsin.

At the time the synthesis of the *M.leprae* 18 kDa protein was being considered, the only people carrying out regular chemical protein synthesis were Kent and Clarke-Lewis. Kent kindly provided us with a copy of the cycles that had been used for the synthesis of interleukin 3 and were currently in use in his laboratory for the synthesis of the HIV protease enzyme. It was felt that FMOC chemistry had not progressed sufficiently to enable a synthesis of this length to be undertaken and so a synthesis using this method was not attempted. The newer NMP/HOBt cycles were relatively untried in the synthesis of larger protein molecules and hence these cycles were not used. In addition Boc chemistry of longer proteins had the best track record and seemed still the least expensive.

Kent's cycles were tested in the synthesis of the peptides in Chapter 2 and it was shown that Kent's synthetic protocols appeared to out perform the standard ABI cycles, including the newer NMP/HOBt chemistry used in the second synthesis of SLP-16. The synthesis of the 37 amino acid peptide RLP-39 (95% coupling yield, with an overall predicted yield of 20%) showed that it would be unlikely that the standard ABI-Boc chemistry cycles without substantial modification, would allow the synthesis of even a 100 amino acid protein fragment of this particular protein due to low coupling yields, as shown by the ninhydrin analysis. The use of Kent's cycles would also provide an independent assement of these protocols and there is little evidence of their use outside Kent's or Clarke-Lewis's laboratories.

The strategy used for the total synthesis was to synthesise smaller peptide fragments of the total protein such as SLP-1 to 3 (Chapter 2). As preparation for the 148 amino acid synthesis this strategy allowed the setting of instrument variables, such

as flow rates, to be evaluated. The next stage would involve the synthesis of larger 50 mer fragments which are described in this Chapter. An examination of the coupling yields would determine if there were any difficult coupling areas that needed closer examination. Following this the full 148 amino acid protein synthesis was attempted.

3.2 Experimental

3.2.1 General Experimental.

HPLC purification protocols, amino acid analysis, sequencing, Boc amino acid derivatives and solvents used were as previously described for the synthesis of the peptides in Chapter 2.

3.2.1.1. Gel filtration

Gel filtration was carried out using a Superose 12 preparative grade HR 16/30 Pharmacia column (packed in house with 10,000 theoretical plates) attached to Pharmacia P-500 pump. The detector was a Pharmacia UV-1 detector fitted with a 280 nm filter or a CE212 Variable Wavelength (Cecil Instruments, UK) and the chart recorder was a Pharmacia single pen, or a Sekonic SS 250F chart recorder. A Pharmacia MV-7 injector was used and the fraction collector was a Pharmacia Frac 100. The Sephadex columns were gravity fed and used the same detectors , chart recorders and fraction collector described above

3.2.1.2. Ion Exchange Chromatography.

Ion exchange chromatography was carried out using a Pharmacia FPLC instrument with a Mono Q HR 5/5 column. The eluents were A: 50 mM ammionium hydroxide:acetonitrile (9:1) and B: 1M sodium chloride with 10% acetonitrile. A gradient of 0-30%B in 35 min then 30-100% B in 5 minutes was used.

3.2.1.3. Polyacrylamide Gel Electrophoresis.(PAGE)

Polyacrylamide gel electrophoresis (PAGE) was performed with the SDS discontinuous system as described by Laemmli[180] consisting of the discontinous Ornstien[181] and Davis[182] buffer system with SDS. The PAGE gels were of 15% polyacrylamide concentration.

Samples (0.5 mg) were dissolved in sample buffer (10 mM Tris.HCl, 1 mM EDTA, pH 8) to which was added 20% SDS, 25 µl and the resulting solution boiled for 5 min and then centrifuged at 13500 rpm for 10 min using a microcentifuge (MSE Microcentrifuge, Kempthorne Medical Supplies NZ). Usually 25 µl of this sample was loaded onto the gels. The gels were run at constant current of 22 mA for approximately 2.5 h. The gels were removed from the glass plates and placed in Coomassie staining

solution (30% MeOH, 10% acetic acid and 0.625g Coomassie Brilliant Blue R (Sigma USA)) and shaken gently overnight. The following day the gel was destained using 30% MeOH and 7.5% acetic acid solution for 2h.

Molecular weight (MW) markers used were egg albumin (45000), glyceradehyde -3-phosphate dehydrogenase (36000) and low molecular weight markers as supplied in the Pharmacia (Sweden) Electrophoresis Calibration Kit (part number 17-0551-01) at 17200, 14600, 8240, 6380, and 2560 kDa.

3.2.1.4. Polyacrylamide Gel Electrophoresis using the Phast System

The gels used were precast 20% homogeneous gels purchased from Pharmacia, NZ. Samples (0.5 mg) were dissolved in sample buffer (10 mM Tris.HCl, 1 mM EDTA, pH 8) to which was added 20% SDS 25 μl and boiled for 5 min then centrifuged at 13500 rpm for 10 min using a microcentifuge (MSE Microcentrifuge, Kempthorne Medical supplies NZ). The gels were run according to the standard method (Pharmacia Separation File 111) at 250V, 10 mA, 3.0W at 15°C for 95Vh after the samples had been loaded with the automatic sample loader. The gels were developed using the standard silver staining process for SDS gels (Development File 210)

3.2.1.5. Mass Spectrometry

Mass spectrometry on protein samples was carried out courtesy of A Jones (Bond University) using an API III Biomolecular Mass Analy ser. This instrument was a triple quadrapole mass spectrometer equiped with ion evaporation and an ion spray probe. The instrument was calibrated using PPG/NH4Ac. The instrument had a mass range of 1-2400.

Mass specrometry on the tryptic digest fragments was carried by liquid secondary ion mass spectrometery (SIMS) on a VG70-250S double sector mass spectrometer couresy J.Shaw and J. Allen, DSIR, PN. The peptides were dissolved in a glycerol matrix which was acidified with HCl and bombarded with a stream of cesium ions to produce molecular ions which were accelerated by a fast atom gun. Detection was according to their charge to mass ratio. The instrument was calibrated according to the different molecular ions of glycerol ((Glycerol)$_1$H+, (Glycerol)2H+, etc). More accurate mass calibration was obtained from peaks produced by the bombardment of cesium iodide.

3.2.1.6. Tryptic Digests.

Trypsin was purchased from Sigma (Bovine Pancreas TCKP treated 12,000 BAEE units/mg of protein). HPLC was performed using the analytical system described previously equipped with a Waters UK6 injector and a SynChropak RP-P column (C-18, 300A pore, 6-5m partical size) purchased from Synchrom Inc (Indiana USA). The solvent system used was 0.1% formic acid in H_2O for buffer A and 0.1% formic acid in 80% CH_3CN for buffer B[183].

Synthetic or recombinant 18kDa was dissolved in ammonium bicarbonate (50 mM 990 µl) and the pH adjusted if necessary to pH 8. Trypsin was added to give a ratio of trypsin to protein of 1:100 (w:w). After 18h at 33-35°C the pH was adjusted to pH 2 by the addition of 2M HCl (approximately 100 µl). Samples of the digest were injected manually onto the HPLC column and a gradient of 0-10-35% B was run over 0-10-80 min and the eluted peaks trapped. The collected fractions were dried *in vacuo* (Speed vac, Savant) then redissolved in 50% formic acid for loading onto the sequencer for identification. The major peaks were also subject to mass spectrometry

3.2.2. PEPTIDES SLP-17, SLP-18, SLP-19.

3.2.2.1. Peptide 101-148.SLP-17

The first synthesis of this peptide was carried out using Kent's protocols on a 0.3 mMol scale (393 mg of 0.76 mMol/g of Boc-alanine-PAMresin) using solvents purchased from ABI. DCM was inadvertantly placed in the DMF bottle position and as a result the coupling yields for amino acids in cycles 17-18 averaged 70%. The average coupling yield per amino acid was 95.41% with an overall yield for the peptide on the resin of 9.1%. The yield of the dried peptide resin was 0.74 g, a weight increase of 400 mg which was considerably down on the expected weight increase of 1.55 g

The second synthesis of this peptide was carried out using Kent's protocols on a 0.3 mMol scale (393 mg, 0.76 mMol/g of Boc-alanine-PAM resin), with solvents purified in house as described in Chapter 2. The peptide resin (326 mg) was treated to remove the DNP group, using the β-mercaptoethanol method, then the N-terminal Boc group was removed as discussed in Chapter 2. The peptide was cleaved using a modified Lo-Hi HF procedure of Tam *et.al* [125]. The peptide resin was placed in the reaction vessel with anisole (9 ml) and HF (1 ml) was distilled into the vessel. The

mixture was then stirred for 1.75h at 0°C. After evaporation of the HF and anisole the resin was washed with ether (3 x 20 ml) and DCM (3 x 20 ml) prior to drying by sucking air through the funnel. The dried resin was returned to the HF vessel with anisole (1 ml) and HF (9 ml). After stirring at 0°C for 45 min, the HF was evaporated in vacuo and the peptide precipitated in ether (20 ml). The peptide and resin were filtered from the ether and washed with a further 2 x 30 ml of ether. The peptide was then dissolved by sequential addition of H_2O (20 ml), 5% acetic acid (AcOH, vol:vol, 2 x 40 ml), 20% AcOH (2 x 40 ml) and finally with 50% AcOH (2 x 20 ml). After diluting the AcOH peptide solution to give 5% AcOH, the solution was lyophilised to give 78.4 mg of crude peptide.

The crude peptide (78 mg) was chromatographed on the preparative HPLC with a gradient of 0% - 60% B run over 4 h after the elution of the solvent peak. Fractions were isolated across the peptide peak and checked on the analytical HPLC before fractions were combined on the basis of their retention time giving a total of 18.9 mg of purified peptide. A sample of the purified peptide was hydrolysed and sequenced to determine the amino acid composition. From this it was shown that the wrong sequence had been synthesised, with a glutamine instead of a glutamic acid at postion 119.

Another sample of this peptide was removed during the synthesis of the total protein (326 mg). After removal of the DNP group and the N-terminal Boc group the peptide was cleaved using a normal 90% HF cleavage method as described in Chapter 2, with anisole (1 ml) as the scavenger. After washing the peptide and resin with ether the peptide was dissolved as discussed above and the peptide solution lyophilisedto give 97 mg of crude peptide. The crude peptide (47 mg) was purified by preparative HPLC with a gradient of 0% - 60% B over 2h run after elution of the solvent peak. Fractions across the peptide peak were checked on the analytical HPLC and combined on the basis of their retention times to give four fractions with a combined peptide weight of 28.3 mg.

Amino acid analysis and sequencing was carried out on the purified peptide to determine the structure of the peptide. A SDS-PAGE gel was run to check the molecular weight and purity of the peptide.

3.2.2.2. Peptide 50-100 SLP-18

This peptide was synthesised using Kent's protocols on a 0.4 mMol scale (0.588 g of 0.72 mMol/g of Boc-alanine-PAM resin) using solvents purchased from ABI. The peptide was cleaved from the resin after removal of the N-terminal Boc group, using an

ordinary (90%) HF cleavage with anisole (1 ml) as the scavenger. After stirring at 0°C for 1h the HF was evaporated off the resin and the peptide precipitated with ether filtered, and washed with ether. The peptide was progressively dissolved by using an increasing concentration of AcOH as discussed above, up to glacial AcOH as the peptide would not completely dissolve. After lyophilisation 79.4 mg of crude peptide was isolated.

Preparative HPLC was carried out on 74 mg of the crude peptide, with a gradient of 0% to 60% B run over 4 h, after elution of the solvent peak. The fractions from the peptide peak were checked on the analytical HPLC and combined on the basis of their retention time to give four fractions with a combined weight of 35.3 mg of peptide. Amino acid analysis and sequencing of the N-terminal of the peptide were carried out to determine the peptide structure. A SDS-PAGE gel was run to check the molecular weight and purity of the peptide.

3.2.2.3. Peptide 1-50, SLP-19.

This peptide was synthesised using Kent's protocols on a 0.3 mMol scale (0.396 g of Boc alanine PAM resin with a substitution of 0.76 mmol/g), using solvents purchased from ABI. During the synthesis it was observed that the in-line filter for the amino acids was blocked at cycle 12,and that there was solvent in the cartridges for phenylalanine. After replacing the inline filter, the synthesis was allowed to continue. At cycle 40 the DIEA ran out and had to be replaced with triethylamine. Peptide resin samples for phenylalanine were missing, giving no indication of the % coupling for these amino acids. Amino acid analysis of the peptide resin showed that the two phenylalanines did not couple to 100%. After completion of the synthesis and the removal of the N-terminal Boc group the peptide was cleaved from the resin using the Lo-Hi procedure developed by Tam *et.al* [125]

The Boc group was removed from the peptide resin (300 mg) as discussed in Chapter 2 then the peptide resin was placed in the HF reaction vessel with *p* -cresol (800 µl), *p* -thiocresol (200 µl) and dimethylsulphide (6.5 ml). HF (2.5 ml) was distilled into the reaction vessel and the mixture stirred at 0°C for 1.5h. After evaporating the HF for 20 min, the resin was washed with ether (20 ml) and filtered before washing further with ether (2 x 20 ml) and DCM (4 x 20 ml) and drying. The dried resin was returned to the reaction vessel with *p* -cresol (800 µl) and *p* -thiocresol (200 µl). HF (9 ml) was distilled into the reaction vesel and the mixture stirred at 0°C for 45 min, before the HF was evaporated from the resin in 20 min. The peptide was then precipitated with ether (20 ml) filtered and washed with three further ether washes. The peptide was dissolved

in 6M guanidine.HCl, 25 mM Tris, pH8 (20-30 ml) and checked on the analytical HPLC. The guanidine peptide solution was then gel filtered on a Superose 12 column in six runs. From each run, three major fractions were isolated. The first peak eluted was peptide, the second DMS and the third the remaining non volatile scavengers.

Preparative HPLC was carried out on the isolated peptide peaks with a gradient of 20-40% B run over 2.75 h at a flow rate of 3 ml/min, after elution of the solvent peak. Peptide fractions were isolated and combined according to their retention times on the analytical HPLC. This gave a total of 19.4 mg of purified peptide from two of the gel filtration runs. Amino acid analysis and sequencing were carried out on the purified peptide to determine the peptide structure. SDS-PAGE gel showed that the peptide had a band with the correct molecular weight plus two other higher molecular weight components.

3.2.3 PROTEIN SYNTHESIS.

3.2.3.1. The Synthesis of the 148 Amino Acid 18 kDa Protein.

The total synthesis of the 18 kDa, 148 amino acid protein was carried out on an ABI 430A peptide synthesiser, in a single continuous process lasting 5.5 days with constant monitoring of the instrument in case of an instrument failure. The amino acids were purchased pre-packed in their cartridges from ABI. DMF, TFA, DIEA, DCC, HOBt and ninhydrin reagents were purchased from ABI to ensure the highest purity of reagents. Analar grade DCM and MeOH were double distilled. The first distillation of DCM was from calcium hydride.

The programmer for the instrument would only allow entry of 99 amino acids, which necessitated splitting the synthesis into two halves of 72 cycles and 75 cycles. Dummy end cycles were used at the end of the first stage and dummy begin cycles used to restart the second stage. These cycles prevented the transfer of MeOH and DCM into the reaction vessel, that occurs in the normal begin and end cycles, thereby preventing the shrinking of the protein resin beads. The change over between the two stages took approximately 45 min, this being the time for the compiler to compile the 75 amino acids for the second half of the synthesis.

The synthesis was carried out on a 0.4 mMol scale, using 0.558g of 0.72 mMol/g of Boc-Ala-PAM resin. The low substitution was used in an effort to minimise volumes and achieve a high concentration of reactants, with each amino acid routinely

double coupled to reduce the chances of a failed coupling. This loading level was done nevertheless in an effort to produce as much protein as possible, with the recognition that some of the protein resin would have to be removed at various stages throughout the synthesis and that losses would occur due to resin sampling.

Due to the scale of the synthesis, two major samples of approximately 350 mg of protein resin were removed at cycle 47, residue Arg^{101} and at cycle 89, residue Arg^{59}. These samples would provide cleaved protein samples to monitor the success of the total synthesis up to 48 amino acids and up to 90 amino acids. At these two residues the synthesiser was paused and the reaction vessel filled with DMF then drained until the resin was saturated with DMF. The reaction vessel was opened, the sample removed, and the reaction vessel resealed. The vessel was then flooded with DMF flowing out the top of the reaction vessel to waste, ensuring that all of the air had been removed from the vessel. After draining the vessel was reflooded and drained again. This process was used in an effort to prevent a decreased or failed coupling in the next step of the synthesis, as a result of the presence of oxygen or moisture in the reaction vessel. Even so it is noticeable, that there was a decrease in the coupling yield after the removal of the second large sample. This observation is in accordance with similar observations of Kent [151] when a synthesis is interupted.

Resin samples were removed at the completion of every step up to cycle 101 where sampling was stopped, apart from cycle 122 and 136. In the initial stages of the synthesis, it was noticed that after 20 cycles the size of the resin samples were approximately 11 mg instead of the normal size of 7 mg. At the first major stop the reaction vessel cycles were reprogrammed to remove 1 min instead of two minute resin samples, reducing the resin sample size to approximately 8 mg. Despite the removal of the increased amount of resin during resin sampling, it was necessary to remove the two larger samples to ensure proper vortexing and mixing of the protein resin in the various reagents, and to prevent the resin from filling the 20 ml volume of the reaction vessel.

The resin sampling was stopped at cycle 101, residue Ile^{48} as the ninhydrins showed a sharp decrease in the % coupling and increasing sample size to over 10 mg. The decrease in the coupling yield could be due to the limit of detection for the assay being surpassed, which occurs at residue 60-70[121,133b]. Stopping routine resin sampling also prevented any further losses of the protein over the last 47 residues, as the weight increase per amino acid addition from this point on would not be keeping up with the size of the resin samples. Resin samples were removed at cycle 102, 122 and 136 so that protein resin amino acid hydrolysis and protein resin sequencing could be carried out to determine the % coupling yields. The resin samples taken at cycles 20,42,61 and

83 were split in half, to allow the protein resin to be sequenced and hydrolysed, as well as being used in the ninhydrin assay. At the completion of the synthesis there was **2.46** g of protein resin.

3.2.3.2. Cleavage #1

The protein resin (300 mg) was treated to remove the DNP group with β–mercaptoethanol then N-terminal deprotected as described in Chapter 2. The cleavage method is based on the Lo-Hi (S_n2,S_n1) method developed by Tam *et.al* [125].

The dried resin was placed in the HF reaction vessel with *p* -cresol (800 µl), *p* -thiocresol (200µl) and dimethyl sulphide (DMS, 6.5 ml). HF (2.5 ml) was distilled into the reaction vessel and the protein resin stirred at 0°C for 1.5 h. After the HF was evaporated *in-vacuo* for 20 min, ether (20 ml) was added, the resin was washed and filtered. After two further ether washes the resin was washed with DCM, dried and returned to the HF vessel. *p* -Cresol (800 µl), *p* -thiocresol (200µl) was added and HF (9 ml) was distilled into the reaction vessel. After 45 min at 0°C the HF was evaporated for 20 min and the resin thoroughly washed with ether (4x 20 ml). The protein was progressively dissolved in 5% acetic acid (AcOH, vol : vol, 3 x 40 ml) then 20% AcOH (2 x 40 ml), with mixing of the AcOH protein solution and resin. The AcOH solution was diluted with water, to give an approximately 5% AcOH solution, and lyophilised to give 67.9 mg of crude protein.

The first gel filtration step of the crude protein was a rapid desalting step used to separate the non volatile organic scavengers and exchange any HF salts away from the protein. The crude protein (10 mg) was loaded onto the Superose 12 column in 10% acetic acid using 0.1M acetic acid as the eluent. After the protein failed to elute from the column, the concentration of the acetic acid was increased up to 10% then 20%. Finally as the protein seemed bound to the column, the protein was eluted using 70% formic acid and lyophilised to give 10.2 mg of proteinaceous material. The recovered protein was dissolved in 2M guanidine.HCl along with the rest of the crude protein and gel filtered in 2M guanidine.HCl, loading 16 mg at a time on the Superose 12 column. Despite using guanidine.HCl, this procedure was essentially a desalting step and not a proper gel filtration.

The protein peaks isolated from the column were combined and loaded onto the preparative HPLC (as described in Chapter 2). The solvent peak was eluted at 10%B before applying a gradient of 10-37% over 2 h at a flow rate of 3 ml/min. Fractions

were taken across the protein peak and were combined on the basis of their retention times and lyophilised to give a total of 24 mg of protein. The fractions were subject to AAA and checked on a SDS polyacrylamide gel (SDS-PAGE) for an idea of the size distribution of the isolated material.

3.2.3.3. Cleavage #2

The protein resin (300 mg) was cleaved using the method as described for cleavage #1. The reaction time for the Lo-step was extended to 2 h and the Hi-step to 1.5 h to ensure that the protein was completely deprotected and cleaved from the resin. The protein was progresively dissolved in 5% AcOH (3 x 40 ml) to 20% AcOH (3 x 40 ml) solutions, with a final wash of the resin with a 70% AcOH solution (2 x 10 ml) to ensure the protein was totally dissolved from the resin. The AcOH protein solution was diluted with water to give a 5% AcOH solution and lyophilised to give 18.6 mg crude protein. The cleaved protein resin weighed 55.4 mg after HF cleavage, giving an expected yield of protein of approximately 240 mg.

The protein was loaded directly onto the preparative HPLC and after elution of the solvent peak, a linear gradient of 20-50% B over 2h at 1.5 ml/min was run. Fractions were combined according to their retention time on the analytical HPLC and lyophilised. It was noted that some of these fractions had a distinct yellow tinge, probably as a result of the DNP group still being bound or associated with the protein.

3.2.3.4. Cleavage # 3

The Lo-Hi HF cleavage was performed as described previously for Cleavage 2. After the ether washes, the protein was dissolved in H_2O then in 5% AcOH (2 x 30 ml), to 20% AcOH (3 x 30 ml) then to 70% AcOH (2 x 20 ml) as a final wash, and lyophilised to give 141.3 mg (54 %) of crude material. The weight of the cleaved resin was 43.2 mg, leaving unaccounted 117 mg of protein.

To see if the unaccounted protein had not precipitated in the ether and had been washed through the sinter of the funnel, the ether fraction was washed with saturated sodium bicarbonate to neutralise any residual HF, then separated. The ether layer was evaporated *in vacuo* while the bicarbonate wash was reacidified with acetic acid and lyophilised. The lyophilised material was dissolved in 50% acetic acid and desalted on a Sephadex G-25 column (46 x 560 mm) in 50 % acetic acid at approximately 3 ml/min. No protein was detected at 280 nm (LKB Uvicord 280 nm detector) or in analytical

HPLC of fractions from the gel filtration column. There was also no sign of any protein in the residual oil from the ether fraction, also checked by analytical HPLC.

Two 50 mg samples of the crude protein were dissolved and loaded on the preparative HPLC column from which the major protein peak was trapped to give 32 mg of protein material. A sample (0.2 mg) was sequenced to determine if the N-terminal amino acids were presentin the correct order. From this it was detected that benzyl ether protecting groups on the two threonines at position 5 and 21 were present in the molecule.

3.2.3.5. Removal of the Residual Protecting Groups using Hard Acid Soft Base.

This deprotection scheme is based on the cleavage method published by Yajima *et.al.*[184]

A sample of the protein (11.7 mg) was treated with 1M trimethylsilane trifluorosulphonate-thioanisole (1:1) in TFA (4 ml) with *m* -cresol (15μl) and ethanedithiol (12μl). After stirring in an ice bath for 2 h, anhydrous ether (50 ml) was added which gave a fine white precipitate. The ether solution was centrifuge in sealed vessels at 3000 rpm for 20 min then the ether layer was decanted off from the solid white pellet. The pellet was resuspended in fresh ether and centrifuged for a further 20 min, after which the ether was decanted off and the pellet suspended in H_2O. The pH was adjusted to 8.5 with triethylamine and the resulting suspension stirred for 30 min at 0^oC with β-mercaptoethanol (400 μl) and 1M ammonium fluoride (400 μl). The solution was reacidified to pH 5 with 10% aqueous AcOH and desalted on the Superose 12 column using 1M AcOH at 0.5 ml/min. The first 280 nm peaks from the gel filtrations were combined and lyophilised.

The isolated protein fractions were then loaded on to the preparative HPLC column. After elution of the solvent peak, a gradient of 20-50% B in 45 min was applied. Analytical HPLC was used to identify the fractions containing the protein, which were then combined and lyophilised to give 5.8 mg of protein. Sequencing of this material showed no sign of the benzyl protecting group on the N-terminal threonines.

Therefore the bulk of the crude protein from the third cleavage was combined with the best fractions isolated from the second cleavage and treated in the same manner as described above. After desalting using the Sephadex G-25 column with 50% acetic acid at a flow rate of 3 ml/min, five fractions containing protein were isolated. These

samples showed no differences on the analytical HPLC, so they were combined and gel filtered using the Superose 12 column and 50% acetic acid at 1ml/min as the eluent. The protein peak was then loaded onto the preparative HPLC where a linear gradient of 20-50% B in 80 min was used, after elution of the solvent peak. The protein peak eluted off the HPLC was trapped and lyophilised to give a total of 7.4 mg of protein per run. A sample from one of the fractions from the preparative HPLC was sequenced and showed no sign of the protecting group on the threonine.

A sample of the protein isolated from the preparative HPLC was dissolved in 6M guanidine.HCl and stirred gently at 30°C for 30 min before being dialysed using 1000 molecular weight (MW) cut off dialysis tubing (Spectropor 6, Gibco NZ) against decreasing concentrations of guanidine HCl (4M-2M-Mili-Q water). All steps were carried out for 24 h at 4°C, after which no apparent difference in the HPLC profile was detected.

SDS-PAGE on the Phast System, was run on the fractions isolated from the preparative HPLC. The gel was silver stained and showed no sign of any synthetic protein corresponding to the recombinant 18 kDa protein with every synthetic protein lane showing a broad band travelling further than the recombinant protein.

The remaining material was subjected to a tryptic digest as described above but after several analytical HPLC runs it became clear that a C-18 and not a C-4 column was required and that the best solvent system was the formic acid acetonitrile system. An attempt to isolate tryptic fragments using the preparative C-18 column failed with no tryptic fragments being detected. Preparative chromatography on future tryptic digests was therefore changed to an analytical C-18 HPLC column, as described below.

The faint yellow colour present with the protein material could be due to DNP still attached to the histidines. Therefore a sample of the protein from this cleavage was treated with 20% β-mercaptoethanol in 6M guanidine.HCl for 1h, then the whole sample was desalted on the Sephadex G-50 column (46x 370 mm) using 6M guanidine.HCl, 25 mM Tris.HCl at pH 7. A sample of the desalted protein was dialysed using 1000 MW cut off tubing as described above, and compared on SDS-PAGE with untreated material which was either dialysed or undialysed. Two distinct bands were present, one at the 18 kDa mark and one corresponding to twice that molecular weight.

3.2.3.6. Cleavage #4

See flow diagram figure P.1 for the purification protocol in the pouch at the back of the thesis.

The cleavage of protein resin (298 mg) was as described above for cleavage 3 with the exception that the Lo HF step was altered so that the HF was evaporated over 1 h. During this hour, after evaporation of most of the HF and DMS, the remaining 3 mls of DMS had warm H_2O applied to the vessel to ensure the complete removal of the DMS and the HF. The resin was then washed as before and returned to the HF vessel with *p* - cresol (500µl) and 1,4-butanedithiol (500µl) to ensure the complete removal of the formyl group from tryptophan[138]. The HF was then evaporated for 20 min and the protein dissolved in 5-20% acetic acid with a final wash of 50% acetic acid which gave 192 mg of crude protein after lyophilisation.

The crude protein was dissolved in 50% Acetic acid and desalted on a Sephadex G-50 (fine) column (46x 370 mm) with a flow rate of 1 ml/min. This gave a single major protein peak which was cut into two fractions of 64 and 65.5 mg respectively. The first fraction, which gave the best result on the polyacrylamide gel was sequenced for 28 residues. This showed that the benzyl group had been removed from the threonines during the HF cleavage. The bulk of these two fractions were stored at - 18°C to be further purified once a purification scheme to overcome the problems with the streaking pattern and aggregation on the gel had been found. AAA on the crude material and the two fractions showed that the synthetic product closely agreed with the expected ratios of the protein.

A small sample (5 mg) of the protein from the second half of the protein peak from the gel filtration column was dissolved in 6 M guanidine.HCl, 25 mM Tris.HCl at pH7 and dialysed using 1000 MW cut off tubing against decreasing concentrations of guanidine (6M-4M-2M-0M) with 25mM Tris at pH7 for 6 h, 24 h, 24 h and 4 x 12 h respectively at 4°C. The protein solution was then lyophilised. A polyacrylamide (20 %) SDS gel electrophoresis of the dialysed material showed no improvement in the streaking pattern or the 18 kDa band or the band at around twice the molecular weight.

3.2.3.7. Cleavage #5

See flow diagram figure P.2 for the purification of the protein in the pouch at the back of the thesis.

This cleavage method is adapted from that published by Tomisch *et.al* [185] for the cleavage of peptides that aggregate strongly.

The DNP group and the Boc group were removed from the protein resin (300 mg) as described above. The Lo- HF step is as for cleavage four, while the Hi- HF step was altered by the addition of SDS (6.5 g). After evaporation of the HF, the protein resin was washed with ether and the protein dissolved in 5-20% acetic acid. After lyophilisation of the protein SDS mixture, 568.1 mg of material was isolated.

The crude protein was desalted on a Sephadex G-50 column (46x 370 mm) run at 1 ml/min in 0.1M NaCl, 30mM SDS and 25 mM Tris.HCl at pH 8. Fractions from column could not be checked on the normal HPLC system due to the precipitation of the SDS in the acid solution. Therefore HPLC chromatography was performed using a basic solvent system of, buffer A 0.1M ammonium bicarbonate pH 8 and buffer B 0.1M ammonium bicarbonate pH 8:CH_3CN:isopropanol (1:1:1). The protein peak was trapped and sequenced to determine whether the benzyl group had been removed from the threonine during the HF cleavage. The amount of protein in the sample was very small but the benzyl group was detectable on the threonine at position 5.

In an attempt to keep the protein in a disaggregated state while removing the excess SDS, the protein was subject to hydrophobic interaction chromatography (HIC) using a phenyl Superose HR5/5 column attached to a Pharmacia FPLC system. However when the peaks were trapped, no protein was able to be detected on the analytical HPLC.

The protein SDS sample was therefore gel filtered using the Sephadex G-50 column (46 x 370) run at 1 ml/min, using the same solvent system as described above, in an attempt to remove more of the excess SDS. The protein fractions were detected on the analytical HPLC. The protein fractions were then further gel filtered on a Sephadex G-100 column (15 x 280 mm) at a flow rate of 0.3 ml/min in 0.1M NaCl, 30mM SDS and 25 mM Tris.HCl at pH 8, with greater resolution of the protein peak, which was compared on the analytical HPLC with the G-50 gel filtration samples. The peak from the gel filtration was trapped and checked on a SDS-PAGE gel, which showed most of

the protein precipitating at the gel interface between the stacking gel and the running gel. The amount of protein running on the gel was found to be dependent on the % of SDS in the samples. From these gels it was noticeable that there is a major band at the 18kDa mark that compared to the recombinant material and that the protein is streaked down the lanes from the top of the gel to below the 18 kDa marker. There is no sign of the dimeric and trimeric material found in the previous cleavages.

The rest of the gel filtered protein material was dissolved in the eluent used in the gel filtrations and ultra filtered through a 1000 molecular weight (MW) cut off membrane (Amicon ultrafiltration unit.USA) to ensure the SDS is washed away from the protein as much as possible. Two 50 ml volumes of the eluent were filtered through the membrane then two 50 ml volumes of H_2O. After this the retentate which held the protein was able to be chromatographed on the normal TFA based HPLC system. The protein at this stage was a very bright yellow colour. Filtration of the protein through a 10 000 MW membrane did not separate the bright yellow colour from the protein which was probably due to residual DNP being bound or tightly associated with the protein. After further purification using the analytical HPLC system, the protein peak showed no improvement in protein purity on SDS-PAGE.

3.2.3.8. 90 mer protein (Arg^{58}-Ala^{148})

See flow diagram figure P.3 for the purification protocol in the pouch at the back of the thesis.

After the problems with the previous cleavages it was decided to try purification of a protein fragment that does not contain some of the more sensitive amino acids such as tryptophan or a large number of methionines.

The DNP group was removed using the thiophenol method for 1h and the Boc group was removed as described previously. The Lo HF cleavage was carried out with *p* -cresol (1ml), DMS (6.5 ml) and 2.5 ml of HF at -5 - 0°C for 1.5 h The HF and DMS were evaporated until 1 ml of solution remained. This residual solutionwas filtered from the resin in the initial ether wash. The washed and dried resin was returned to the HF vessel with anisole (1ml) and HF (10 ml) was distilled into the reaction vessel which was kept at -5 - 0°C for 1h. After evaporation of the HF, the resin was washed with ether and the peptide dissolved in 6M guanidine.HCl and 10 mM Tris.HCl at pH7 (4 x 10 ml).

The protein was then gel filtered, in a solution of 6 M guanidine.HCl and 10 mM Tris.HCl at pH7, on a Sephadex G-50 column (46 x 370 mm) run at 0.3 ml /min and the protein peak as detected at 280 nm was trapped This column was run in the same manner as for the previous desalting columns in order to rapidly remove most of the non volatile organic scavengers from the HF cleavage. Analytical HPLC on the fractions from the gel filtration showed a much cleaner product compared to all of the previous 18 kDa protein samples.

The protein peak from the gel filtration column was loaded in two equal portions directly onto the preparative HPLC. After elution of the guanidine.HCl from the column in the solvent peak, a gradient of 0-60% B in 60 minutes at a flow rate of 2ml/min was run. The protein peak in both preparative runs was divided in half. The identical fractions between the runs were combined and lyophilised. The combined samples of protein from the HPLC runs were then redissolved in 6 M guanidine.HCl and 10 mM Tris.HCl at pH7 and loaded in two runs onto a Sephadex G-50 column (15 x 300 mm) run at 0.3 ml/min using 6 M guanidine.HCl and 10 mM Tris.HCl at pH7. This gave two peaks in both runs which were both trapped and dialysed, using 1000 MW cut off tubing, against a sample buffer of 10 mM Tris.HCl and 0.1 mM ethylenediamine tetraacetic acid (EDTA) (2 x 3l) for 24h.

The protein solution was then dialysed against sample buffer with 0.6% sucrose (2 x 3l) for 24h and finally against a 1.6% sucrose in sample buffer solution until the protein solution had concentrated to approximately 3 ml. The protein solution was separated into Eppendorf tubes and centrifuged (MSE Microcentrifuge, Kempthorne Medical supplies NZ) at 13500 rpm for 10 min to separate the solid material that precipitated out in the dialysis tube.

SDS-PAGE was run on the four solu ble protein solutions and the dissolved solid material isolated from the centrifugation. This showed that the second peak in both gel filtrations gave protein of the correct molecular weight, while the first peak contained material that only faintly stained in Coomassie. The solid material was the aggregated high molecular weight material which caused the streaking in the lane of the previous gels of the 18 kDa protein.

The bulk of the crude 90 amino acid peptide, which had been stored at 4°C, was then gel filtered on a Sephadex G-50 column (46 x 370 mm) in 6 M guanidine.HCl and 10 mM Tris.HCl at pH7 at 0.3 ml/min. The protein peak was trapped and stored at 4°C. A 5 ml sample of the crude 90 amino acid peptide was treated as described above except the dialysis samples were concentrated by ultrafiltration through a 1000 MW cut off

membrane then centrifuged. The polyacrylamide (15%) SDS gel of these samples gave the same result as before. Combining the two runs and lyophilising the protein solution gave 3.2 mg of protein some of which could still be salt.

The associated yellow colour present in the previous attempts at purifying the 18 kDa protein was present with the purified 90 amino acid protein after the first gel filtration. However it appears at this stage to be effectively separated from the protein during the HPLC or the second gel filtration, or could be associated with the material which precipitated during the dialysis of the pure protein solutions.

Mass spectrometry on this sample showed two key components at 6586 and 4172 mass units. After carrying out analytical HPLC and trapping the protein peak, sequencing revealed 3 fragments, Arg^{59}-Ala^{148}, Arg^{59}-Phe^{87} and Asn^{88}-Ala^{148}, of which the last fragment constituted approximately 25% of the total material. SDS-PAGE run on on this sample confirmed the presence of the 3 fragments. The protein was gel filtered on the Superose 12 column with 10 mM Tris.HCl pH 8, using a Pharmacia 214nm detector at a flow rate of 0.5 ml/min. Three fractions all containing varying quantities of 90 amino acid peptide material were isolated as shown by SDS-PAGE.

Time course amino acid analysis on the purified 90 amino acid peptide was carried out in an effort to accurately determine the ratio of the amino acids

3.2.3.9. Cleavage #6

See flow diagram for the purification scheme (figure p.4 in the pouch at the back of the thesis).

The protein resin was pretreated with thiophenol to remove the DNP groups [134] after which the Boc group was removed as described previously. The formyl group on tryptophan was removed by preswelling the protein resin in DMF then adding piperidine to give a 10% solution. After stirring at 0°C for 2h the resin was washed with DMF, DCM and methanol, dried and placed in a dessicator overnight.

The Lo-Hi HF cleavage was performed as described previously using *p* -cresol (800μl) and *p* -thiocresol (200 μl), DMS (6.5 ml) and HF (2.5 ml) at -5-0°C for 1.5h. After the evaporation of the HF and DMS the resin was washed as before, then returned to the reaction vessel. The Hi HF step with *p* -cresol (800μl) and *p* -thiocresol (200 μl) and HF (10 ml) was stirred at -5-0°C then evaporated before being washed with ether. The protein was dissolved in 6M guanidine.HCl and 10 mM Tris.HCl at pH7 (2 x 25

ml and 2 x12 ml). Half of this protein solution was stored at 4°C temporaily. Half of this solution was gel filtered immediately on a Sephadex G-50 column (46 x 370 mm) at 0.3 ml/min to give three fractions. The first fraction was loaded immediately on to the preparative HPLC and the small protein peak cut into two fractions which after lyophilisation gave a total of 1 mg of material.

The other two fractions (45 ml) were combined, as the analytical HPLC traces showed very little difference between the two fractions and were concentrated to 10 ml by ultrafiltration using a 1000 MW cut-off membrane. Two 5 ml portions of this sample were injected onto the preparative HPLC, where once the solvent peaks of salt had been eluted, a gradient of 0-60% B over 60 min at a flow rate of 2 ml/min was applied to the column. The eluted protein peaks were divided into two fractions across the top of the peak and the first and second fractions combined separately between the runs and lyophilised. The combined first fraction was then dissolved in 6M guanidine.HCl and 10 mM Tris.HCl at pH7 and gel filtered on a Sephadex G-50 column (15 x 300 mm). This gave a single peak which was divided into 3 fractions. These fractions were dialysed, using 1000 MW cut off tubing, against sample buffer (10 mM Tris.HCl 1 mM EDTA pH 8, 2 x 2l) for 24 h then against a 1.6% sucrose in sample buffer solution (2 x 2l) for 24 h. The concentrated samples were collected and centifuged.Samples of the supernatent from the fractions were removed, and loaded on SDS-PAGE gel.

The second HPLC combined fraction was also gel filtered on a Sephadex G-100 column to give four fractions. The samples were dialysed slowly down the salt gradient from 6M guanidine.HCl and 10 mM Tris.HCl at pH7 to 4M guanidine.HCl and 10 mM Tris.HCl at pH7 (1 x 1l) for 24 h to 2M guanidine.HCl and 10 mM Tris.HCl at pH7 (1 x 1l) 24h and finally against sample buffer (10 mM Tris.HCl 1 mM EDTA pH 8, 4x 2l) for 24h. The dialysed samples were then concentrated to 1-2 mls by ultra filtration using a 10000 MW membrane, as no precipitate was present in the dialysis tubes in this dialysis or the previous dialyses of sample one from the HPLC. A sample of the retentate was then removed and run on SDS-PAGE.

The samples for each lane were then hydrolysed in an attempt to determine what was causing the streaking pattern in each lane. A time course AAA at 24, 48, and 72h was carried out

One further sample of stored protein solution from the initial gel filtration was subject to preparative HPLC to give two fractions and lyophilised as before. These samples were then gel filtered on a Superose 12 column using 6M guanidine.HCl and 10 mM Tris.HCl at pH7 at a flow rate of 0.3 ml/min. Two fractions were isolated from

both gel filtrations and these were dialysed as before for 48h per step with the final dialysis for 2x 24 h then 2 x 12h. The dialysed samples were then concentrated using a 10000 MW membrane to concentrate the samples to around 2-3 ml. These were checked on SDS-PAGE

The first peak from the gel filtrations was filtered through a 10000 MW membrane with three water washes (50 ml) to remove the guanidine.HCl and then lyophilised. After treatment with the Yajima hard acid soft base cleavage method[184], as described previously (cleavage 3), the protein was gel filtered on the Superose 12 column. The protein peak was trapped as two fractions which were then subject to preparative HPLC using the analytical C-4 Vydac column with the protein peak being trapped and checked on SDS-PAGE, which showed that protein was detectable over the entire lane as before.

3.2.3.10. Cleavage #4 Samples Purified Using the Methods Developed in Cleavage #6

See figure P.1 for flowchart for material from cleavage 4.

The gel filtered crude protein fractions from cleavage 4 were dissolved in 6M guanidine.HCl and 10 mM Tris.HCl at pH7 and gel filtered on the Superose 12 column at 0.3 ml/min. The second fraction from the gel filtration gave one major broad peak which was divided in two fractions. Preparative HPLC was carried out on these two fractions with elution of the solvent peak, before a gradient of 0-40% B in 1h at a flow rate of 2 ml/min was run. The resulting two protein peaks were trapped and lyophilised. These two peaks were then checked by SDS-PAGE, which showed that the protein had separated into the aggregated material and the 18 kDa product. The same procedure was repeated for the first fraction from the original gel filtration with the same result.

3.2.3.11. Western Blots of the Synthetic 18kDa Protein.

The Western blots were done courtesy of R. L. Prestidge at the Auckland Medical School. Unless otherwise indicated all reactions were carried out at room temperature.

The protein samples were made up in reducing sample buffer, boiled for 2 min and applied to a 12% SDS-Page mini gel (Bio Rad Mini gel format.) The gels were run at 15mA each for 75 min and then western blotted onto nitrocellulose at 100 mA/26V overnight. The blots were blocked in TBS/10% fetal calf serum for 1h and rinsed in

TBS. Incubations with primary antibodies were done at the following dilutions in TBS/0.5% Tween 20 for 1h.

> α1-20 serum extracted from B10.BR mice 1:10
> α16-35 serum extracted from B10.BR mice 1:10
> α76-95 serum extracted from Balb/c x Balb.B F1 mice 1:20
> L5 acites 1:100
> α(RLP-23-thyroglobulin conjugate) serum extracted from Balb/c x C57 B1/10 F1 1:10

Blots were washed three times in TBS/Tween for 5 min each and then incubated with biotin-horse anti mouse IgG 1:200 for 1h and washed as above. After incubation with streptavidin-biotin horse radish peroxidase 1:400 in TBS for 1h the blots were washed with TBS 3 x 10 min then developed with 30mg 4-chloro-1-napthol /10ml ethanol, 40 ml 50 mM Tris.HCl pH 6.8 and 50 μl H_2O_2 (30%) for 10-60 min. The blots were then washed with water and dried.

Some pieces of nitrocellulose were stained for total protein with indian ink by the method of Glenny[186].

The T cell assay is as described in Chapter 2 for the T-cell proliferation assay of the peptides.

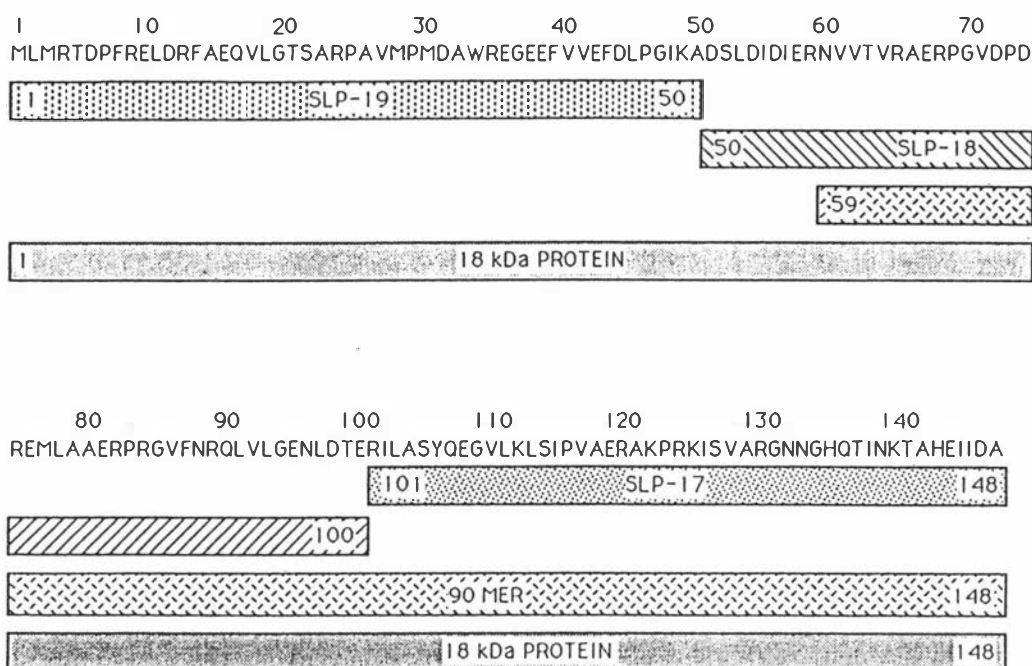


Figure 3.3.1. Protein sequence showing the peptides 1-50, 50-100, 101-148, the 90 amino acid fragment 59-148 and the full length 148 amino acid 18 kDa protein synthesised in this chapter.

Amino Acid	Expected	Peptide Resin	F1	F2	F3
Asp	5	9.27	7.50	5.95	5.84
Thr	2	2.70	3.10	2.26	2.20
Ser	3	1.41	1.62	2.71	2.80
Glx	5	5.35	4.63	5.2	5.43
Pro	2	1.57	1.45	1.90	2.04
Gly	3	3.00	3.00	3.00	3.00
Ala	5	7.69	5.23	5.19	5.44
Val	3	2.43	1.91	2.56	2.65
Ile	6	8.90	4.72	4.07	4.32
Leu	3	1.63	0.69	2.06	2.55
Tyr	1	0.33	0.23	0.72	0.91
His	2	ND	2.79	2.13	2.02
Lys	4	4.04	3.34	3.85	3.93
Arg	4	3.11	3.22	3.81	4.01

Table 3.3.1. Amino acid analysis table for peptide 101-148 synthesis #2. For the peptide resin sample His is not detected due to DNP-His.

3.3 RESULTS AND DISCUSSION.

Figure 3.3.1 shows the peptides synthesised in this Chapter. Peptides 1-50, 50-100, and 101-148, were synthesised as separate peptides. Protein fragment 59-148 was isolated from the total protein synthesis by removal of some peptide resin from the reaction vessel. Protein purification flow charts, the protected amino acid side chain sequence and the protein sequence are in the pouch at the back of the thesis.

3.3.1 SYNTHESIS of LARGE PEPTIDES SLP-17, SLP-18 and SLP-19

3.3.1.1. C-terminal Peptide SLP-17, 101-148.

The initial synthesis of this peptide failed due to a mixup in the solvents with a DCM bottle being misplaced in a DMF bottle position. As a result, from cycle 5 onwards, the yields drop dramatically from 99.5% to 95% until finally there were four couplings during cycles 15-18 that averaged 70%. Despite changing the DMF bottle back at cycle 17, the coupling yields averaged around 97% for the rest of the synthesis. In hindsight, to prevent wastage the synthesis should have been aborted at cycle 18. Amino acid analysis on the peptide resin showed that the amino acid ratios bear no relation to the expected ratios. The earlier lower % couplings from cycle 5-15 could well have been as a result of the DNP group affecting the ninhydrin assay, as shown in the synthesis of the peptides SLP-13 and SLP-14. No attempt was made to cleave and purify this peptide.

The next synthesis of this peptide proceeded smoothly with a weight gain for the resin of 854 mg (plus 182 mg of peptide resin samples) compared with with an expected yield of 1.55 g. The coupling yields for the amino acids averaged 98.97% with an overall expected yield of 60.52% (Figure 3.3.2). AAA of the peptide resin showed most of the amino acids to be in the correct ratios for the peptide (Table 3.3.1).

Prior to the HF cleavage the DNP group and the N-terminal Boc group were removed using the β-Mecaptoethanol method, and TFA respectivly as described earlier in Chapter 2. A Lo-Hi HF cleavage was carried out on 300mg of peptide resin to give 78.4 mg of crude peptide , which was considerably lower than the expected yield of 230mg. Analytical HPLC of the crude lyophilised peptide (Figure 3.3.3) showed the presence of two major broad peaks. Preparative HPLC was carried out on the crude sample to give 3 combined fractions of a total of 23.4 mg. AAA (Table 3.3.1) on each combined fraction

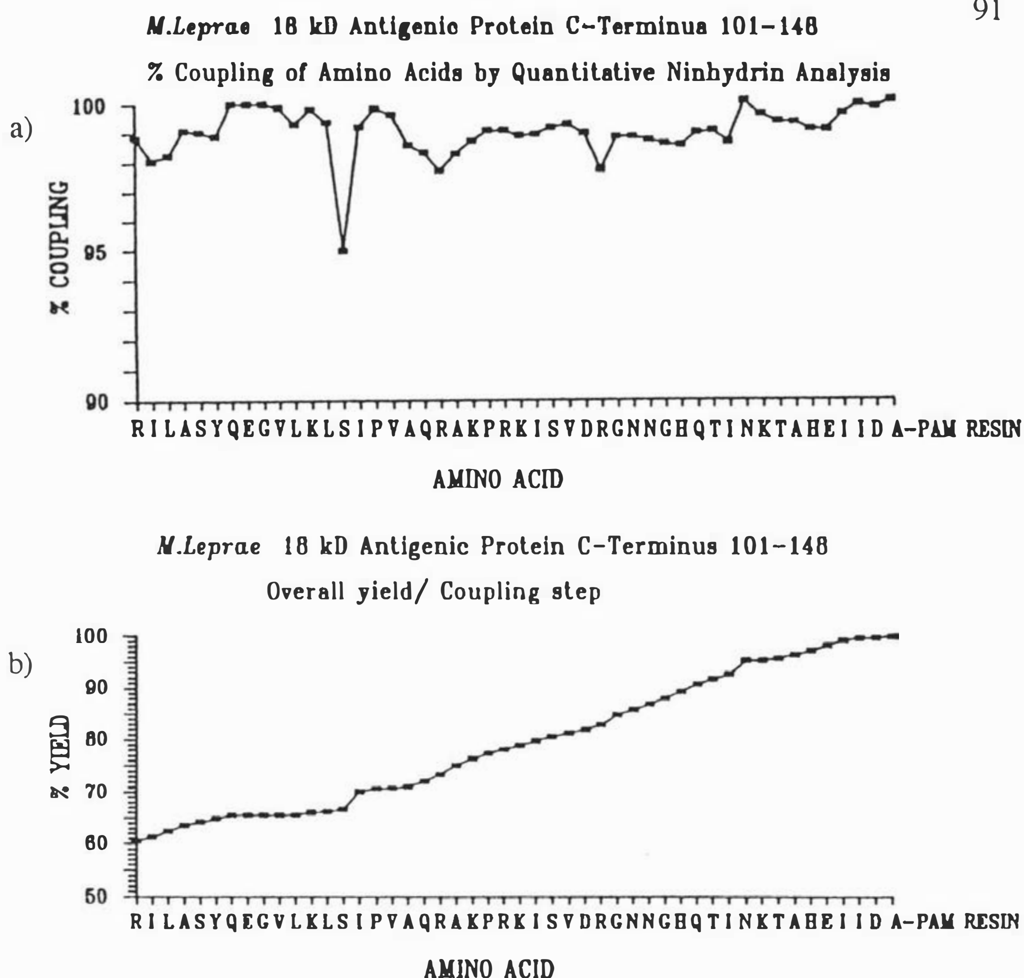


Figure 3.3.2. Ninhydrin analysis for peptide 101-148 synthesis #2, a) % coupling yields for each amino acid. b) Cumulative yield for the peptide on the resin.

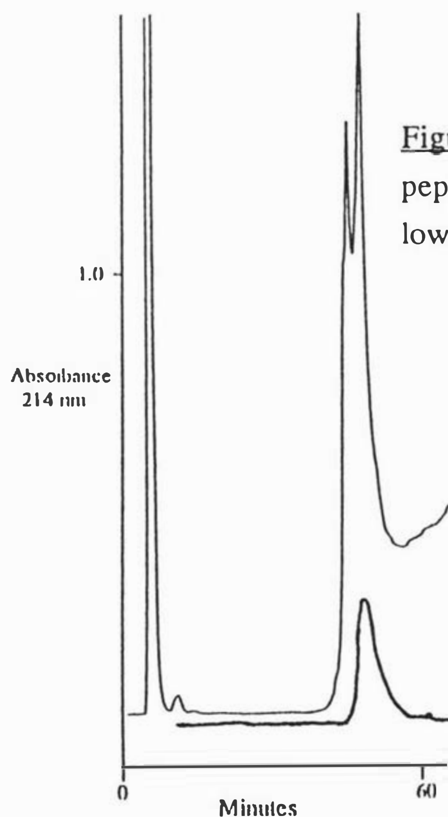


Figure 3.3.3. Analytical HPLC of the crude peptide 101-148 from synthesis 2 with the lower trace being the 280 nm profile.

Amino Acid	Expected	F1	F2	F3	F4
Asx	5	5.06	4.98	4.90	4.70
Thr	2	2.00	1.94	1.98	1.80
Ser	3	2.30	2.63	2.62	2.51
Glx	5	4.95	5.03	4.84	4.85
Pro	2	1.87	1.88	1.98	1.86
Gly	3	3.00	3.00	3.00	3.00
Ala	5	4.51	4.96	4.79	4.76
Val	3	2.80	2.83	2.89	2.82
Ile	6	4.27	4.65	4.61	4.70
Leu	3	2.26	2.77	2.70	2.55
Tyr	1	0.56	0.98	0.79	0.89
His	2	1.92	1.98	1.58	1.41
Lys	4	3.72	3.80	3.69	3.67
Arg	4	3.64	3.80	3.47	3.76

Table 3.3.2. Peptide 101-148 isolated from the total protein synthesis. F1-F4 refers to the combined purified fractions from the preparative HPLC

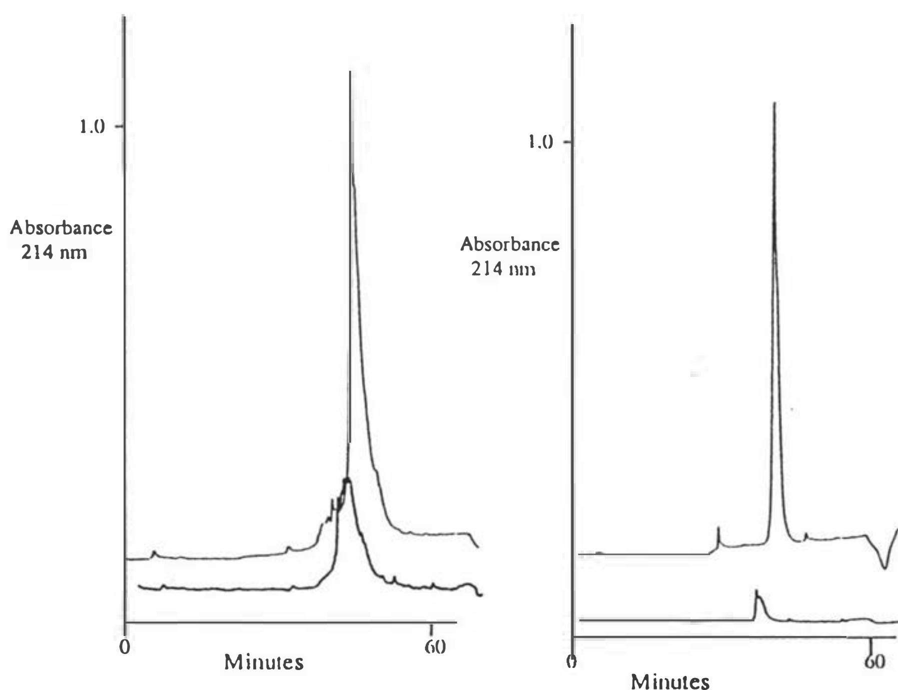


Figure 3.3.4. Analytical HPLC of a) crude and b) purified peptide 101-148 isolated from the peptide resin removed at position Arg 101 during the total synthesis of the 148 amino acid 18 kDa protein synthesis. The lower trace is the 280 nm profile.

showed that the first fraction had no correlation with the expected amino acid ratios but correlates well with a 17 amino acid fragment Asn^{141}-Ala^{148}. For the other two fractions (F2-F3) the amino acid ratios agree with those expected apart from Asx which is consistently high. Ion exchange chromatography showed that the purest peptide was occurring in the later eluting fractions from the preparative HPLC.

The purest fraction was sequenced in order to determine the presence of any protecting groups still attached to the peptide and to verify the peptide structure. The results of the sequence analysis showed that there had been a mistake made in the programming of the synthesiser and a Gln had been substituted for a Glu at position 119 or cycle number 29. This error is not detected in the AAA due to the AAA detection of Glx instead of the individual amino acids glutamine and glutamic acid. This error had occurred despite rigorous checks prior to the start of the synthesis. Therefore no further studies were made with this peptide.

Authentic synthetic 101-148 peptide was obtained from the total synthesis of the 18kDa protein, where a 326 mg sample of the peptide resin was removed at cycle 47 Arg^{101}. The average coupling of the peptide was 98.81% with a predicted yield of the peptide on the resin of 56.92%. However, with on resin sequencing (see section 3.3.2.1), the average coupling % was 99.54% with an overall yield of 76.7% predicted for this stage of the synthesis. The resin sequencing also confirmed the correct sequence of amino acids over the entire peptide and AAA of the peptide resin showed that the amino acids were present in the correct ratios.

The peptide resin was cleaved using the Lo-Hi HF method as described earlier to give 97 mg of crude peptide, still considerably down on the predicted yield of peptide from the resin. Preparative HPLC was carried out on 47 mg of the crude peptide. Four fractions were isolated on the basis of their retention times on the analytical HPLC and combined to give 28.3 mg of purified peptide. Analytical HPLC on these combined fractions indicated only a single peak (Figure 3.3.4). A sample of one of these fractions was sequenced for 26 residues to check the structure of the peptide to ensure that no protecting groups were present. SDS-PAGE showed that the purified peptide exhibited a band spread from 15 kDa with the major concentration of peptide ranging from 10 kDa to just below 5 kDa (Figure 3.3.5, lane 4). This suggests that the peptide is existing as mainly a dimer with some trimer formation which is stron gly held together. AAA on the purified peptide showed that the synthetic peptide ratios agreed with the expected ratios (Table 3.3.2).

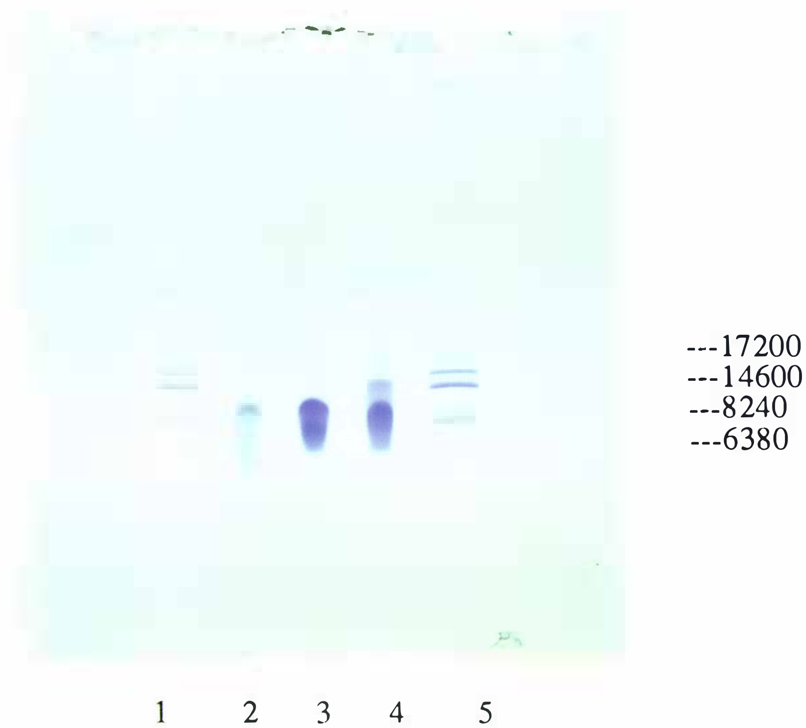
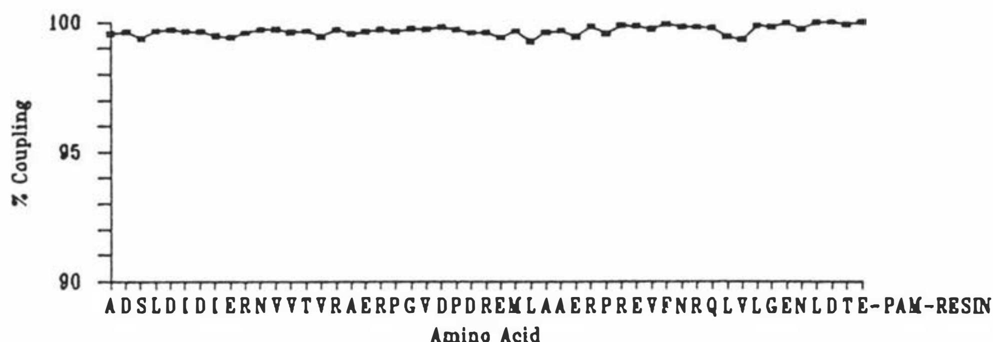


Figure 3.3.5. SDS-PAGE gel showing the HPLC purified 50 amino acid peptides. Lane: 1 MW standards 17200, 14600, 8240, and 6380 as stated in the experimental. Lane 2: 1-50, Lane 3: 50-101, Lane 4: 101-148. Lane 5: MW standards.



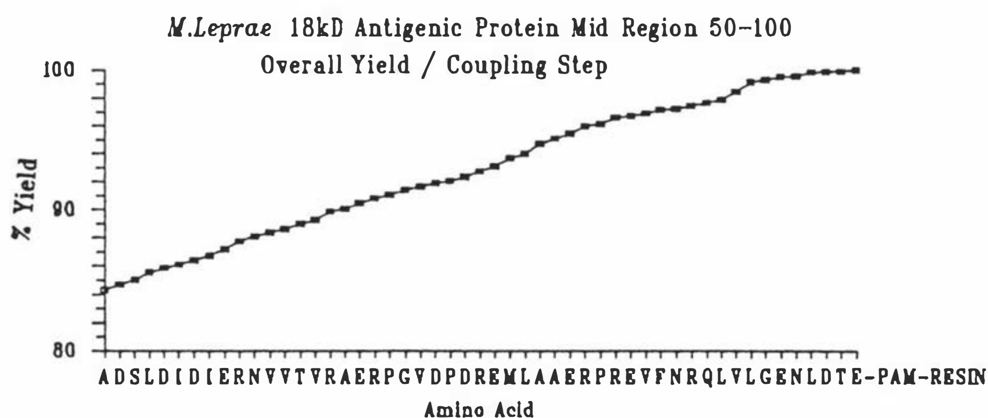


Figure 3.3.6. Ninhydrin analysis for peptide 50-100, a) % coupling yields for the amino acids, b) cumulative yield for the peptide on the resin.

Amino Acid	Expected	Peptide Resin	F1	F2	F3	F4
Asx	9	7.90	8.96	9.13	9.31	9.26
Thr	2	1.86	2.16	2.06	2.00	2.12
Ser	1	ND	0.52	0.75	0.83	0.76
Glx	7	7.27	8.24	7.59	7.55	7.47
Pro	3	2.85	3.56	3.36	3.33	3.53
Gly	3	3.00	2.45	3.07	3.05	2.25
Ala	4	3.36	4.00	4.00	4.00	4.00
Val	6	5.55	4.75	5.50	4.25	4.79
Met	1	0.42	1.13	0.86	1.01	1.03
Ile	2	1.31	1.43	1.77	1.71	1.72
Leu	5	4.88	5.35	5.66	5.57	5.52
Phe	1	0.97	0.86	1.03	0.76	0.89
Arg	7	7.22	8.60	8.26	8.04	8.17

Table 3.3.3. Peptide 50-100 F1-F4 refers to the combined purified fractions from the preparative HPLC. ND, not detected.

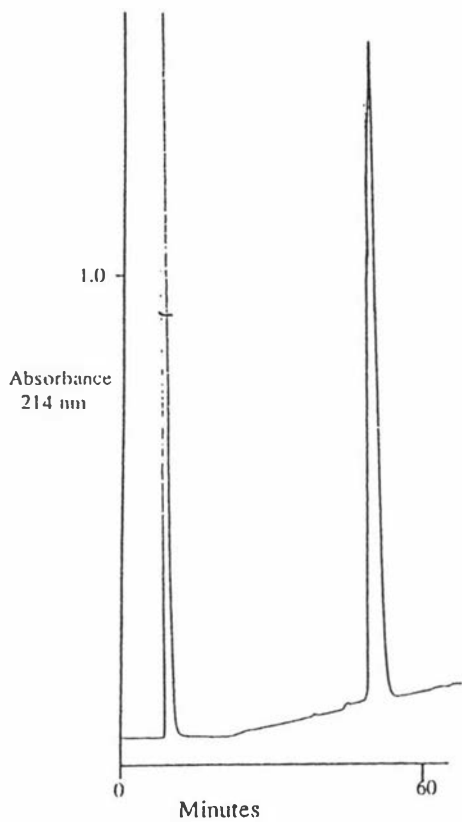


Figure 3.3.7. Analytical HPLC of purified 50-100.

3.3.1.2. Mid Region Peptide SLP-18, 50-100.

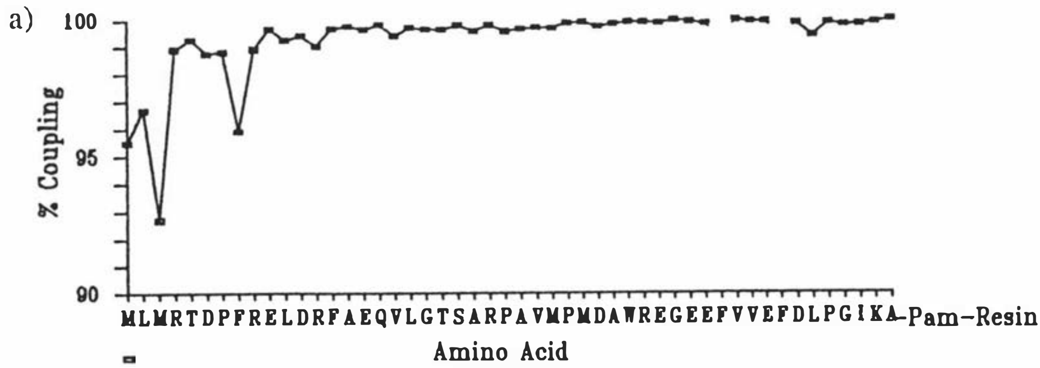
The synthesis of the mid region 51mer 50-100 proceeded with no difficulties. The lack of any special protecting groups meant that the only potential problem was the Asp-Pro bond, which is susceptible to cleavage in dilute acid solutions (discussed in detail later in this section). The synthesis was carried out on a 0.4 mMol scale with average couplings of 99.66% with an overall predicted yield for the peptide of 84.16% (Figure 3.3.6). AAA of the peptide resin showed that the synthetic peptide on the resin had the expected amino acid ratios (Table 3.3.3).

After the removal of the N-terminal Boc group, an ordinary single step HF cleavage (90% HF) was carried out on 300 mg of peptide resin to give 79.4 mg of crude peptide. The peptide was very difficult to dissolve, probably due to the large number of hydrophobic groups on this peptide, which also encompases the troublesome region of peptide SLP-8 (46-65). Preparative HPLC on 74 mg of the crude peptide gave a total of 35 mg of purified peptide. The fractions from across the peptide peak were combined on the basis of peptide retention time on the analytical HPLC (Figure 3.3.7). Analytical ion exchange chromatography of the combined fractions showed that the fractions that contained most of the peptide were the later eluting fractions on the HPLC. AAA on the peptide fractions showed that all the amino acids were present in the correct ratios (Table 3.3.3). The best fraction (F 4 in Table 3.3.3) as detected by ion exchange chromatography was sequenced for 38 residues to determine the correct formation of the peptide structure and to demonstrate the lack of any protecting groups attached to the peptide. This fraction was also run on SDS-PAGE gel to determine its purity (Figure 3.3.5, lane 3). The peptide showed a similar pattern to peptide SLP-17 but without the presence of the trimer material.

3.3.1.3. N-terminal Peptide SLP-19, 1-50.

This peptide was the first of the 3 larger peptides synthesised, and care had to be taken due to the tryptophan and the N-terminal Asp-Pro bond. The peptide was synthesised on a 0.3 mMol scale and on completion of the synthesis the peptide resin weighed 1.40 g. For both phenylalanine amino acids there was solvent present in the cartridges and blockage of the in-line filter for the amino acid injector,probably due to the amino acid not dissolving completely (see Chapter 2, same batch of Vega phenylalanine). After changing the in-line filter, the synthesis procceded smoothly until cycle 40 where the DIEA ran out and was replaced with triethylamine (TEA). It is noticable from the ninhydrins that there is a considerable drop off in the coupling efficency over the last few amino acids after the

MASSEY UNIVERSITY
LIBRARY



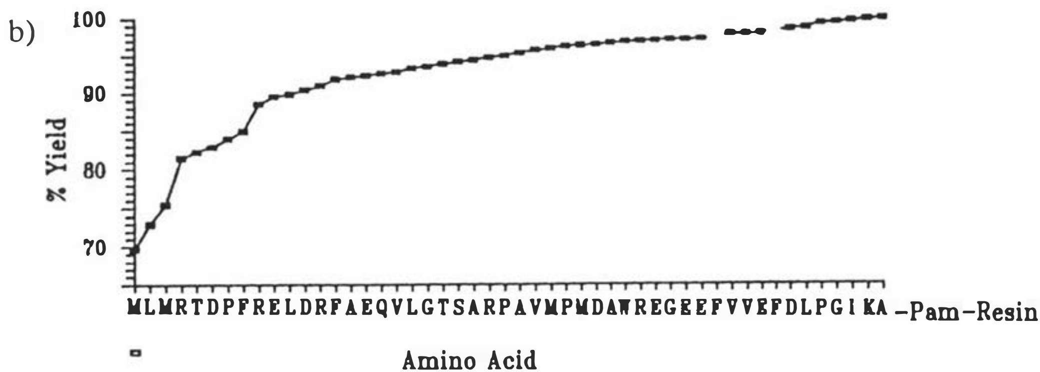


Figure 3.3.8. Ninhydrin analysis for peptide 1-50, a) % coupling yield for the amino acids, b) cumulative yield for the peptide on the resin. There are two missing samples for the phenylalanines at position 39 and 43. The decrease in the coupling yield at the N-terminal is due to the change in neutralising solvent from DIEA to TEA.

Amino Acid	Expected	Peptide Resin	F1	F2	F3
Asx	4	3.62	3.43	3.64	3.6
Thr	2	1.51	1.58	1.67	1.7
Ser	1	0.47	0.94	0.97	1.02
Glx	7	6.56	6.53	6.61	6.32
Pro	4	3.45	1.74	1.79	4.74
Gly	3	2.95	2.94	2.98	2.91
Ala	5	5.00	5.00	5.00	5.00
Val	4	3.90	2.75	2.55	2.19
Met	4	1.72	3.01	3.16	2.94
Ile	1	1.30	0.70	0.69	0.76
Leu	4	3.37	3.22	3.35	3.06
Phe	4	2.52	2.35	2.45	2.87
Lys	1	1.55	0.73	0.69	0.74
Arg	5	4.41	4.70	4.90	5.71

Table 3.3.4. Amino acid Analyses for peptide 1-50 SLP-19. Low Val content due to Val-Val and Val -Leu. F1, F2 and F3 refer to the combined fractions from the preparative HPLC. Tryptophan (1) not detected.

change of the neutralising base reagent (Figure 3.3.8). Unfortunately there were no peptide resin samples to enable the determination of the coupling yield for the two phenylalanines at position 39 and 43, probably as a result of there being very little activated amino acid solution in the reaction vessel. By excluding the probable failed coupling of phenylalanines 39 and 43 and the drop off as a result of the TEA the expected yield of the peptide would have been approximately 85%.

From the AAA of the peptide resin (Table 3.3.4) it can be seen that the value for the phenylalanine is low, indicating incomplete coupling occurred for the first two phenylalanines. The indications are from the ninhydrins and AAA are that the two phenylalanines at the position 8 and 14 coupled as expected.

After the removal of the N-terminal Boc group, the peptide resin was cleaved by the Lo-Hi HF method. The crude peptide was dissolved in 6M guanidine.HCl 25 mM Tris pH8.6 and gel filtered on the Superose 12 column in 6M guanidine.HCl to remove the organic scavengers used in the HF cleavage. The isolated peptide peak was then chromatographed on the preparative HPLC and the fractions isolated from across the peptide peak (Figure 3.3.9). AAA on the purified peptide indicated that most of the amino acids were present as expected except for the phenylalanines. In this case only three phenyalanines were present compared to an expected four (Table 3.3.4).

SDS-PAGE (Figure 3.3.5, lane 2) showed a band of the correct molecular weight 5 kDa with to bands higher at 10 kDa and 15 kDa with the major amount of peptide at 15 kDa. The N-terminal of the peptide was sequenced for 20 residues to determine the peptide structure and gave no sign of any side chain protecting groups still present on the peptide.

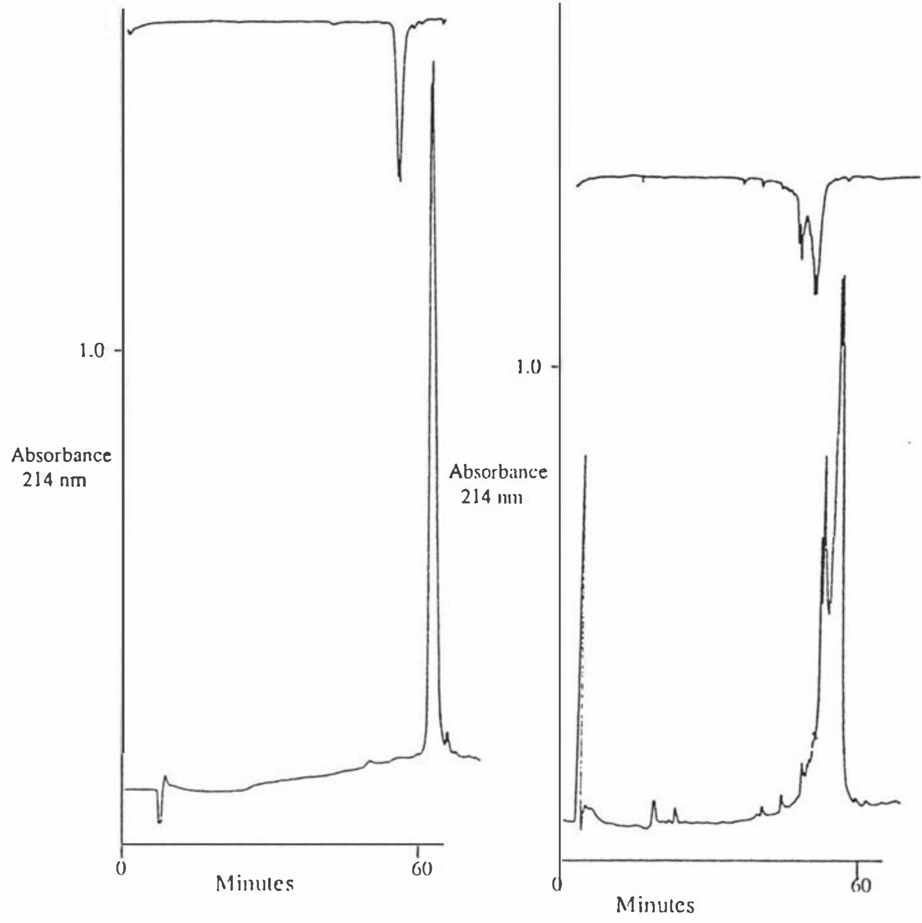
3.3.1.4. Asp-Pro Bond Acid Stability.

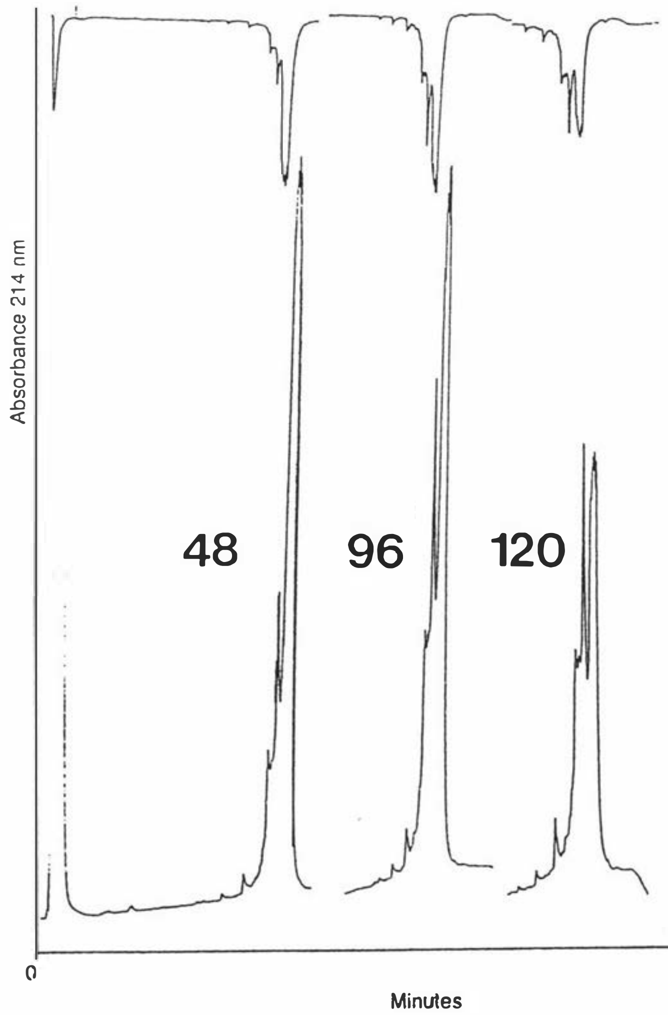
An important consideration for the total synthesis of the 18 kDa protein was the acid stability of the Asp-Pro bond. Exposure of an Asp-Pro bond for long periods of time to dilute acid concentrations can cause this bond to hydrolyse[187]. It was for this reason that the 1-50 mer peptide was dissolved in 6M guanidine rather than the dilute acetic acid normally used to dissolve the crude peptide.

In order to determine the relative rate of cleavage of the Asp-Pro bond, hydrolysis was first carried out using the method of Marcus *et.al* [187]. The purified 1-50 peptide was treated with 0.015 M HCl at 100°C for 30 min. The analytical HPLC (Figure 3.3.9) showed that the original single peptide peak had been hydrolysed to give two major broad

Figure 3.3.9. Analytical HPLC traces, a) purified peptide 1-50, b) peptide 1-50 hydrolysed with 0.015N HCl for 30 min at 100°C. The top trace is the 280 nm trace.

Figure 3.3.10. Analytical HPLC of peptide 1-50 hydrolysed for 48, 96 and 120 hours respectively with 10% acetic acid at pH 2.5 at 40°C. Top trace is 280 nm trace.


1.0
Absorbance
214 nm
0
Minutes
60
1.0
Absorbance
214 nm
0
Minutes
60


Absorbance 214 nm
48
96
120
0
Minutes

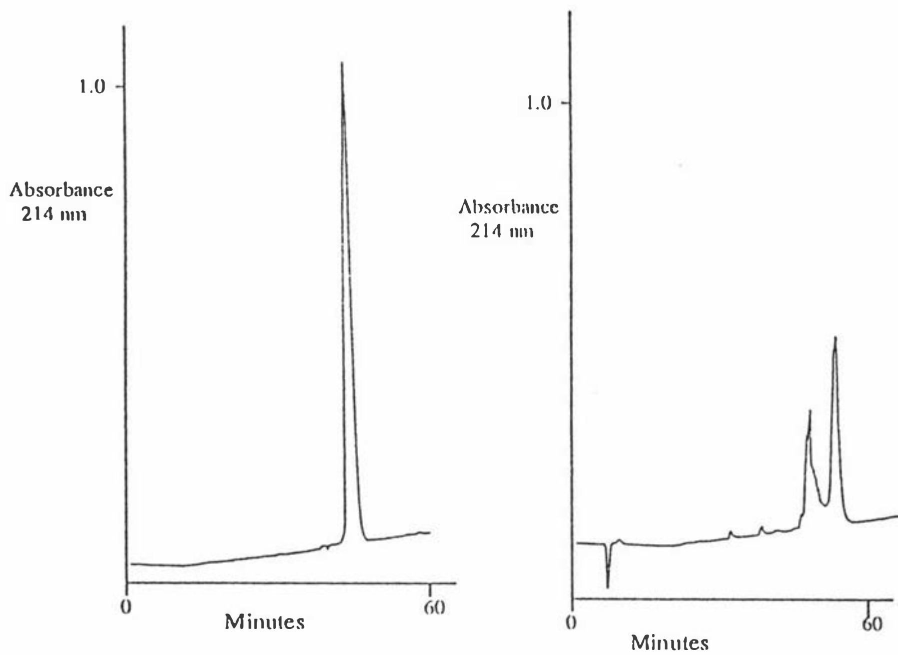


Figure 3.3.11. Analytical HPLC of purified 50-100 and of peptide 50-100 hydrolysed with 0.015N HCl for 30 min at 100°C.

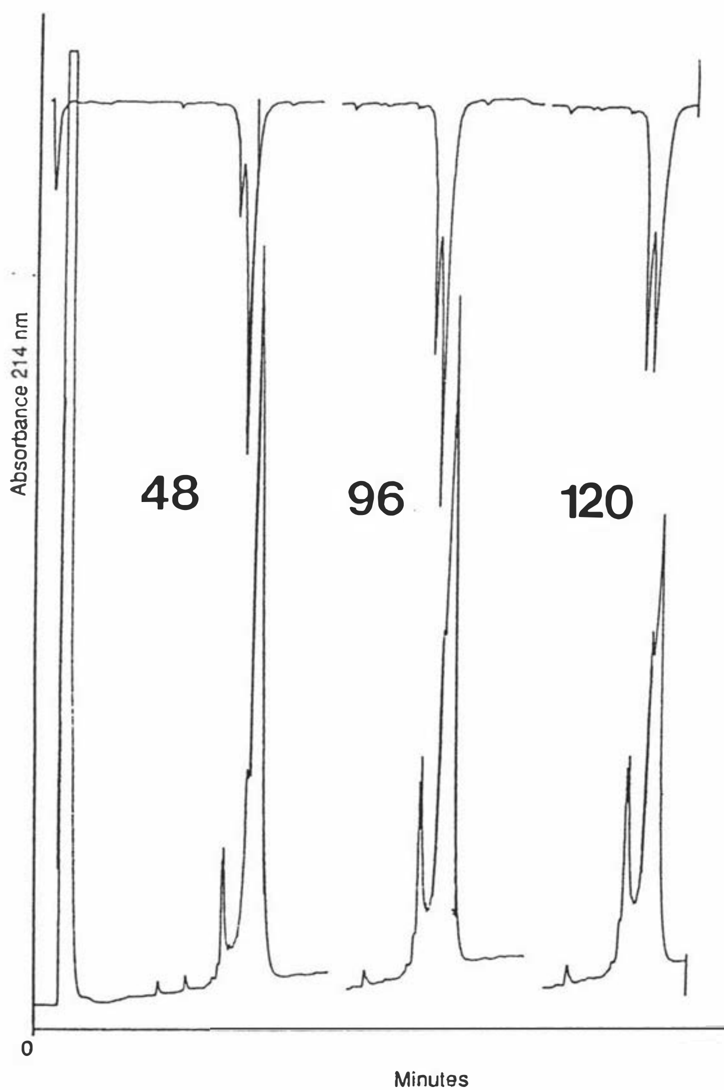


Figure 3.3.12. Analytical HPLC of peptide 50-100 hydrolysed with 10% actetic acid at pH 2.5 for 48, 96 and 120 hours respectively. The top trace is the 280 nm trace.

peaks. It was decided that the hydrolysis with dilute acetic acid could be monitored by this method. Therefore a sample of the peptide was treated with 10% acetic acid adjusted to pH 2.5 with pyridine[188] at 40^oC. Samples from the peptide hydrolysis were removed at 48, 96 and 120h and immediately injected onto the analytical HPLC (Figure 3.3.10). Based on the inte grated peak areas, there was 24% cleavage of the Asp-Pro bond at 48h, 46% cleavage at 96h and an undetermined amount at 120h, due to the presence of other peaks interfering with the integration of the peaks.

For peptide 50-100 the position of the Asp-Pro bond in the middle of the peptide chain should confer some degree of acid stability to the bond compared to the 1-50 peptide, where the Asp-Pro bond is in an exposed position. Therefore, during the HF cleavage, the crude peptide was dissolved in acetic acid and lyophilised immediately to reduce the peptides exposure to the acid. The hydrolysis with 0.015 M HCl gave two peaks on the analytical HPLC (Figure 3.3.11). The hydrolysis with acetic acid showed that the hydrolysis of the peptide bond proceeded more slowly than in the N-terminal 1-50 peptide, with 8.4%, 20.8% and 26.32% of the peptide, as determined by the intergrated peak area, being cleaved at 48,96 and 120h respectively (3.3.12)

3.3.1.5. Summary.

Despite the operator errors in these lead up syntheses of the larger peptides, the results were sufficently encouraging to carry out the 148 amino acid 18 kDa protein synthesis, without further expensive 50 amino acid peptide syntheses. The human error would be reduced in the protein synthesis, particularly in the programming of the sequence, by having a third person totally unfamiliar with the project confirm the correct sequence. The problem of the amino acids not dissolving was removed by buying the amino acids already in their cartidges from ABI.

The coupling yields for the 50 amino acid peptides, as judged by ninhydrin analysis, particularly for peptide 50-100 (99.66%), were encouraging. If the poor coupling yields for the N-terminal residues for the 1-50 peptide, were attributted to the TEA, the overall yield of the peptide would be improved considerably by using DIEA. However any conclusion on this peptide must be tempered with caution, due to the apparent failed coupling of the phenalanines at position 39 and 43. For the C-terminal peptide (101-148) the % coupling yields were a lot lower than the other two 50 amino acid peptides. This could probably be attributed to the DNP histidine interfering with the ninhydrin assay along with fact that in the second synthesis of this peptide, all solvents were purified on the premises and perhaps were not as pure as those purchased from ABI. The alternative method of sequencing the peptide resin to determine the coupling yields, was employed during the synthesis of the

protein. This will help to overcome the problem of the low ninhydrin absorbance levels that occur during the early stages of the synthesis, as a result of the DNP group.

At this stage it's the chemical yields of the peptides that is the major concern. The aggregation of the peptides is certainly a problem as seen in the SDS-PAGE gel. If these peptides are to be used for studies on their immunology, or structural propeties then further work on purifying the peptides would need to be done. For example it is hoped that soon the C-terminal or the N-terminal peptide will be the subject of an NMR study on its structure. Further purification of the peptide such as ion exchange chromatography, and gel filtration under denaturing conditions, with slow dialysis from 6M guanidine.HCl will hopefully unfold the protein and allow it to refold correctly.

The Asp-Pro hydrolysis experiments indicated that the N-terminal Asp-Pro bond was less acid stable than the Asp-Pro bond in peptide 50-101. Provided the protein was not exposed to dilute acetic acid for more than 20 hrs at room temperature, the rate of cleavage of this bond should be minimised.

The resin samples must be regularly monitored with the ninhydrin assay in the early stages of the total synthesis, so that at the first sign of a failed coupling the synthesis can be terminated, without too much loss. A comparison with the ninhydrin analysis of the second 101-148 peptide synthesis was deemed essential to ensure the results were equal or better than the peptides.

PEPTIDE RESINS							
Amino Acid	128-148	106-148	87-148	66-148	46-148	26-148	1-148
Asx	5.16 (5)	6.09 (5)	8.77 (8)	11.44 (10)	15.54 (14)	17.80 (16)	18.56 (18)
Thr	1.86 (2)	2.15 (2)	3.07 (3)	3.60 (3)	5.08 (4)	5.48 (4)	6.67 (6)
Ser		1.07 (2)	1.43 (3)	1.56 (3)	2.22 (4)	2.72 (4)	2.72 (5)
Glx	2.30 (2)	5.02 (5)	7.53 (7)	10.68 (11)	12.28 (12)	15.98 (16)	18.73 (19)
Pro		2.01 (2)	2.17 (2)	4.49 (5)	4.84 (5)	6.39 (7)	7.88 (9)
Gly	2.00 (2)	3.00	4.00	6.00	7.00	8.00	9.00
Ala	2.47 (2)	4.50 (4)	5.68 (5)	8.64 (8)	9.90 (9)	11.91 (11)	13.91 (13)
Val	0.94 (1)	2.79 (3)	3.75 (4)	5.53 (6)	7.44 (9)	9.37 (12)	11.73 (13)
Met				0.37 (1)	0.35 (1)	1.08 (3)	2.15 (5)
Ile	2.83 (3)	4.99 (5)	5.84 (6)	6.80 (6)	8.45 (9)	8.61 (9)	9.60 (9)
Leu		1.75 (2)	4.58 (5)	6.10 (7)	7.21 (8)	8.17 (9)	10.68 (12)
Tyr		0.65 (1)	0.63 (1)	0.74 (1)	0.82 (1)	0.91 (1)	0.60 (1)
Phe			0.58 (1)	0.67 (1)	0.66 (1)	1.97 (3)	3.27 (5)
Lys	1.22 (2)	4.07 (4)	4.35 (4)	5.03 (5)	5.85 (5)	6.14 (5)	7.80 (5)
Arg	0.86 (1)	2.91 (3)	4.33 (5)	7.55 (9)	8.74 (11)	9.80 (12)	14.47 (16)

Table 3.3.5. Peptide resin hydrolysis of samples removed during protein synthesis . 128-148 refers to the C-terminal 20 amino acids. Amino acid ratios are referenced to Glycine as Alanine is attached to the resin. Tryptophan not detected (1)

3.3.2 PROTEIN SYNTHESIS.

The final weight increase of the protein resin was 2.09 g compared to the expected weight increase of 6.64 g. If the removal of the 104 resin samples (826 mg in total), the two larger samples at the 1/3 complete and the 2/3 complete stages (326 mg and 426 mg respectively), is taken into account the final yield of the protein resin is as expected.

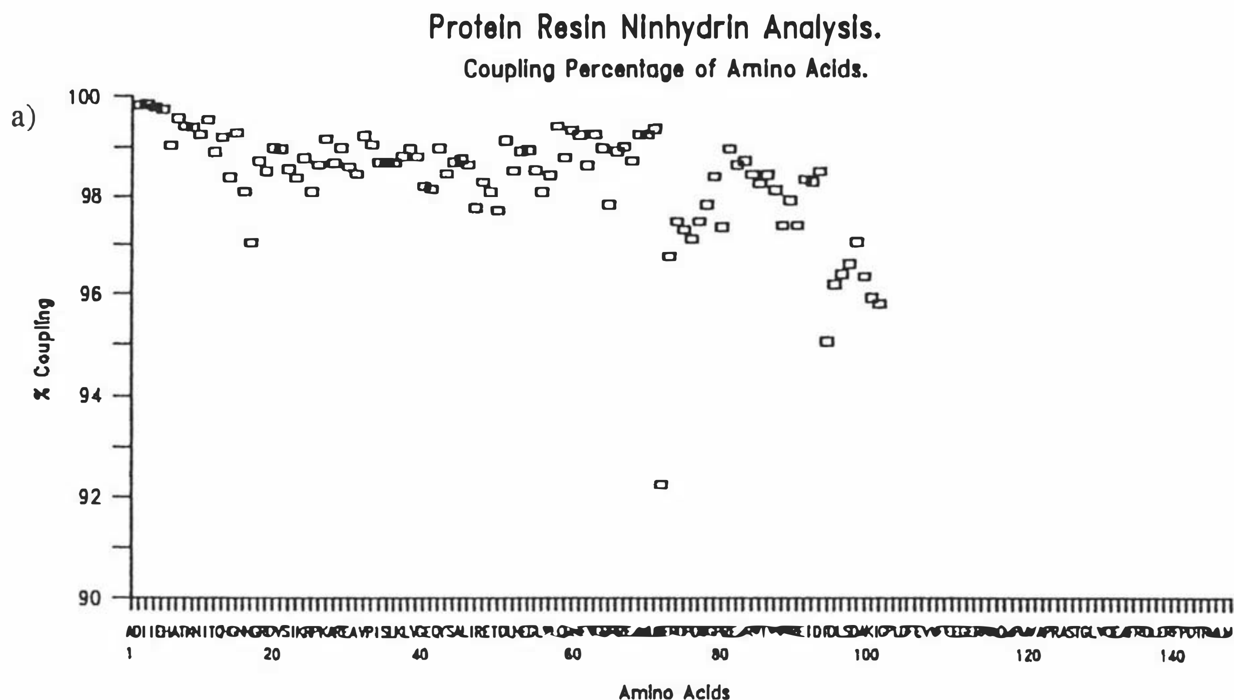
AAA of the protein resin samples (Table 3.3.5) was carried out and the molar ratios determined relative to glycine. All hydrolyses compared favourably with the expected ratios, with the Asx value being consistently one amino acid higher than expected, for the majority of the samples. The harsh conditions of the hydrolysis for the protein resin samples affect serine and methionine which give consistently lower than expected results[127]. As the DNP group was not removed from histidine prior to the hydrolysis, no determination of this amino acid could be made and no determination was made, for tryptophan either.

From the ninhydrin analysis over the first 100 amino acids, the average coupling yield per amino acid was 98.53 % (Figure 3.3.13), with the expected yield of protein for the 100 amino acid of 20.27%. This gives a predicted yield extrapolated for the total 148 amino acids of 8.3%. It was noticeable, after coupling the two His (DNP) residues that the coupling percentages from the ninhydrins dropped from the high 99% values to 97%. This drop in coupling percentages is also visible in the peptides SLP-13 and SLP-14 and the C-terminal 50 amino acid peptide 101-148. The drop off in the % coupling samples is visible after sample 70 and by the time sample 100 is reached the % coupling had dropped dramatically to the 96-95% range.

In long peptides a small background value can be seen which can be estimated from the value of the amino acid after a proline (ie proline NH is exposed), which being a secondary amine does not release a soluble chromophore, therefore any colour development is due to the indiscriminate binding to the protein chain and resin beads.[120,132c]. For the residues after a proline the coupling percentages are 98.63, 98.68, 99.00, 97.40 and 98.94%. This then means that ninhydrin assay has a background average of 1.4%, raising most coupling percentages into the high 99% range, giving the possibility that the yield of protein could be a lot higher than that calculated.

The other method for determining coupling yields uses preview analysis of the peptide resin. This method of sequencing the side chain protected protein-resin on a gas phase ABI 270A sequencer presented problems for over 1.5 years, before eventually

a)



b)

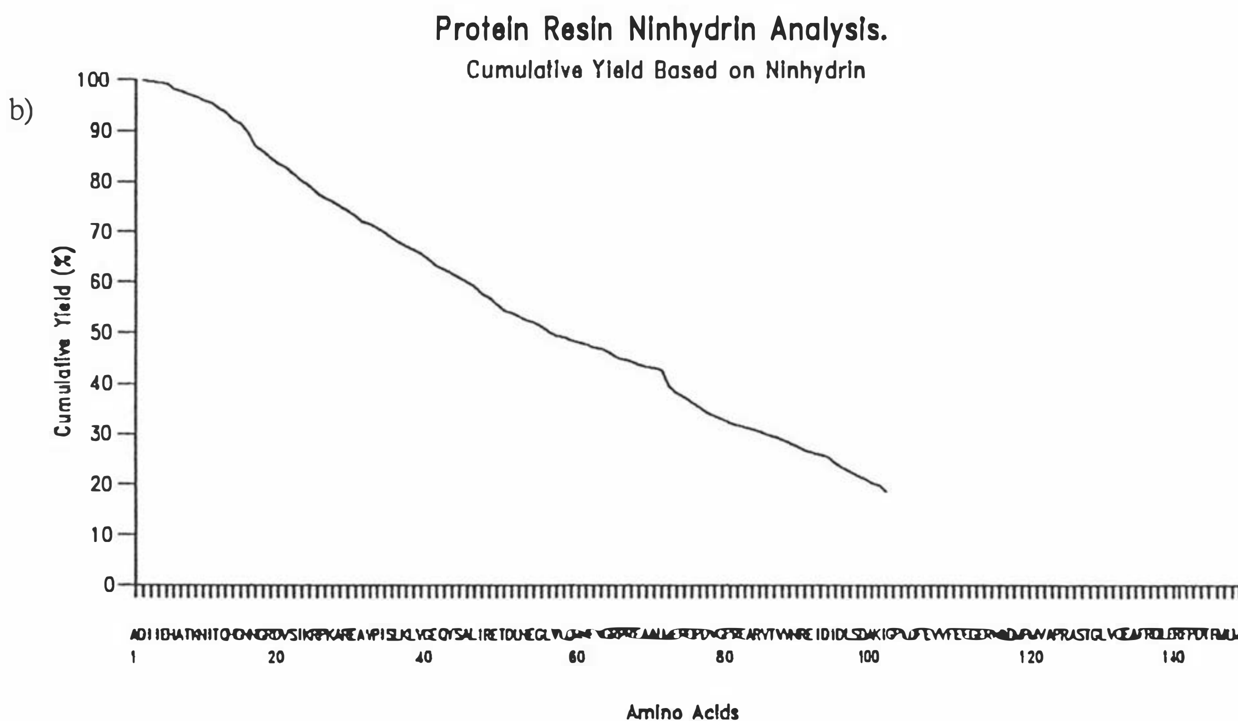
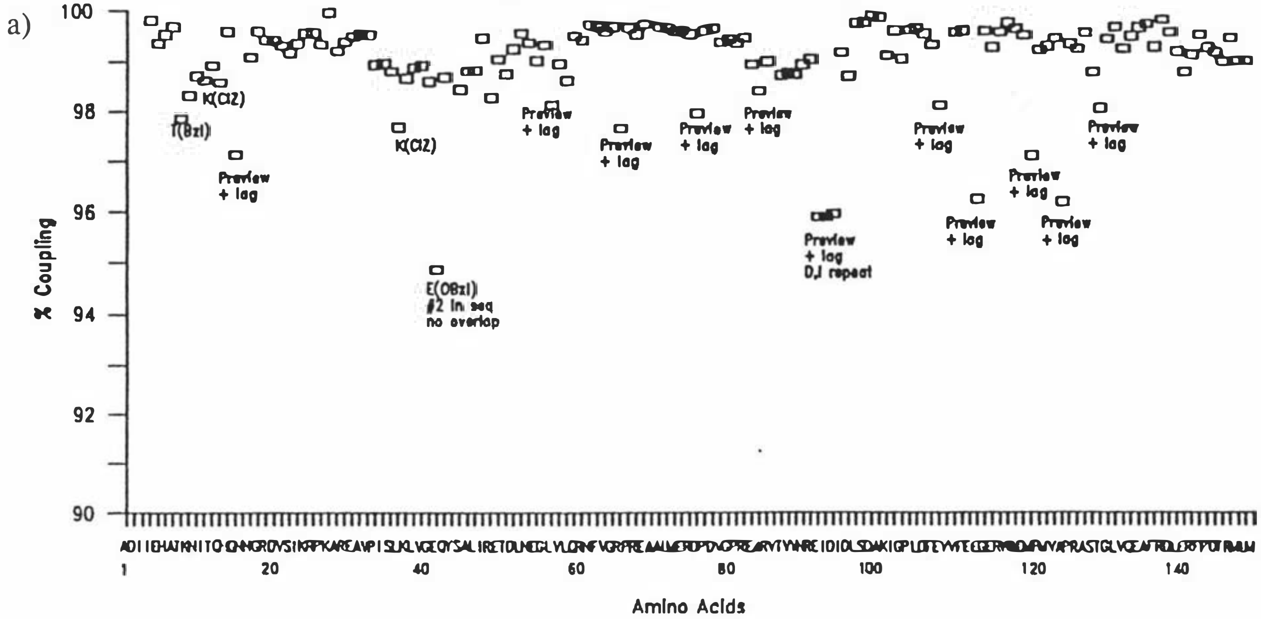


Figure 3.3.13. Ninhydrin assay for the total protein synthesis. a) % coupling yield for the amino acids b) cumulative yield for the protein on the resin calculated from the % coupling yields for the amino acids. There is a decrease in the avereage % couplings for the amino acids from 70-100. From the first histidine DNP amino acid the % coupling decreases to average 97-99% instead of the expected 99-100%.

a)



b)

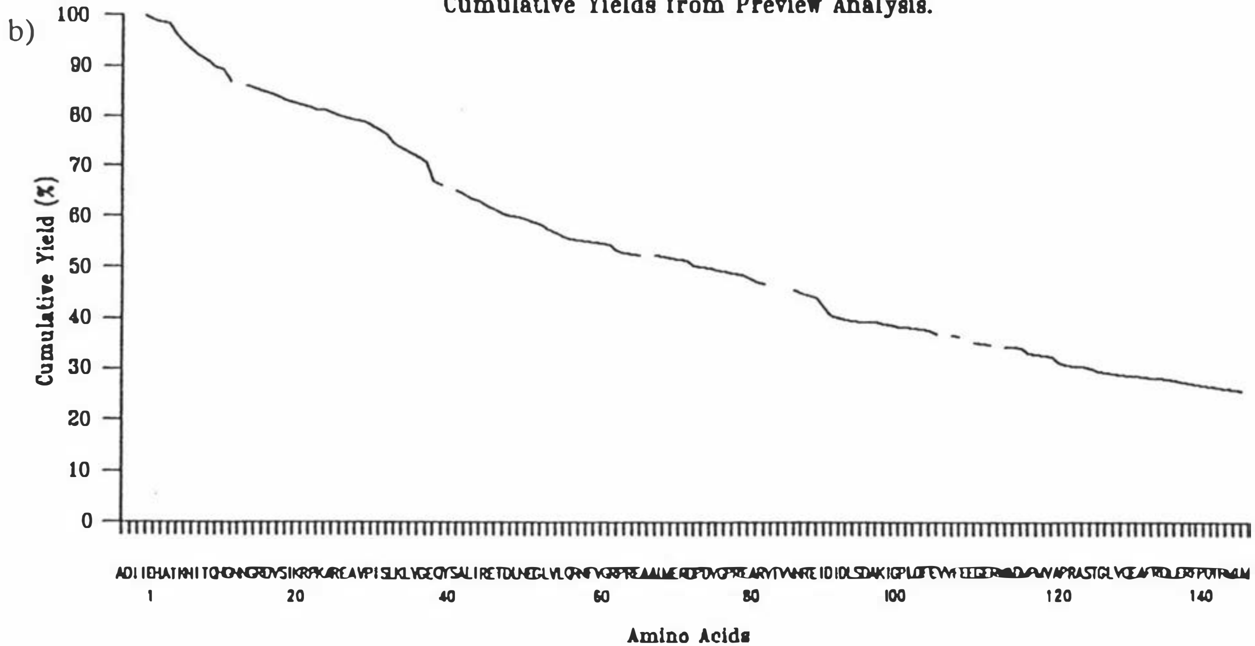


Figure 3.3.14. Protein resin sequence analysis. a) % coupling yields for the amino acids. The lower values are artifactually low due to preview and lag cycle overlap or the presence of side chain protecting groups. b) cumulative yield for the protein on the resin calculated from the % coupling yields. The gaps are due to residues where no coupling yield could be calculated due to the same amino acid in adjacent positions.

being solved. Using the methods published by Applied Biosytems[189] for sequencing peptide resin beads it was noticeable that the lag value (ie amino acid not cleaved in cycle n but in the n-1 cycle) started at the 3rd cycle, became very big at about the 7th cycle and by the 10th cycle the lag value was bigger than the actual value for that amino acid in cycle n. This meant that accurate measurement of peak heights was not possible as most of the amino acid instead of being at cycle n was present in cycle n-1. Despite altering the standard recommended program by increasing the temp and TFA reaction time according to the suggestions made by ABI, the problem persisted.

Finally it was decided that the problem occurred in sequencing from the beads themselves with the reagents not penetrating the beads to react with all the available amino acid for a given residue. The clue to this fact was that there was little lag in the first 2 or three residues, with the lag building rapidly from this point onwards. This points to the fact that not all of the amino acid is being totally cleaved either due to the reagents not penetrating the beads or there being too much amino acid in each step, with the remaining uncleaved amino acid being carried over into the next cycle. This problem combined with the normal build up of lag associated with sequencing rapidly causes the lag to increase to unmanageable proportions, making accurate measurements impossible.

The peptide resin samples were therefore crushed to give a finely suspended powder in DCM/CH_3N. After loading a sample of the powder onto a Zitex filter and drying, another Zitex filter was placed over the top of the resin powder before sealing in the Pyrex reaction chamber. TFA was injected through the reaction chamber to ensure the removal of the Boc group on the N-terminus. After several minutes the reaction chamber was placed in the sequencer and the filters dried with a nitrogen stream before sequencing started.

Using this procedure the value of the lag in cycle n-1 never got above half the peak height of the amino acid in cycle n. The protein resin was sequenced at every 20th residue to provide complete coverage of the protein allowing the entire sequence of amino acids and their integrity to be examined. It is noticeable that when the % coupling values are calculated from the % preview there was variation in the consistency of the samples (Figure 3.3.14). The samples (eg sample at position 62 sequenced 62-41) that had been used previously , suspended then dried down and placed back in storage at -18^oC, gave consistently poorer % coupling values compared to samples that were untouched and stored at -18^oC (i.e. sample at position 83, sequenced 83-61). Most of the low coupling yields that were calculated (Figure 3.3.14a) could be attributed to repeating amino acids eg Gly-Leu-Gly, where the second glycine preview value is a combination of preview and lag from the first amino acid. Hence when calculating the preview for the second glycine the value is higher than normal, leading to a lower % coupling yield for the second glycine. Another

Sample at Residue Number	Preview determined at Residue Number	Cycle Number	Preview (%)	Yield of Correctly Assembled Chains (%)	Average Yield per Step (%)
128	141	16	4.70	95.30	99.68
106	124	19	7.79	92.21	99.55
87	101	16	8.51	91.49	99.41
66	86	22	7.21	92.78	99.64
47	64	19	17.24	82.76	98.95
26	45	21	8.33	91.67	99.56
1	14	15	4.54	95.45	99.67
			OVERALL	49.74	99.49

Table 3.3.6. Quantitative sequence analysis of peptide resin samples. Peptide resins were sampled at the indicated position (first column) and subject to automated Edman degradation on a gas phase sequencer. The premature appearence of the derivitised amino acid (preview) is calculated on the ratio of the peak height at cycle n+1 to the peak height at cycle n. All readings were corrrected for background. The yield of correctly assembled chains is 100-% preview. Average yield per step is calculated as $(\% \text{ preview}/100)^{\frac{1}{n}} \times 100$ where n is the number of cycles over which the preview is determined. The average yields per step

were calculated as follows. $\frac{99.68^{20}}{100} \times \frac{99.55^{21}}{100} \times \frac{99.41^{18}}{100} \times \frac{99.64^{20}}{100} \times \frac{98.95^{18}}{100} \times$

$\frac{99.56^{20}}{100} \times \frac{99.67^{25}}{100}$

reason for the low coupling yield is the second position where the preview is determined from the first cycle or where the residue is a side chain protected amino acid such as lysine (ClZ) where the protecting group inhibits the formation of the phenylthiohydantions.

The average coupling yield from preview analysis of the 137 amino acids for which preview was able to be determined, was 98.96% giving a predicted yield of protein of 21.49% (compared with 98.34 and 8% from the ninhydrin analysis). The higher coupling yields calculated in the same manner as the ninhydrin yields, adds support to the theoretical values which were calculated by adding the background values due to the proline residues.

The % preview determined at any one position gives an indication of the % of failed protein chains. By removing from the calculations all the artifactually low values (as discussed above) then the true % yields can be calculated. Table 3.3.6 shows the % previews determined at approximately the same positions (in order to match lag size). The average % coupling yield for the correctly formed protein chain was 99.49% with a predicted overall protein yield of 49.74%. Sequencing of the total side chain protected protein also provided a check that the protein has been assembled correctly with all the amino acids in their correct order.

The conclusion drawn from these results is that the protein proceeded with no sign of any failed (below 97%) couplings. The % couplings as detected by ninhydrins is low by up to 1% due to the background reading as a result of the non discriminate binding to the protein chain or resin beads. If the artifactual values are excluded from the preview analysis the % coupling averages 99.49%. In this case the overall yield of protein is 49% which compares favourably with the synthesis of interleukin 3[132c] and hGM CSF[169]. Both had average couplings of 99.4% with total yields of 41 and 46.5 % respectively. AAA of the peptide resin samples (Table 3.3.5) excluding the destroyed amino acids show calculated values that are in close agreement with the theoretical values. Sequencing the peptide resin samples showed that all the amino acids were assembled in the correct order on the resin.

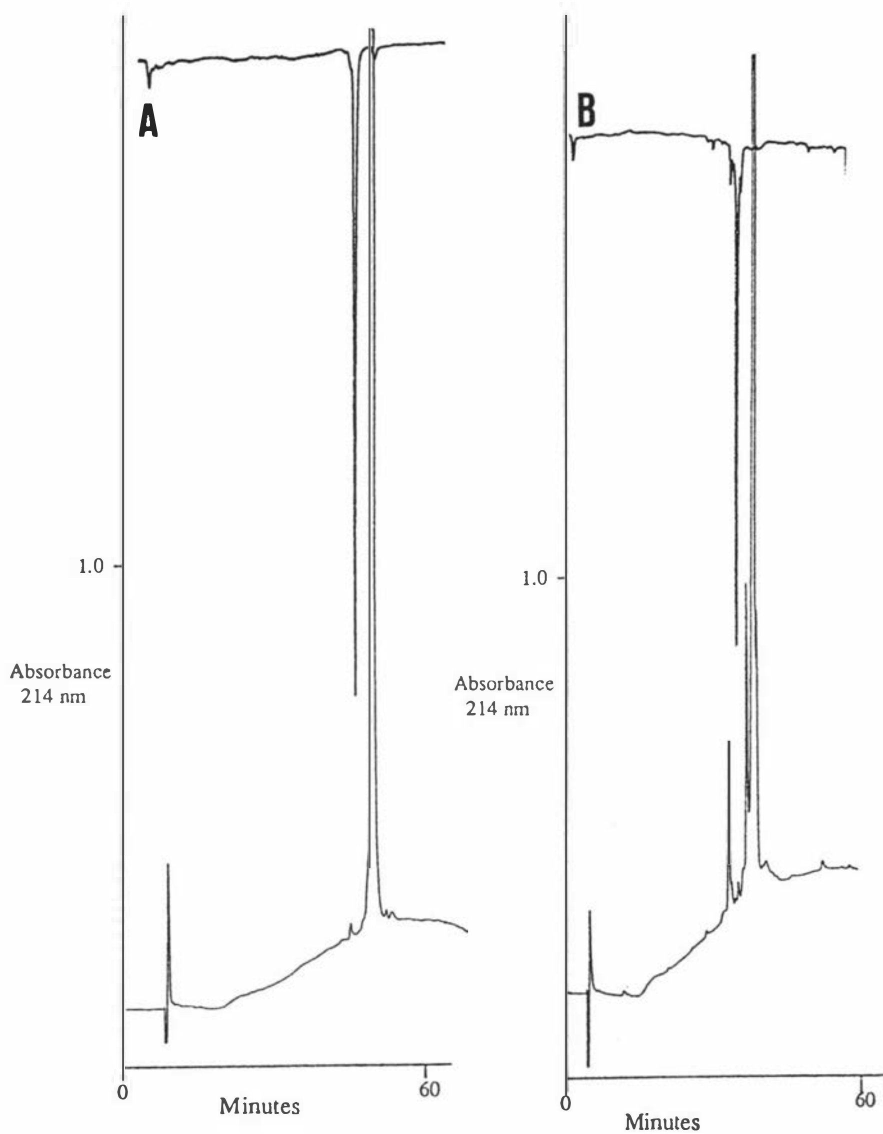


Figure 3.3.15. Analytical HPLC of the recombinant 18 kDa protein as supplied by the Department of Molecular Medicine Auckland Medical School. a) fresh 18 kDa protein. b) After storage at -18°C for 4 months showing the degree of breakdown products. See also figures 3.3.17 and 3.3.48.

3.3.3. PROTEIN CLEAVAGE AND PURIFICTION.

It was decided to directly cleave a sample of the protein resin and to try and work out a purification scheme, despite the possible losses of an expensive product. It was not until the sixth cleavage that a method was sufficiently developed to enable the isolation of 'purified ' protein. The 90 amino acid fragment removed from the synthesis was purified prior to the sixth cleavage and will be discussed in sequence. The 50 amino acid fragment also removed from the protein synthesis was treated separately as described previously. In most cases the SDS-PAGE gels are run with a sample of the 18 kDa recombinant protein for comparison. An analytical HPLC of the recombinant protein is shown in Figure 3.3.15.

3.3.3.1. Cleavage #1

From the Lo-Hi HF cleavage of 300 mg of protein resin, 67.9 mg of crude protein was isolated, a much lower than the expected yield (See cleavage 3 for a better yield recovery). Analytical HPLC of the crude material showed a broad peak spanning the corresponding position of the recombinant 18 kDa product. During initial gel filtration of the crude protein on a Pharmacia Superose 12 HR 30/60 with 0.1N acetic acid,the protein failed to elute off the column even though the concentration of the acetic acid was increased to 10% then 20% acetic acid. The protein was eventually eluted from the column with 70% formic acid, which after lyophilisation gave only 16.2 mg of protein.

Gel filtration was eventually carried out using 2M guanidine HCl, 25 mM Tris HCl at pH7. Fractions taken across the protein peak showed a slight improvement in the analytical HPLC profiles as compared with the crude sample. The protein was then loaded onto the preparative HPLC column and fractions containing the protein were combined according to their retention times on the analytical HPLC. The AAA (Table 3.3.7) on the purified fractions gave consistent results with the expected ratios of amino acids for the 18kDa protein, over all fractions. This showed that their was no visible enhancement in purity in any of the isolated fractions. Samples of each fraction were sent to Auckland Medical School to be assayed for antibody recognition.

SDS-PAGE showed that all fractions of the isolated protein streaked in their lanes with one fraction (lane 3) giving a highly concentrated band around the 18 kDa mark (Figure 3.3.16). The synthetic protein material was detected in all the lanes from the stacking gel interface. The streaking and presence of high molecular weight material could be due to

Amino Acid	Expected	F1	F2	F3	F4
Asx	18	18.01	18.03	17.77	17.72
Thr	6	5.21	5.49	4.90	5.19
Ser	5	5,34	5.17	5.08	4.98
Glx	19	18.83	19.65	19.00	18.76
Pro	9	8.27	8.57	8.76	8.57
Gly	9	9.41	9.79	9.40	9.25
Ala	13	13.00	13.00	13.00	13.00
Val	13	8.52	9.36	7.98	8.49
Met	5	3.20	3.80	3.75	3.73
Ile	9	6.60	6.78	5.36	6.07
Leu	12	10.28	10.48	9.53	10.41
Tyr	1	1.39	1.08	1..23	1.12
Phe	5	3.30	4.00	3.95	3.74
His	2	2.32	2.08	1.84	1.83
Lys	5	5.34	5.05	4.47	4.55
Arg	16	15.63	16.44	15.15	15.72
Trp	ND	ND	ND	ND	ND

Table 3.3.7. Amino acid analysis of 18 kDa protein from Cleavage #1. F1-F4 refers to the combined fractions from the preparative HPLC. ND Not detected.

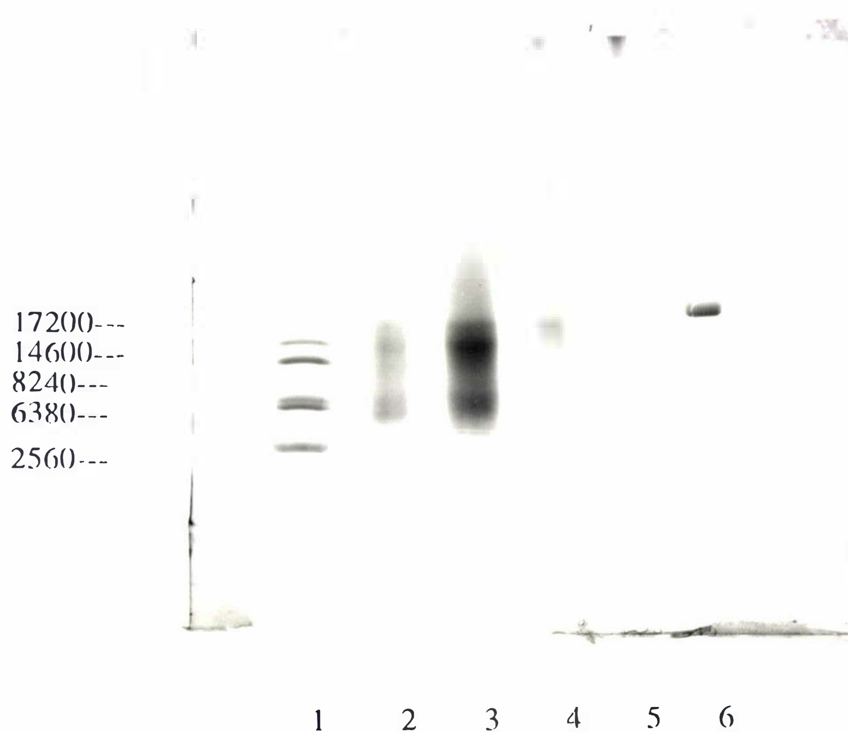


Figure 3.3.16. SDS-PAGE of synthetic 18 kDa protein purified by preparative HPLC. Lane 1: MW standards, 17200, 14600, 8240, 6380, 2560. Lane 2: Fraction from HPLC run # 2. Lane 3: Fl from preparative HPLC run #1. Lane 4: F2 from preparative HPLC run #1. Lane 5: F3 from preparative HPLC run #1. Lane 6: Recombinant 18 kDa.

aggregation of the protein even though SDS was present. ELISAs (as discussed in Chapter 2) on the synthetic protein showed that the synthetic product was recognised by the L5 antibody[190].

3.3.3.2. Cleavage #2

From the HF cleavage of 303 mg of protein resin only 18.6 mg of crude protein was isolated. A check on the weight of the cleaved resin (55 mg) revealed that most of the protein (expected 250 mg) had been removed from the resin. Due to the way the HF cleavage was done, it is most likely that the ether wash contained the remaining protein material. Unfortunately the ether wash was discarded due to the fuming of the ether caused by residual HF, which had been washed from the reaction vessel. It was noticed that the ether solution was a cloudy colour suggesting that the protein had not properly precipitated and was washed through the glass sinter in the ether wash. Hence approximately 230 mg of protein was lost. The remaining 18 mg of protein was subjected to preparative HPLC and stored at -18°C, until it was used along with material from cleavage 3 with the hard acid soft base cleavage method.

3.3.3.3. Cleavage # 3.

HF cleavage on 301 mg of protein resin gave 141 mg of crude protein. The cleaved resin weighed 43 mg leaving unaccounted 117 mg of protein. This time the ether wash was saved, neutralised with a saturated solution of sodium bicarbonate and the ether layer separated. The ether layer was evaporated *in-vacuo* to provide a thick yellow oil. The bicarbonate layer was reacidified with acetic acid and lyophilised before being redissolved and chromatographed on a G-50 column using 50% acetic acid. No protein was detected at the end of the gel filtration or in the ether layer by analytical HPLC. Therefore extreme care must therefore be taken to ensure that the protein precipitates properly before filtering off the ether, to ensure that all the protein in the reaction vessel precipitates properly in the ether solution and does not pass through the filter.

Preparative HPLC was carried out on 2 x 50 mg quantities with fractions taken from across the protein peak to give 32 mg of protein. A sample of this protein was sequenced confirming the integrity of the N-terminal and the correct structure. From the sequencing it was detected that the cleavage had failed to remove the benzyl groups from the threonines at positions 5 and 21. This implies that the protecting group could perhaps be present on other threonines further along the protein chain. There was no sign of any other protecting groups on the first 23 amino acids which includes Arg (Tos), Asp (OBzl) and

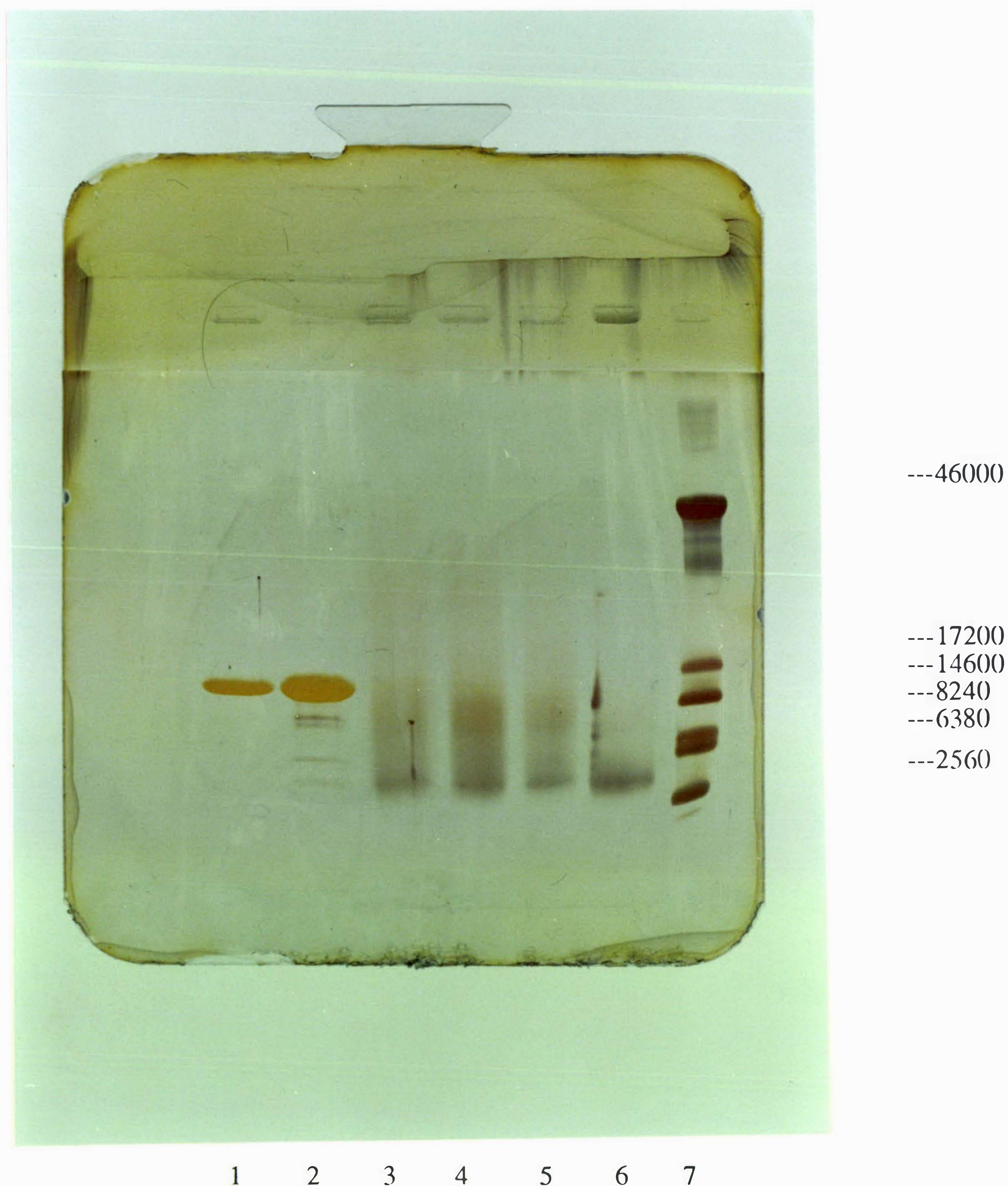


Figure 3.3.17. SDS-PAGE gel Phast System. Lane 1 and 2: Recombinant 18 kDa. Lane 3, 4, 5, 6: Synthetic 18 kDa protein from cleavage 3 after treatment with hard acid: soft base and purification by preparative HPLC. Lane 7: MW standards 46, 17200, 14600, 82590, 6380, 2560. There is present in lane 2 indications of breakdown products or contaminants associated with the recombinant protein.

Glu (OBzl) as determined from sequencing. It is peculiar that the benzyl group was not present on the cleaved 50 amino acid peptides and is present despite the Lo-Hi HF cleavage procedure.

The problem then remained as to how to remove the benzyl group without causing further damage to the peptide. Rather than repeating the HF procedure on the crude material, it was decided to evaluate the cleavage method proposed by Yajima[185], the hard acid-soft base method . TFMSA was ruled out as probably not being sufficiently strong enough, especially as the Lo-Hi HF did not work. After treating a sample of the crude peptide with trimethylsilane trifluorosulphonate/ thioanisole, it was noted that in the gel filtration no size differentiation as seen previously of the protein occurred, although the salts were readily separated. By dissolving the protein in 6M guanidine.HCl to unfold the protein and gel filtering, it was hoped that some resolution could be achieved, thus allowing the separation of any smaller terminated proteins. As is noted below, 6M guanidine.HCl produces better gel filtered material than acetic acid. After the gel filtration, 9.1 mg of protein was obtained.

After the hard acid-soft base cleavage of the bulk of the protein sample, the crude protein was dissolved and gel filtered in acetic acid as soon as possible with the isolation of 5 fractions of protein material. SDS-PAGE showed that no conclusion could be made as to the relative degree of purification of any of the samples. The protein samples were then gel filtered to give 2 fractions, which after lyophilisation were chromatographed on the HPLC to give four fractions. One of these fractions was dissolved in 6M guanidine.HCl and slowly dialysed down a guanidine.HCl gradient to be finally dialysed against water. No improvement in the HPLC profile of the sample was noticed. SDS-PAGE using the Phast System with silver staining (Figure 3.3.17) showed that the majority of the material (lanes 3-6) were streaking below the position of the cloned 18 kDa samples (lane 1 and 2), mainly in the region 8-14 kDa. Sequencing of the N-terminal of the protein isolated from the HPLC showed no sign of the benzyl group on threonine and an intact N-terminus.

3.3.3.4. Cleavage # 4.

With the knowledge that the hard acid-soft base cleavage method would cleave any residual benzyl groups, it was decided to produce more 18 kDa protein in an effort to investigate why there was no apparent sign of material on SDS-PAGE which is of a similar molecular weight to that of the recombinant protein.

A further cleavage was carried out in order to provide protein which could be subjected to a tryptic digestion (see flow chart P.1.in pouch at the back of the thesis), as a

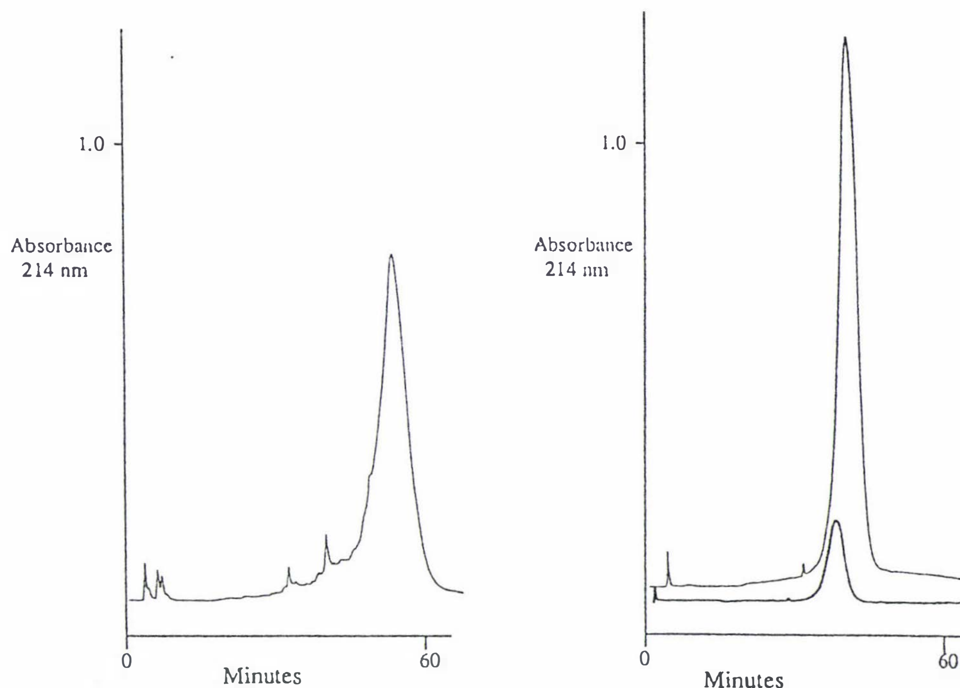


Figure 3.3.18. Analytical HPLC a) crude 18 kDa protein from cleavage 4. b) Gel filtered synthetic 18 kDa protein from cleavage 4 gel filtration, fraction 1.

Amino Acid	Expected	Crude	F1	F2
Asx	18	17.79	18.05	17.94
Thr	6	6.22	5.62	6.13
Ser	5	5.01	5.08	5.17
Glx	19	18.33	19.68	18.62
Pro	9	7.00	8.20	7.51
Gly	9	8.88	9.27	9.28
Ala	13	13.0	13.00	13.00
Val	13	10.38	10.38	11.38
Met	5	2.74	3.42	2.81
Ile	9	9.47	7.64	9.95
Leu	12	9.85	10.95	10.80
Tyr	1	1.43	1.50	1.55
Phe	5	2.86	3.70	3.20
His	2	2.75	2.18	2.79
Lys	5	6.26	5.54	6.80
Arg	16	13.38	16.34	14.69
Trp	1	ND	ND	ND

Table 3.3.8. 18 kDa Protein from Cleavage # 4. Crude refers to the cleaved unpurified protein, F1 and F2 refer to the two gel filtration fractions from the G-50 sephadex column. ND not detected.

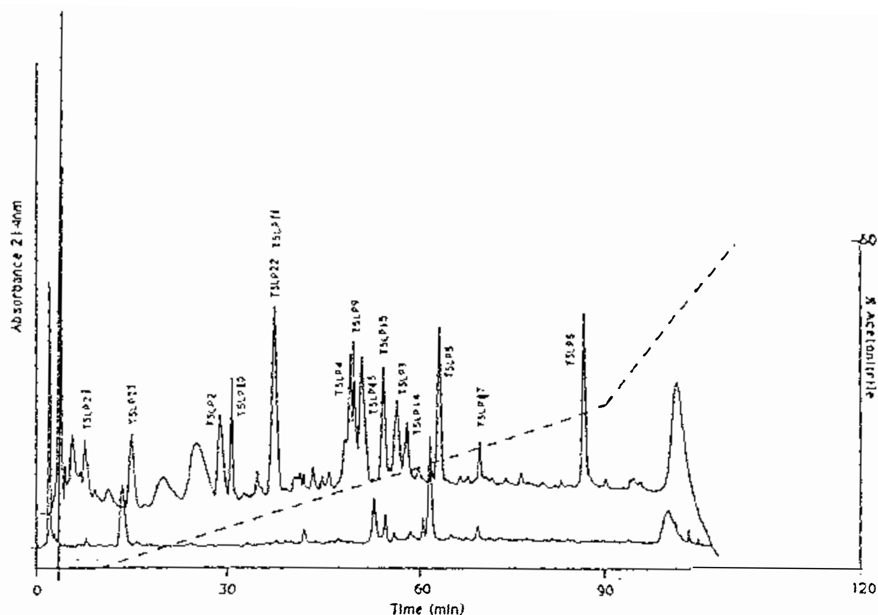


Figure 3.3.20. Tryptic map of gel filtered (fraction 1) synthetic 18 kDa protein from cleavage 4. C18 Synchropak column. Buffer A 0.1% formic acid, 2% CH_3CN. Buffer B 0.1% formic acid 80% CH_3CN. All major peaks (numbered) were trapped and sequenced (Figure 3.3.21).

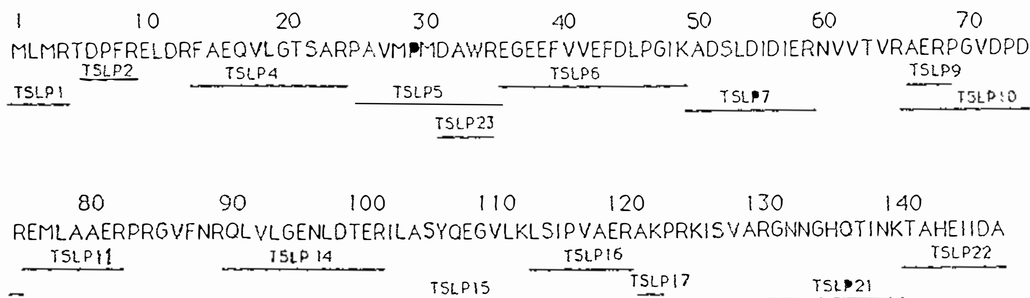


Figure 3.3.21. Protein sequence showing the identified tryptic fragments from digest Figure 3.3.20.

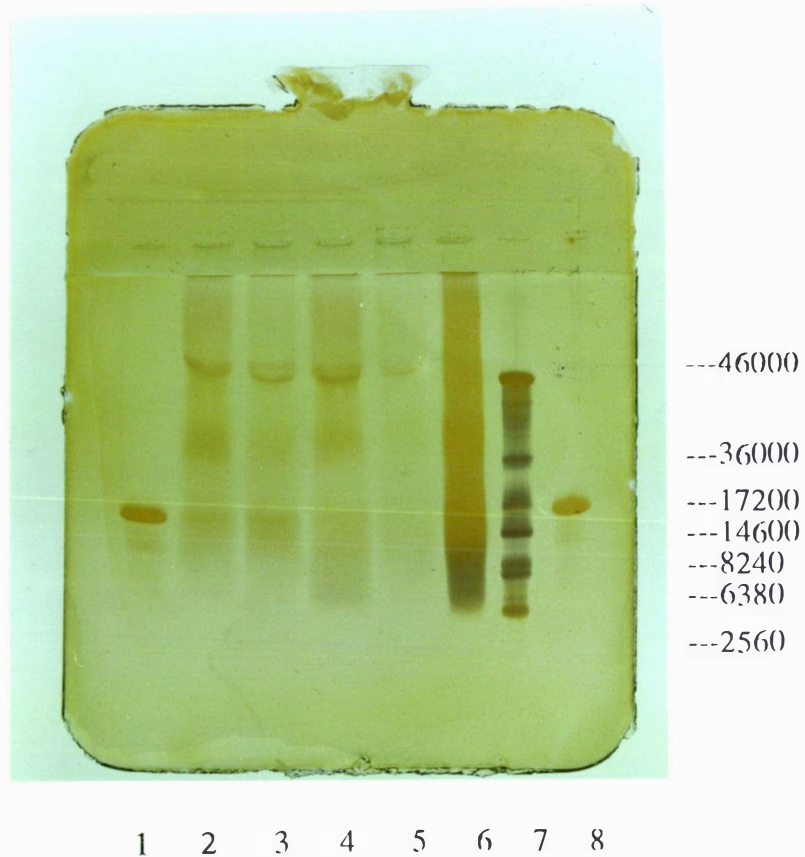


Figure 3.3.19. SDS-PAGE gel of cleavage 4. Lane 1 recombinant 18 kDa. Lane 2: β-mercaptoethanol treated and dialysed synthetic 18 kDa protein. Lane 3: half loading lane 2. Lane 4: Dialysed synthetic 18 kDa protein. Lane 5: half loading of lane 4. Lane 6 Crude 18 kDa protein from cleavage 4 F1. Lane 7:MW standards 46, 36, 17200, 14600, 82590, 6380, 2560. Lanes 2,3,4 and 5 show a band at 48 kDa which is an aberration of the gel. Aggregated protein is detectable at 40 kDa. See also Figure 3.3.25.

further assurance as to the integrity of the synthetic product. After the HF cleavage the protein was gel filtered on a G-50 column and two fractions were isolated and checked on the analytical HPLC (Figure 3.3.18). AAA on the two fractions showed that both had amino acid ratios in accord with the predicted ratios (Table 3.3.8). SDS-PAGE showed that the samples streaked from the stacking gel through the running gel with concentrations at the 45 kDa and the 30-24 kDa region as well as the 18 kDa region. The two bands at the higher molecular weights are thought to be aggregates of the 18 k Da protein as these two regions are approximately two and three times the expected molecular weight (Figure 3.3.19, lane 6).

Tryptic digestion of a sample of the gel filtered protein was carried out and the resulting fragments were chromatographed on a HPLC to give a similar pattern to the tryptic map of the recombinant 18 kDa protein, as supplied by the Molecular Medicine group[190]. The peptide fragments were isolated after chromatography on a C18 analytical column using a formic acid based buffer system[183] (Figure 3.3 20). These fractions were then sequenced for identification.(Figure 3.3.21). All the major peaks were identified as fragments from the 18 kDa protein, with none of the fragments containing an side chain protected amino acid. However several of the protein fragments were still unaccounted for and it is thought that the majority of these occur in the solvent peak. It was decided to leave further work on isolating the peptide fragments from tryptic digests until the problem of the high molecular weight bands was better understood.

In an effort to overcome the aggregation problem a sample of F2 from the initial gel filtration of the crude protein was dissolved in 6M guanidine.HCl and dialysed slowly down a decreasing salt gradient with a final dialysis against four changes of H_2O. This was in an effort to unfold the protein, and then allow it to slowly refold as the guanidine.HCl concentration was gradually decreased, thereby breaking up the 'aggregated' products. SDS-PAGE gel (Figure 3.3.19, lanes 2-5) on this material showed a strong band at about 34 kDa with protein being detected all down the lane to the 2 kDa mark, indicating that the protein was not able to be unfolded then refolded.

3.3.3.5. Cleavage #5 with SDS.

The big question was how to overcome the aggregation of the protein and to overcome the smearing of the protein down the entire lane?. Sequencing, AAA and tryptic mapping has identified the presence of the desired sequence. Certainly it appears at this stage as though the protein is aggregating at a very early stage either during the HF cleavage or when the protein is dissolved immediately after the cleavage. One method that appeared

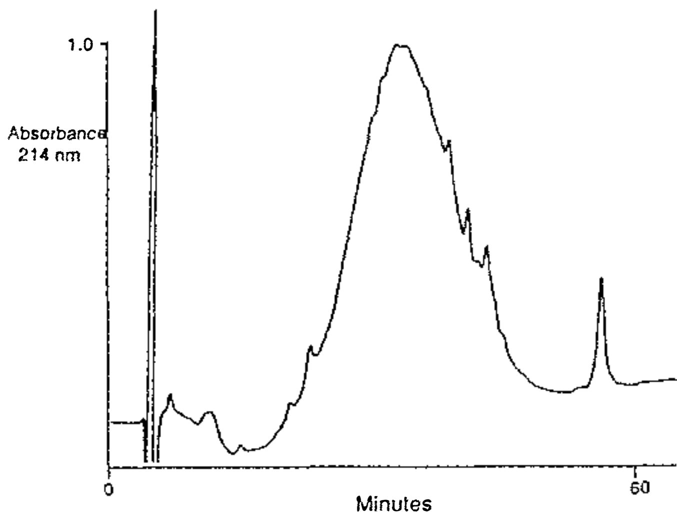
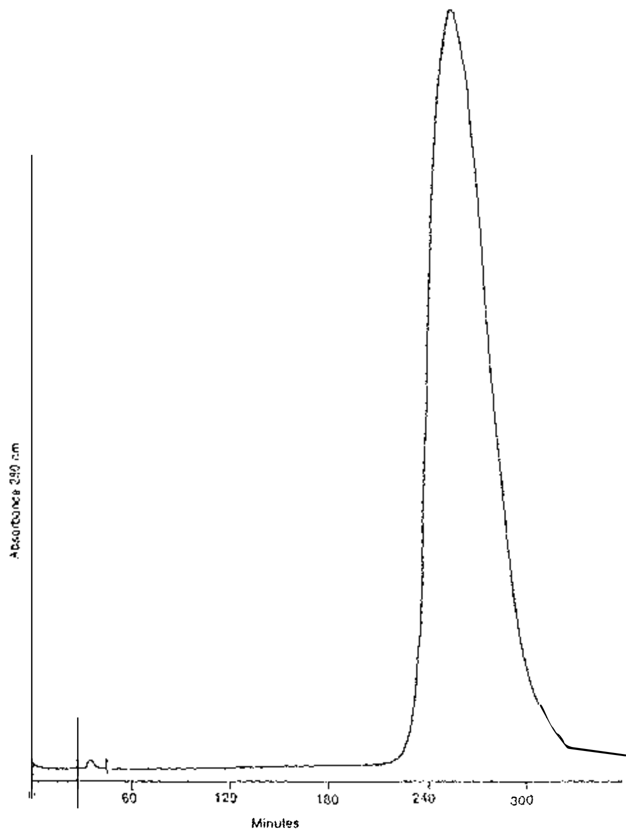


Figure 3.3.22. Crude synthetic 18 kDa protein cleavage 5. C4 Vydac column Solvent system buffer A 0.1 M NH_4HCO_3, pH 8, 2% CH_3CN and buffer B 1:1:1 buffer A:CH_3CN: isopropanol. Gradient 30-100% B in 30 min



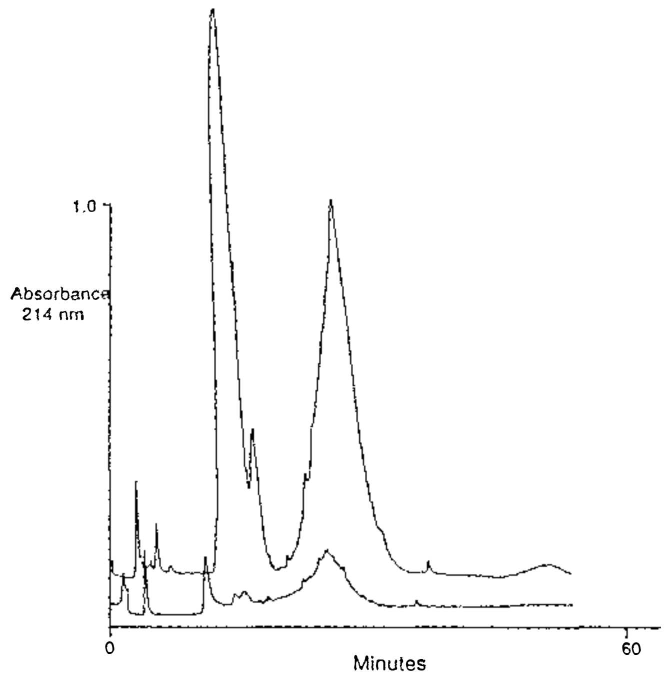


Figure 3.3.23. a) Gel filtration on G-100 column of SDS protein complex. b) analytical HPLC in normal (TFA) buffer system. First peak is SDS material.

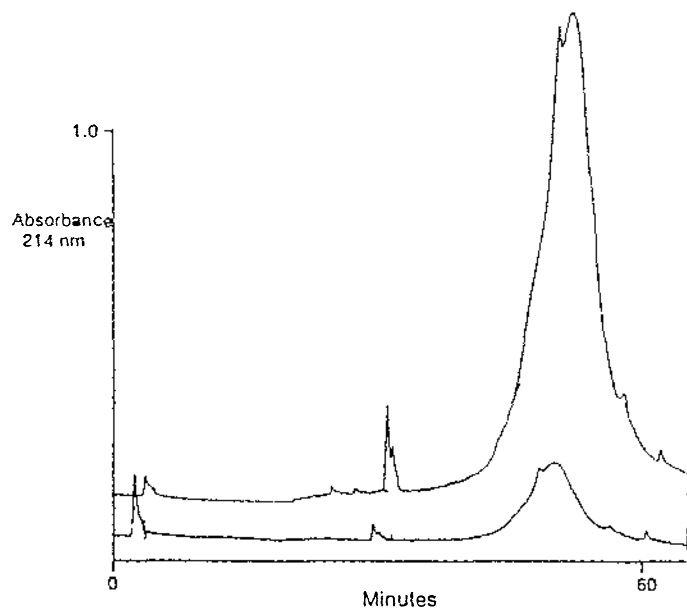


Figure 3.3.24. Analytical HPLC of ultrafiltered material from cleavage 5 after gel filtration.

to be viable was a method reported by Tomich *et.al* [185] who were investigating peptides that aggregate. By adding SDS to the HF cleavage vessel they reported that they were able to prevent the peptides aggregating prior to the purification of the peptides on HPLC where the SDS was exchanged for a volatile detergent .

In applying the same technique to the 18kDa protein, the major problem might be in the purification, which was a single step HPLC in Tomich's case where their peptides were 15-25 amino acids long. An analytical HPLC of the crude protein after the HF cleavage run in a bicarbonate buffer system showed the protein eluted as a very broad peak (Figure 3.3.22). This procedure therefore, would not be sufficient enough for the 18 kDa protein and it was decided to try to retain the SDS with the protein until the final steps of the purification (see flow chart P.2. in pouch at the back of the thesis). A SDS-PAGE on the crude protein after HF cleavage (Figure 3.3.25, lanes 2 and 3) showed that the bulk of the protein was running well below the expected molecular weight . After two initial gel filtrations using a G-50 column, to remove the scavenger material and excess SDS, the protein was then gel filtered on a G-100 column. This was necessary as the size of the SDS-protein complex appears to be greater than the exclusion limit of the G-50 column[191], causing the protein to elute in the void volume of the column.

The gel filtration on the G-100 column in 0.1 M NaCl, 25 mM Tris HCl, pH 8 gave a broad peak that was divided into two fractions (Figure 3.3.23). Preparative HPLC on the analytical column showed that the protein had two peaks in each fraction (Figure 3.3.23). SDS-PAGE using the Phast System three gels showed no detectable protein in any of the lanes, after three separate gels were run. There was an indication that the protein was not proceeding past the stacking gel face. It was noted later that this batch of precast gels had a gap between the stacking gel and the running gel which explains the strange high molecular weight band in the previous gel. (Figure 3.3.20) Therefore a SDS-PAGE slab gel was run (Figure 3.3.26) and this showed streaking from the stacking gel down the lanes. It appeared from this gel, that the concentration of SDS was inhibiting the running of the protein as additional SDS was added to these samples causing an increased amount of protein to precipitate at the stacking gel and running gel interface. The protein samples that ran on the gel showed a predominant band that corresponded to the recombinant 18kDa band, with no sign of the bands forming at the higher molcular weights.

By ultra filtering the remaining crude material, it was postulated that the excess SDS and the yellow colour due to the DNP group would be separated from the protein. The filtration through a 1,000 molecular weight (MW) membrane did not separate the yellow colour, neither did filtration through a 10,000 MW membrane. Preparative HPLC (Figure 3.3.24) of the protein isolated in the retentate from the 10000 MW filtration gave a protein

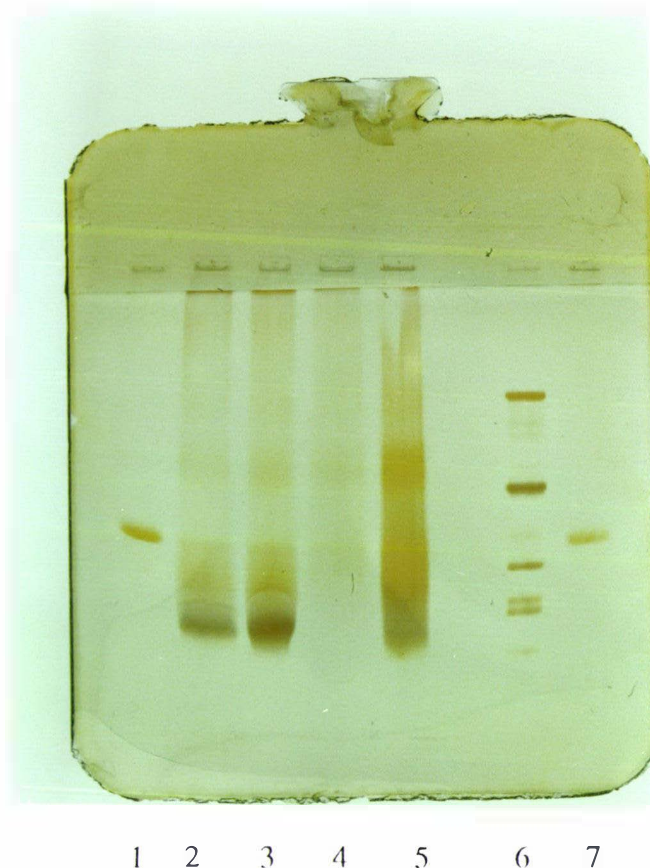


Figure 3.3.25. SDS-PAGE gel of synthetic 18 kDa from synthesis 5. Lane 1: recombinant 18 kDa protein. Lane 2 : crude SDS protein, Lane 3 : twice the concentration of lane 2. Lane 4 : half the concentration lane 2., Lane 5 : material from cleavage 4 F1 at half the concentration of figure 3.3.18 showing more clearly the concentration of aggregated material at the 40 kDa region. Lane 6 : MW standards 46, 36, 17200, 14600, 82590, 6380, 2560. Lane 7: recombinant 18 kDa.

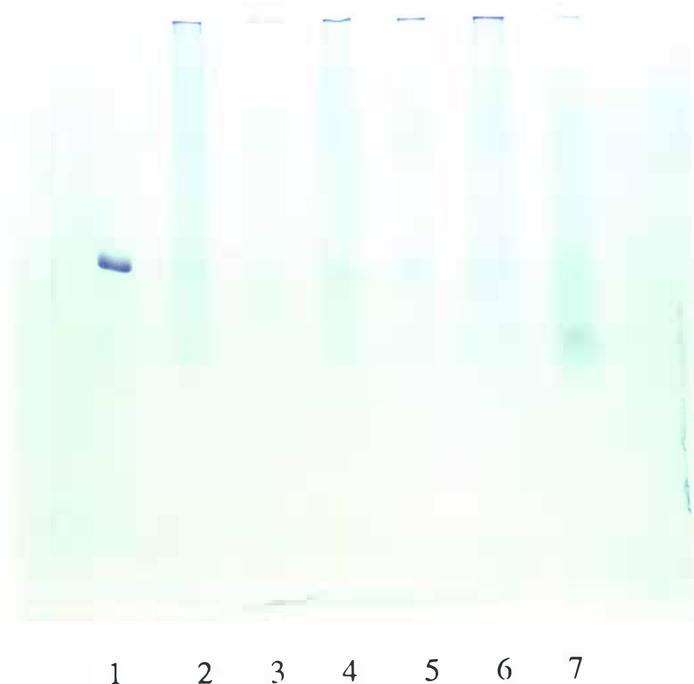


Figure 3.3.26. SDS-Page gel of gel filtered and dialysed synthetic 18 kDa protein from cleavage 5. Lane 1: recombinant 18 kDa. Lane 2 : dialysed synthetic 18 kDa protein. Lane 3 : half the concentration lane 2. Lane 4 : material as for lane 2 with extra 10 µl of SDS. Lane 5: as for lane 4 with 15 µl of SDS. Lane 6 as for lane 4 with 20 µl of SDS. Lane 7 synthetic 18 kDa protein after concentration and ultrafiltration through 10,000 MW membrane.

peak which was trapped , lyophilised and checked by SDS-PAGE (Figure 3.3.26,lane 7). The streaking pattern remained as for the previous gel but with the removal of the precipitation at the gel interface.

The SDS procedure appeared to have removed the problem of aggregation leaving the necessity of purifing the protein in the presence of large amounts of SDS. The other problem that remained was the fact that there was protein material being detected down the entire lane on a SDS gel. The majority of the protein material would be expected to be at the 18 kDa mark or below if terminated protein chains were present.

3.3.3.6. Cleavage of the 90 Amino Acid 18 kDa Peptide (Arg^{48}-Ala^{148})

Another alternative for the purification of the synthetic protein was to dissolve the protein after the HF cleavage in 6M guanidine HCl. This method was used for the 50 mer peptide 1-50 which was dissolved in 6M guanidine.HCl to protect the N-terminal Asp-Pro bond. This procedure has been also used for the purification of the synthetic AIDS protease enzyme[170,192]. Though this would not prevent the aggregation if it occurs within the HF vessel, any aggregation after the cleavage should be prevented.

To investigate whether or not the more sensitive amino acids in the N-terminal region of the protein were playing a part in the aggregation problems it was decided to do the next HF cleavage on the 90 amino acid protein. This strategy would also conseve the dwindling supply of the synthetic 18 kDa protein In the 90mer the problem associated with the protection of the tryptophan due to the reactive indole side chain is removed. Similarly four of the five methionines which are prone to oxidation are eliminated as is the N-terminal Asp-Pro bond. (See Figure P.3 in the pouch at the back of the thesis for the flow chart to the purification procedure described below).

After the Lo-Hi HF cleavage of the 90 amino acid protein, the protein was dissolved in 6M guanidine HCl and samples were immediately chromatographed on a G-50 gel filtration column (Figure 3.3.27) and analysed by HPLC (Figure 3.3.28). The following day the protein peak from the gel filtration column was loaded on the preparative HPLC column and the protein peak eluted and trapped as two fractions, the front half and the back half of the peak (Figure 3.3.29). These fractions were then lyophilised and redissolved in 6M guanidine HCl and reloaded separately onto a gel filtration column. During this second gel filtration step it was noticed that the protein peak exhibited two major equal sized peaks for both of the HPLC fractions (Figure 3.3.30). These peaks were trapped and dialysed against decreasing salt concentration of guanidine.HCl and finally against the sample buffer that was used to dissolve samples before loading onto PAGE gels. As a final step the

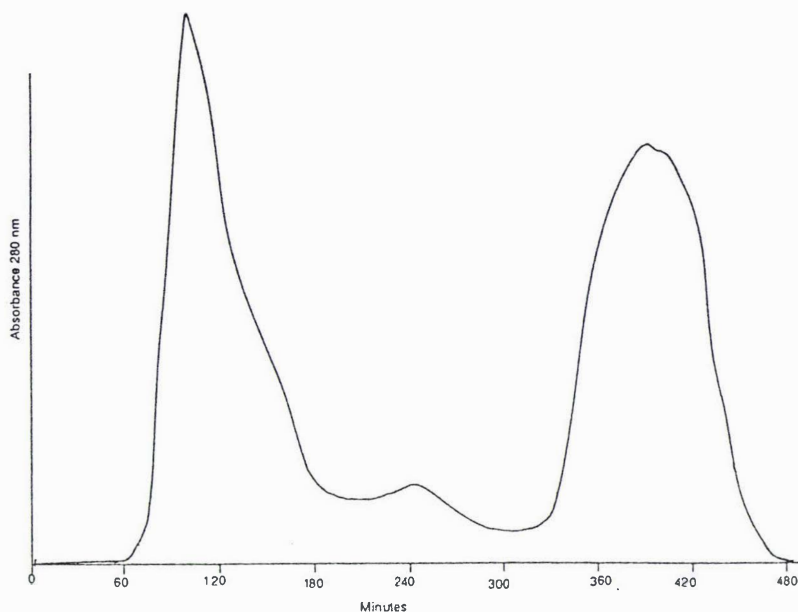


Figure 3.3.27. Gel filtration on G-50 column in 6M guanidine.HCl of crude 90 amino acid protein. Peak 1: protein. Peak 2 : organic scavengers from HF cleavage.

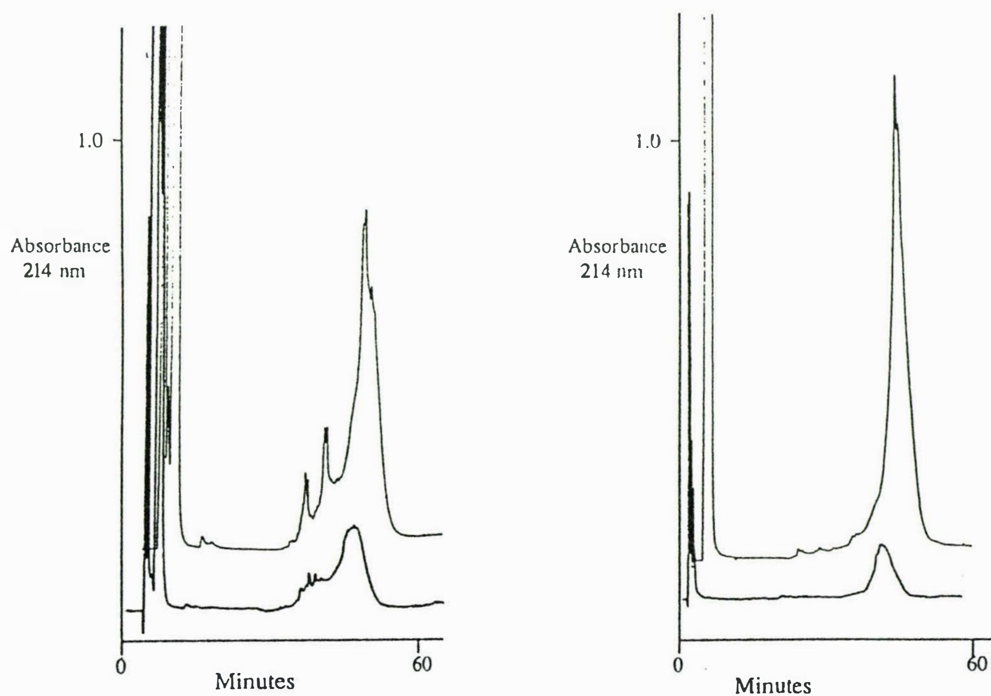


Figure 3.3.28. Analytical HPLC a) crude 90 amino acid protein. b) Protein material after gel filtration (figure 3.3.27).

Figure 3.3.29. Preparative HPLC of 90 amino acid protein from gel filtration (figure 3.3.27) 1st peak solvent peak (guanidine HCl), second peak guanidine.HCl, third peak 90 amino acid protein. The protein peak was divided into two fractions as indicated and loaded separately onto a gel filtration column. See figure 3.3.30 for profile of the second fraction of the protein peak.

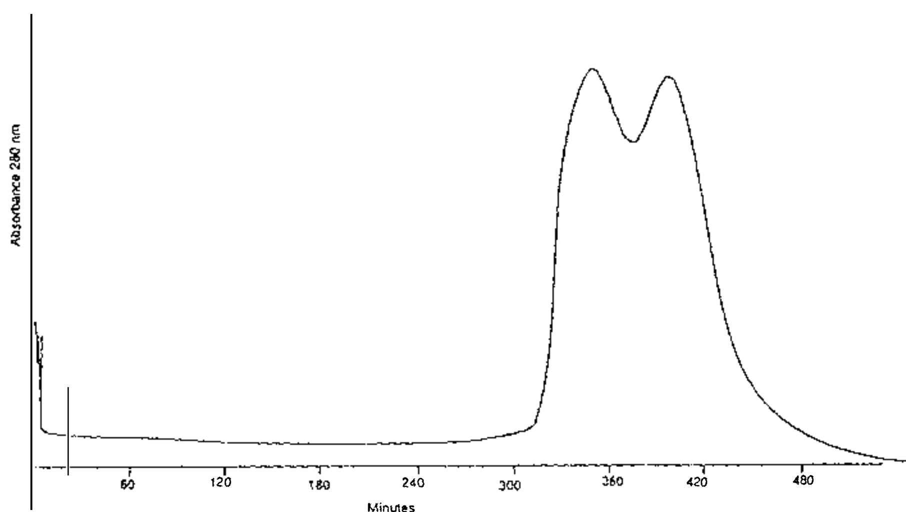


Figure 3.3.30. Gel filtration of 90 amino acid protein after preparative HPLC (figure 3.3.29) showing the two peaks with the protein in the second peak see figure 3.3.31.

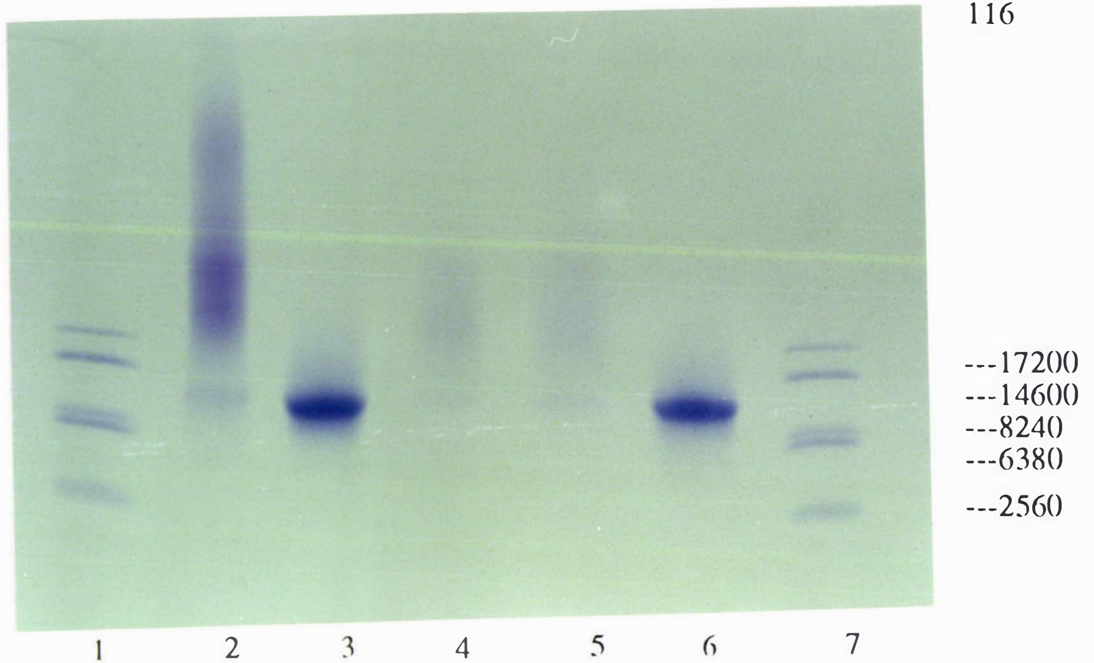


Figure 3.3.31. SDS-PAGE gel of 90 amino acid material isloated after gel filtration (figure 3.3.30) and dialysis. Lane 1 and 7.MW standards 17200, 14600, 8240, 6380, and 2560. Lane 2 : solid material that precipitated during dialysis. Lane 3 : second peak from gel filtration. Lane 4 : first peak from gel filtration. Lane 5 : first peak from gel filtration . Lane 6 : second peak from gel filtration.

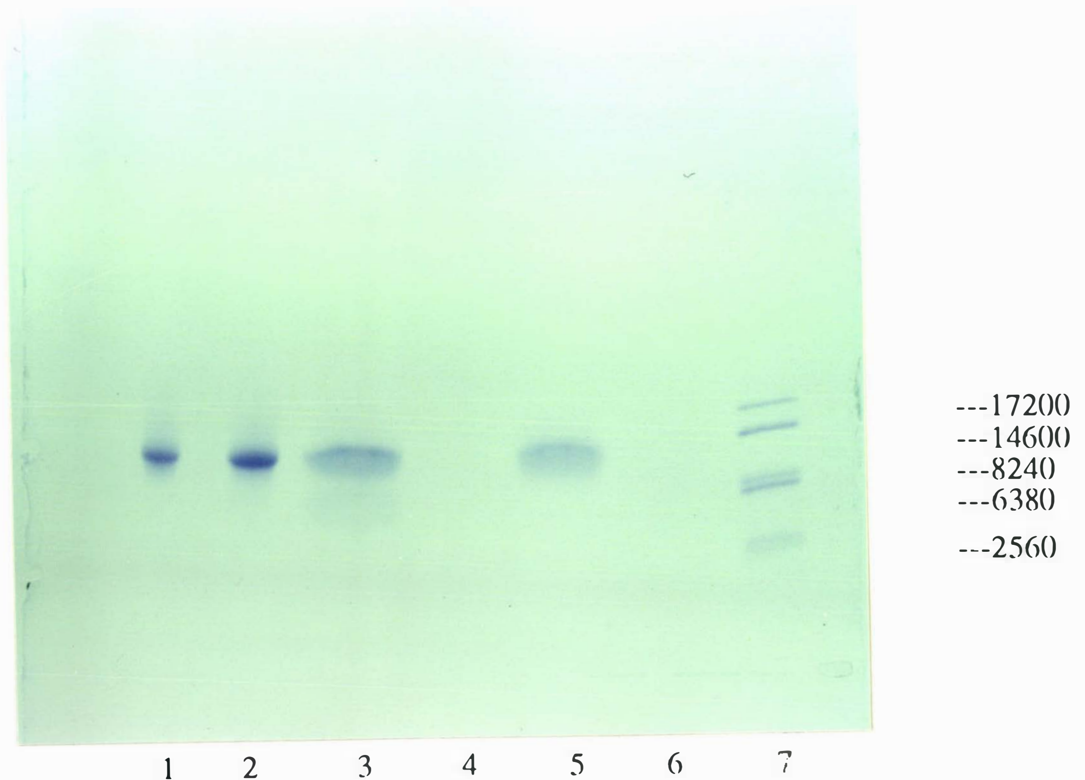


Figure 3.3..32. SDS-PAGE of repeat purification of 90 amino acid protein. Lane 1 : material from lane 3 (figure 3.3.31). Lane 2 : material from lane 5 (figure 3.3.31). Lane 3 : second peak from gel filtration. Lane 4 : first peak from gel filtration. Lane 5 : second peak from gel filtration. Lane 6 : first peak from gel filtration. Lane 7 : MW standards 17200, 14600, 8240, 6380, and 2560.

samples were dialysed against a 1.6% sucrose solution which concentrated the protein solution to approximately 3 ml. The sucrose solution is also used as a high density agent for the running of the slab gels. Removing a sample of the protein solution and loading it onto a SDS-PAGE gel (Figure 3.3.31) showed that in both cases the 90 amino acid protein fragment was essentially pure in the second peak of the last gel filtration (lanes 3 and 6). The first peak had what appeared to be aggregated material as shown by SDS-PAGE gel (Figure 3.3.31, lanes 4 and 5).

A repeat of the above procedure was made to further validate the procedure before it was tried on the full 18 kDa protein (Figure 3.3.32, lanes 3-6)). Time course amino acid analysis of the purified 90mer showed the correct ratios with the expected (Table 3.3.9) The purified protein solutions were then lyophilised and stored at -18°C. Mass spectrometry was carried out on this sample which indicated that two fragments were present with masses of 6586 and 4172. This splitting of the protein was noticed in the gel of the samples from the last purification (Figure 3.3.32 lane 3). As a result some of the 90 amino acid protein was chromatographed on the analytical HPLC column (Figure 3.3.33) with the single protein peak trapped and sequenced to determine where the protein had fallen apart. The sequencing indicated that the protein had split at Phe^{87}-Asn^{88} to give two fragments plus the full length 90 amino acid protein, as shown on the SDS-PAGE gel were the protein is detected at 10 kDa, 6 kDa and 4 kDa (Figure 3.3.34, lane 1).

The breakage of this peptide bond is the normal cleavage site exhibited for the proteases chymotrypsin and thermolysin, which have not been used at all in this work or in the laboratory. A further gel filtration using 10 mM Tris .HCl pH 7 using the Superose 12 gel filtration column was carried out to separate the fragments from the full length 90 mer to give the pure 90mer as shown on a SDS-PAGE gel (3.3.35, lane 3).

Amino Acid	Expected	Peptide Resin	24	48	72
Asp	11	12.04	11.16	9.82	10.01
Thr	4	4.46	3.46	3.59	3.28
Ser	3	1.92	3.20	2.84	2.61
Glu	12	11.14	11.17	10.27	10.00
Pro	5	4.46	4.69	4.57	4.72
Gly	6	6.00	6.19	5.84	5.88
Ala	8	8.68	8.00	8.00	8.00
Val	9	7.54	4.26	7.20	7.31
Met	1	0.36	1.04	0.92	0.92
Ile	6	6.16	3.54	5.29	5.38
Leu	7	4.21	5.93	6.87	6.78
Tyr	1	1.17	0.93	0.96	0.97
Phe	1	0.92	0.75	0.94	0.93
His	2		1.82	1.77	1.74
Lys	4	4.76	3.97	4.19	3.94
Arg	11	9.87	10.77	10.70	10.37

Table 3.3.9. Amino acid analysis of the 90 amino acid protein fragment. The headings 24,48 and 72 refers to the number of hours of hydrolysis to give a more accurate Val and Ile ratio.

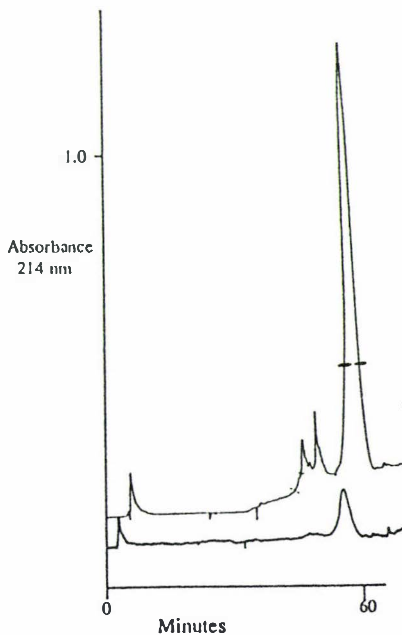


Figure 3.3.33. Analytical HPLC of stored purified 90 amino acid protein. Protein peak was trapped and run on SDS-PAGE (figure 3.3.34) and sequenced.

Figure 3.3.34. SDS-PAGE gel of protein peak trapped from the analytical HPLC (Figure 3.3.33) showing the intact 90 amino acid fragment and the two fragments at 6 and 4 kDa. Lane 2: recombinant 18 kDa protein.

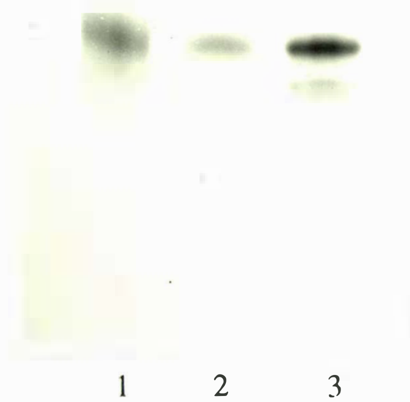


Figure 3.3.35. SDS-PAGE gel of the 90 amino acid protein after gel filtration from figures 3.3.33 and 34. Notice that the intact protein lane 3 still has associated with it the 6 kDa protein fragment.

3.3.3.7 Cleavage #6

Using the purification procedure developed for the 90 mer, the final batch of 18 kDa protein-resin was cleaved (see flow chart P.4 in the pouch at the back of the thesis). In this case the formyl group on the tryptophan was removed prior to the HF cleavage to help minimise any side reactions as a result of carrying this process out at the same time as the HF cleavage. The Lo-Hi HF cleavage was carried out and the product dissolved in 6M guanidine.HCl. The protein was immediately gel filtered on a G-50 column (Figure 3.3.36) to remove the residual scavengers. The protein solution was then filtered through a 1000 MW membrane,which failed to remove the yellow colouration present that could be a result of the DNP group. Preparative HPLC was carried out with the protein peak being divided to give two fractions (Figure 3.3.37).

The gel filtration of the fractions from the HPLC was carried out on a G-100 column in 6M guanidine.HCl. The first fraction from the HPLC was gel filtered giving single peak from which the isolated fractions were dialysed against the sample buffer then concentrated using a 1.6% sucrose solution. These samples were then loaded on a SDS-PAGE gel which showed a diffuse band in one of the fractions corresponding to the 18kDa recombinant protein (Figure 3.3.38). The second fraction from the HPLC was then gel filtered on a G-100 column giving two peaks (Figure 3.3.39), from which the isolated fractions were dialysed slowly down the salt gradient from 6M guanidine HCl to H_2O. The dialysed samples were then concentrated by filtration though a 10000 MW membrane. The retentate was checked on a SDS-PAGE gel (Figure 3.3.40). This gel showed that it was possible to separate the aggregated material (lanes 2 and 3) from the band of protein that corresponds to the recombinant protein (lane 5).

In subsequent purifications this last gel filtration step was switched to the Superose 12 column as the G-100 kept compacting due to the viscosity of the 6 M guanidine.HCl. It was found that the G-100 column run against gravity also did not help the flow rate.

The samples run on the second gel were lyophilised and hydrolysed for identification by amino acid analysis (Table 3.3.10). The AAA indicated that each fraction consisted of the 18 kDa protein and this suggested that the streaks on the gel were 'aggregated' protein or protein which still has some form of chemical modification on the side groups or a chemical rearrangement of the amino acid chain. The term aggregated is used loosely to describe

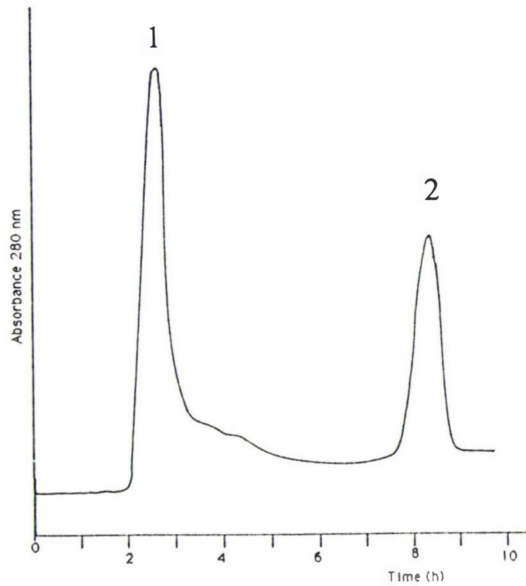


Figure 3.3.36. Gel filtration of crude 18 kDa on a G-50 column in 6M guanidine.HCl. Peak 1 is the protein peak 2 is organic scavenger material from the HF cleavage.

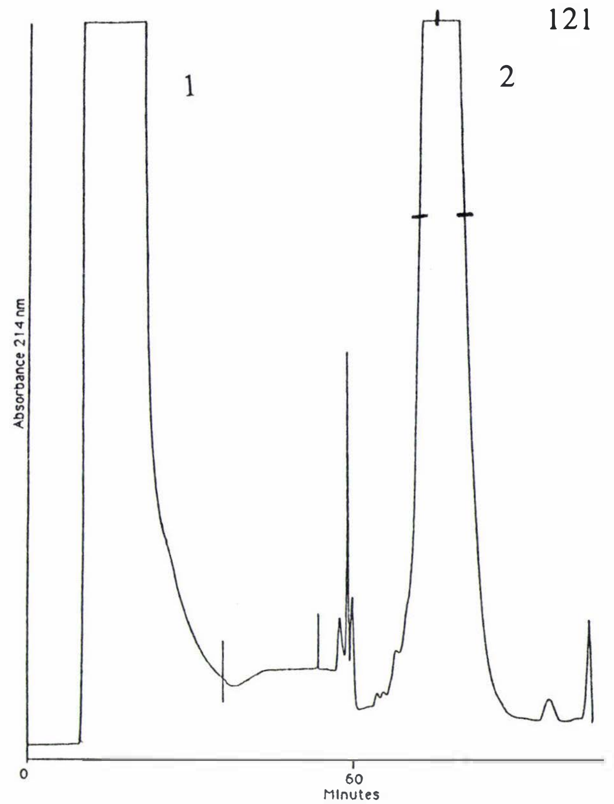


Figure 3.3.37. Preparative HPLC of 18 kDa protein. The first peak is the solvent peak guanidine .HCl. The second peak is the protein peak which was divided into two fractions.

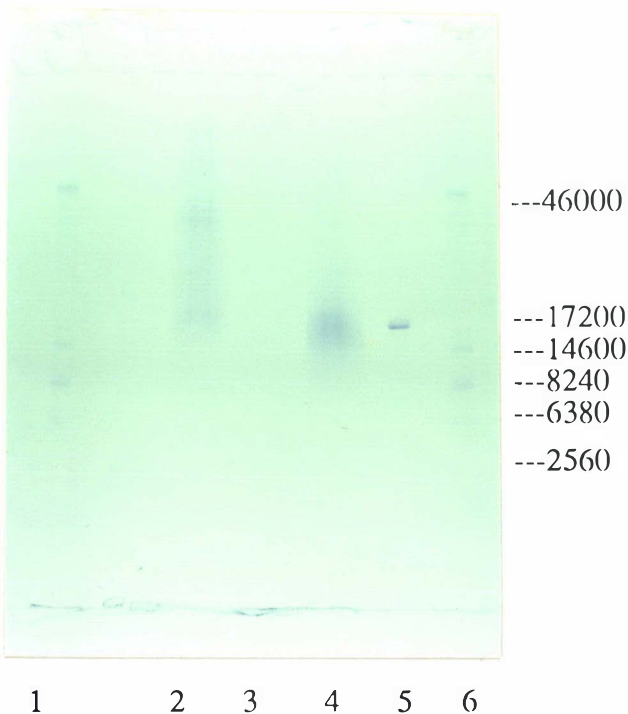


Figure 3.3.38. SDS-PAGE of purified synthetic 18 kDa protein from the first fraction isolated from the preparative HPLC and gel filtered (which gave only one peak, not shown). Lane 1 : MW standards 17200, 14600, 8240, 6380, and 2560. Lane 2 : fraction 1 from gel filtration. Lane 3 : fraction 2 from gel filtration. Lane 4 : fraction 3 from gel filtration. Lane 5 : recombinant 18 kDa protein. Lane 6 : MW standards.

Amino Acid	Expected	F1	F2	F4
Asx	18	19.32	17.68	17.89
Thr	6	5.37	5.36	5.01
Ser	5	8.19	5.67	7.17
Glx	19	20.92	19.21	19.11
Pro	9	7.87	8.25	7.63
Gly	9	14.67	9.94	17.13
Ala	13	13.00	13.00	13.00
Val	13	2.87	8.54	4.49
Met	5	3.45	3.36	3.07
Ile	9	6.16	6.40	5.63
Leu	12	10.99	10.58	10.00
Tyr	1	1.44	1.31	1.43
Phe	5	4.24	4.06	3.84
His	2	2.11	1.97	2.02
Lys	5	5.64	5.07	5.20
Arg	16	14.57	15.17	13.64
Trp	ND	ND	ND	ND

Table 3.3.10. Amino acid analysis of the 18 kDa protein from cleavage 6. F1-F3 refers to the fraction from gel filtration colmn figure 3.3.39. See also SDS-PAGE gel figure 3.3.40. ND not done.

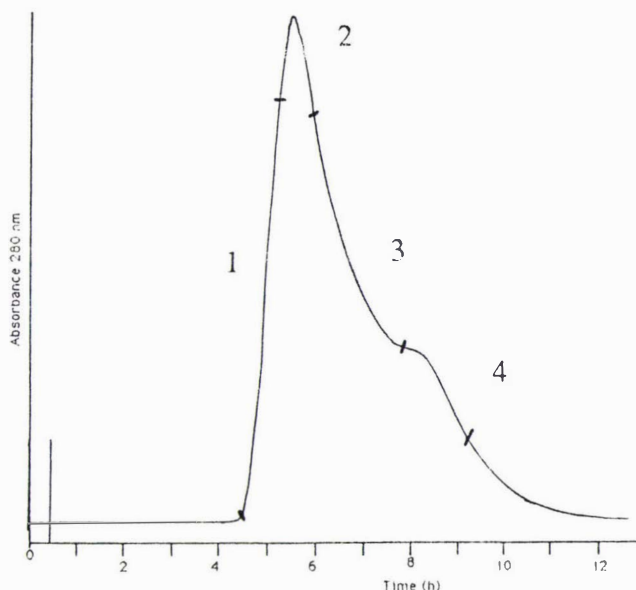


Figure 3.3.39. Gel filtration on a G-100 column in 6M guanidine.HCl of one of the 18 kDa fractions from the preparative HPLC (figure 3.3.37). Fractions were dialysed and run on SDS-PAGE (figure 3.3.40).

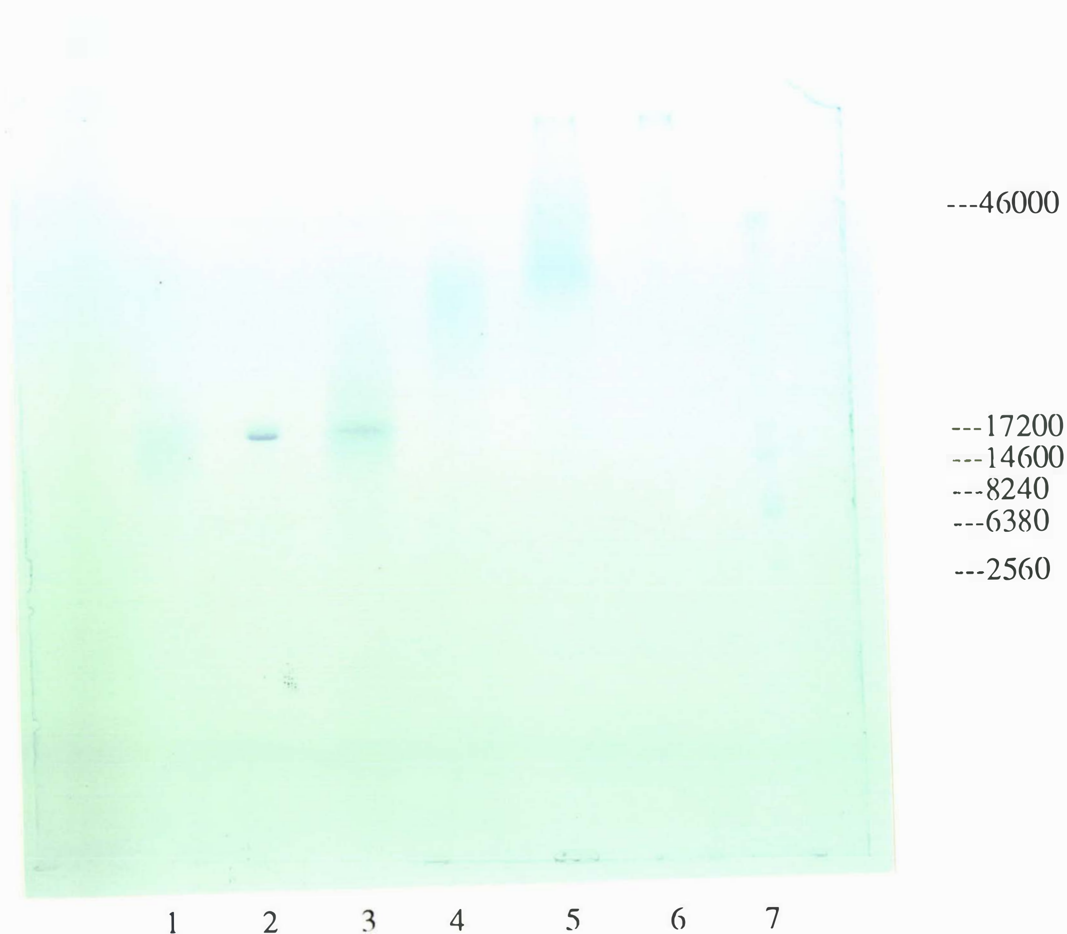


Figure 3.3.40. SDS-PAGE of fractions from the gel filtration (figure 3.3.39) after dialysis. Lane 1 : F3 from figure 3.3.39. Lane 2 : recombinant 18 kDa protein. Lane 3 : fraction 4 from gel filtration. Lane 4 : fraction 3 from gel filtration. Lane 5 : fraction 2 from gel filtration. Lane 6 : fraction 1 from gel filtration. Lane 7 : MW standards 46, 17200, 14600, 8240, 6380, and 2560.

what is detected on the SDS gels. This is because the SDS acting as a detergent is supposed to unfold aggregated products after boiling for 5 min though some proteins are known to still be aggregated after this treatment. From the gels run here it is obvious that if aggregation is occurring, the SDS is unable to separate the high molecular weight products despite them being boiled in the presence of SDS for 5 min. Attempts at increasing the concentration of SDS in the samples to separate the aggregated products only increased the amount of product precipitated at the gel face or did not allow the protein to run on the gel at all (see Figure 3.3.26).

The N-terminal of the sample giving the band (Figure 3.3.40, lane 3) corresponding to the 18 kDa recombinant protein was sequenced and from this it was detected that the benzyl group had not been removed from threonine. Therefore in an attempt to see if the streaking was due to the protecting group a sample of the fractions that indicated 'aggregated' material was treated with trimethylsilane trifluorosulphonate in TFA and thioanisole as for the previous treatment with the hard acid-soft base cleavage. After gel filtration then preparative HPLC using the analytical C4 column (Figure 3.3.41), the protein peak was isolated and lyophilised. A SDS PAGE gel of this material indicated that the sample was still streaking with two concentrated bands appearing at the 18 kDa and 36 kDa marks (Figure 3.3.41, lane 1).

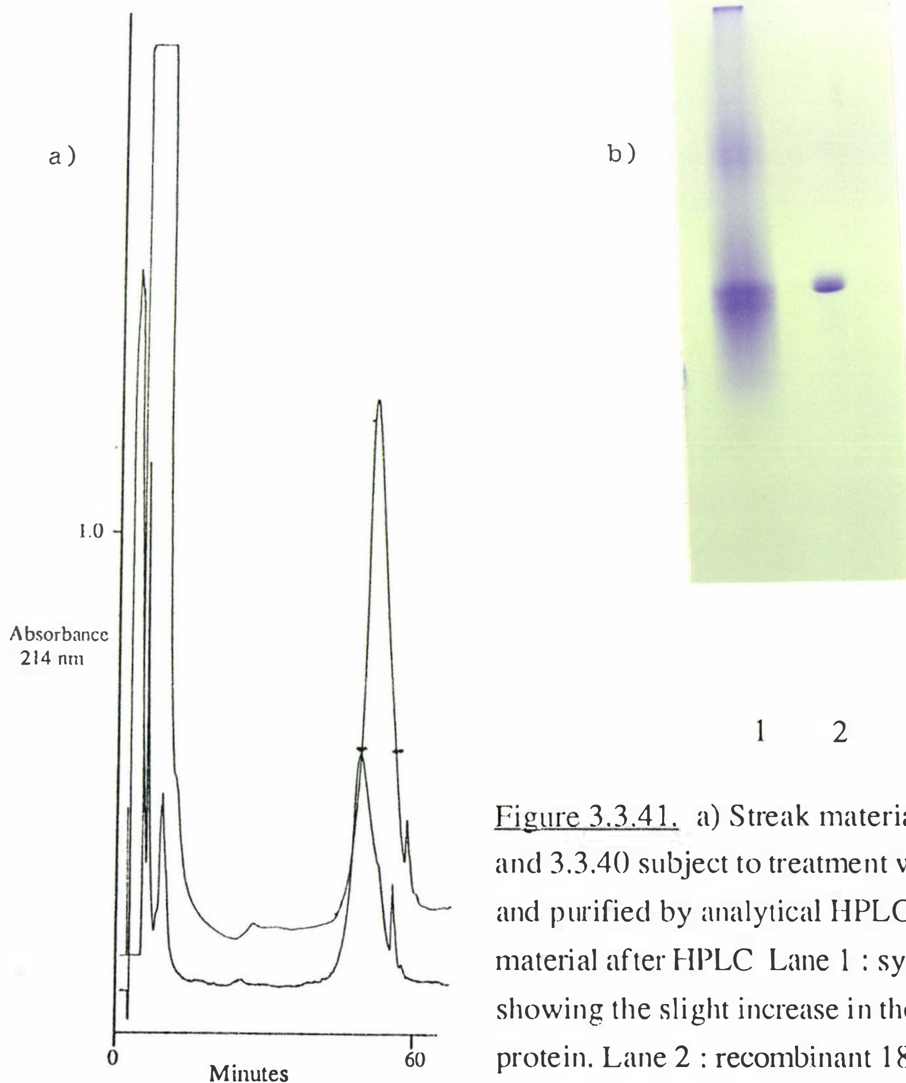


Figure 3.3.41. a) Streak material from figure 3.3.39, and 3.3.40 subject to treatment with hard acid soft base and purified by analytical HPLC. b) SDS-PAGE of material after HPLC Lane 1 : synthetic treated material, showing the slight increase in the amount of 18 kDa protein. Lane 2 : recombinant 18 kDa protein.

Amino Acid	Expected Ratio	Cloned 18kDa 24h	Cloned 18kDa 72h	Synthetic 18kDa 24h	Synthetic 18kDa 72h	Synthetic 18kDa Aggregated 24h	Synthetic 18kDa Aggregated 72h
Asx	18	15.79	16.68	18.57	18.67	17.74	17.97
Thr	6	5.48	5.83	5.34	5.71	4.91	5.54
Ser	5	5.42	5.46	5.02	4.84	4.75	4.72
Glx	19	17.05	18.02	19.28	19.32	18.35	19.02
Pro	9	6.56	6.94	8.39	8.12	8.08	7.87
Gly	9	9.30	9.86	9.69	9.67	9.67	9.70
Ala	13	13.00	13.00	13.00	13.00	13.00	13.00
Val	13	8.42	10.94	8.33	10.64	ND	ND
Met	5	3.31	3.30	3.59	3.40	4.08	3.82
Ile	9	5.57	6.88	6.54	8.35	5.64	8.28
Leu	12	10.26	11.44	10.65	11.79	9.69	10.87
Tyr	1	2.02	2.06	1.44	1.14	1.89	1.08
Phe	5	3.83	4.23	3.71	3.91	4.19	4.04
Lys	5	6.15	6.89	5.43	5.82	4.59	5.52
His	2	1.85	2.03	2.43	2.27	1.95	2.07
Arg	16	11.00	11.73	15.31	15.25	15.86	14.46
Trp	1	ND	ND	ND	ND	ND	ND

Table 3.3.11. Amino acid analysis of cloned 18 kDa protein, synthetic 18 kDa protein pure as identified by SDS-PAGE gel and of the streak material form the synthetic 18 kDa protein as identified by SDS-PAGE gel. The streak material amino acid ratios are close to the synthetic 18 kDa protein ratios. Valine value for the aggregated material was not determined as the sample was contaminated with residual hydrolysis material from a valine resin sample. ND not detected.

3.3.3.8 Cleavage of Stored 18 kDa Protein from Cleavage # 4.

The previously cleaved material without the benzyl group present (cleavage #4, see flow chart Figure P.1 in the pouch at the back of the thesis) was then purified in the same manner to remove the protein from the 'aggregated' material. The protein from fraction 1 which had previously been gel filtered in acetic acid was gel filtered again using 6M guanidine HCl with the resulting protein fractions being trapped (Figure 3.3.42). The protein peak was loaded on to the HPLC and the single protein peak divided into two fractions which were then lyophilised. Loading these samples directly onto a SDS-PAGE gel (Figure 3.3.43) showed that the protein was concentrated in the third peak (lane 4) from the gel filtration column (Figure 3.3.42). A repeat of the above procedure for fraction 2 from cleavage 4 gave a large single peak on gel filtration from which three fractions were isolated. Loading each fraction on the preparative HPLC (Figure 3.3.44) two peaks were detected, with the 'aggregated' material in the first HPLC peak (see Figure 3.3.45 lanes 4 and 5) and the bulk of the protein material in the second peak as shown by SDS-PAGE (Figure 3.3.45, lanes 2 and 3). These samples were not further gel filtered or dialysed as it appeared from the SDS-PAGE gels that the purity of the protein had already reached a stage comparable to the two previous purifications.

Time course AAA (Table 3.3.11) was carried out on the recombinant 18kDa, the purified synthetic 18 kDa and the fraction that was giving the 'aggregated' material. Both the synthetic samples gave ratios that compared closely with the theoretical and the recombinant 18 kDa. UV spectra of the purified synthetic protein and the recombinant 18 kDa protein were identical (Figure 3.3.46), which would indicate that there were no formyl groups on the tryptophan or DNP groups on histidines.

As a further check, tryptic maps (Figure 3.3.47) were made on each of the three samples. Each sample was identical to the other with the exception of the peak heights between the synthetic and the recombinant material. The major peaks were trapped and checked by sequencing to see that they matched the identities from the previous digest of the crude material (Figure 3.3.20). The digested fragments of the purified synthetic material were also checked by mass spectrometry as a further check to make sure that no protecting groups were present . In all the fragments checked by mass spectrometry the masses of the fragment corresponded exactly between the synthetic purified protein and the cloned material with no sign of any other chemical modifications.

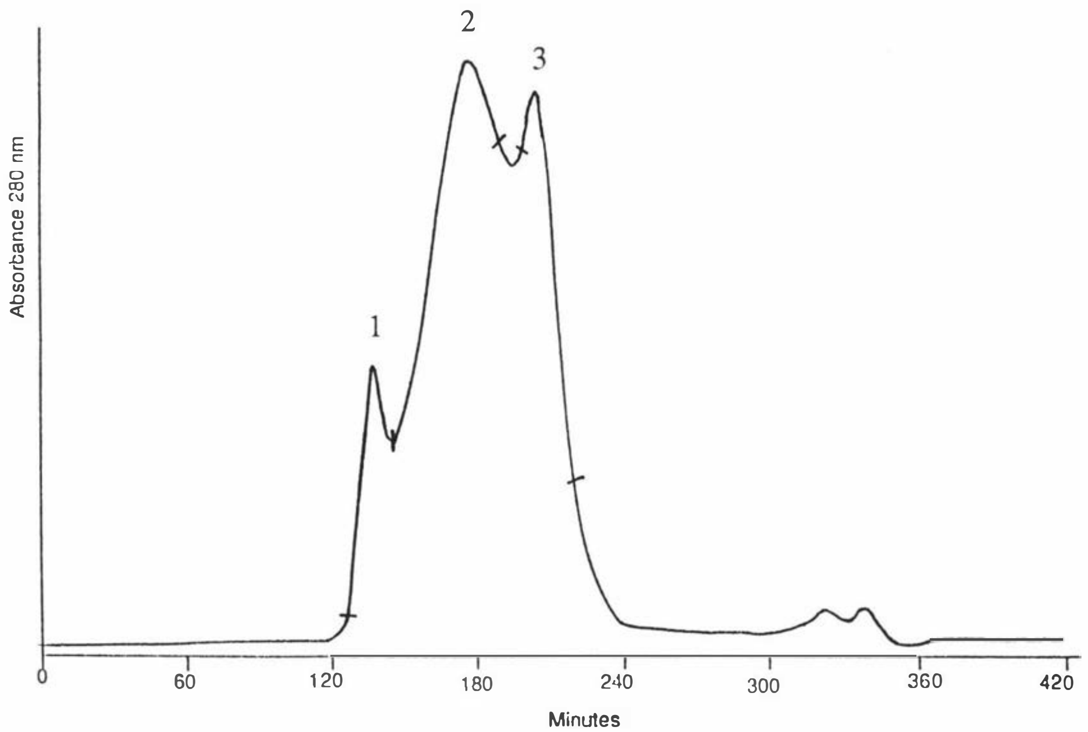


Figure 3.3.42. Gel filtration in 6M guanidine.HCl of stored synthetic 18 kDa from cleavage 4, fraction 1

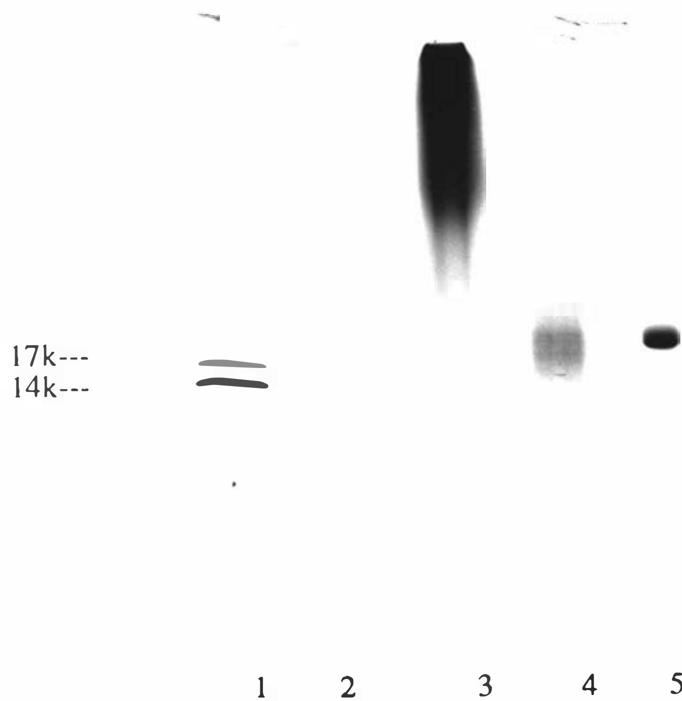


Figure 3.3.43. SDS-PAGE of material from gel filtration fraction 1 (Figure 3.3.42) after preparative HPLC. Lane 1 ; MW standards 17200, 14600. Lane 2 :peak 1. Lane 3 : peak 2. Lane 4 : peak 3. Lane 5 : recombinant 18 kDa protein.

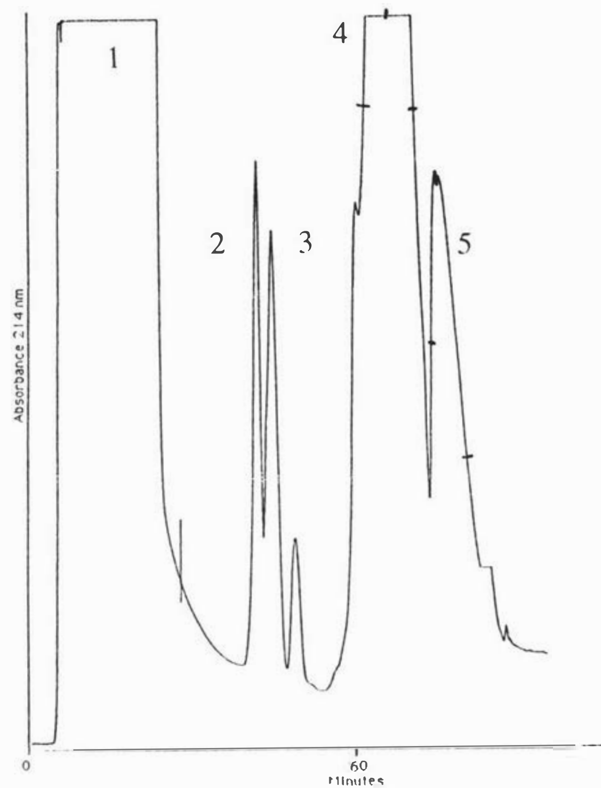


Figure 3.3.44. Preparative HPLC of material from gel filtration fraction 2, Two protein peaks are present. Peak one solvent peak guanidine.HCl. Peak 2 and 3 left over guanidine.HCl. Peak 4 see Figure 3.3.45 lanes 4 and 5. Peak 5 see Figure 3.3.45 lanes 2 and 3.

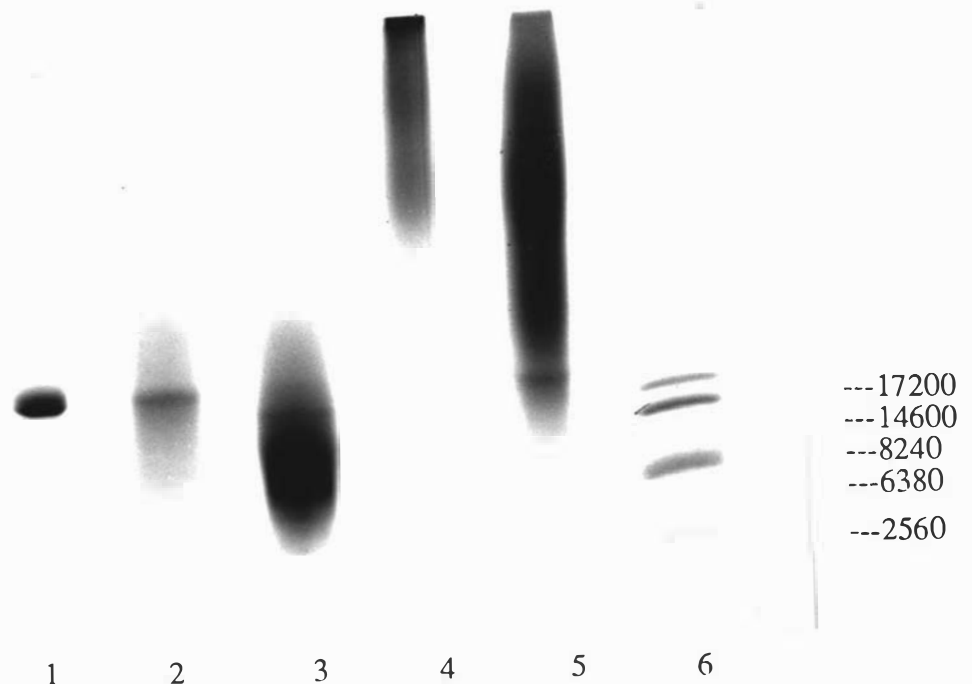


Figure 3.3.45. SDS-PAGE of synthetic 18 kDa from the second gel filtration (not shown) after preparative HPLC (figure 3.3.43). Lane 1 : Recombinant 18 kDa. Lane 2 : fraction 1 from the preparative HPLC. Lane 3 : fraction 2. Lane 4 : fraction 3. Lane 5 : F4. Lane 6 : MW standards 17200, 14600, 8240, 6380, and 2560.

a)

b)

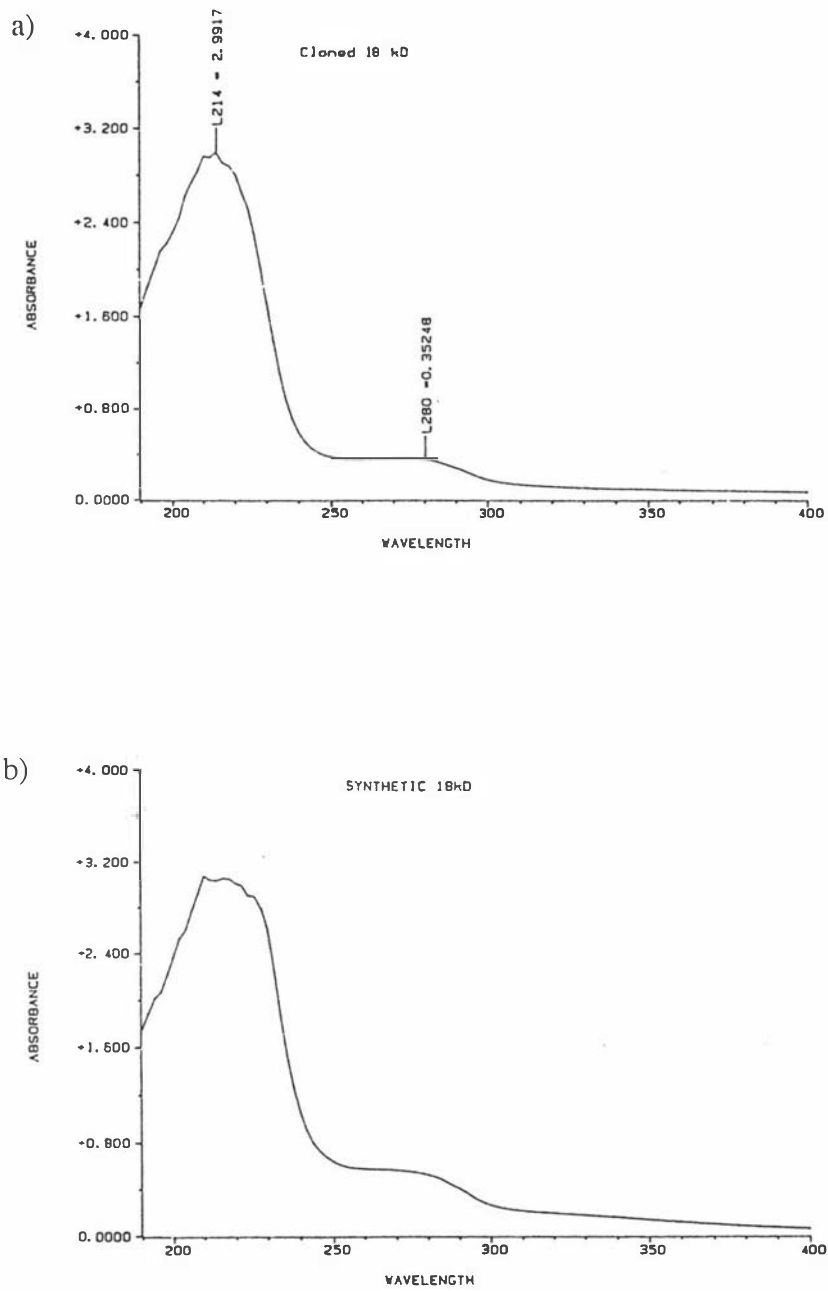
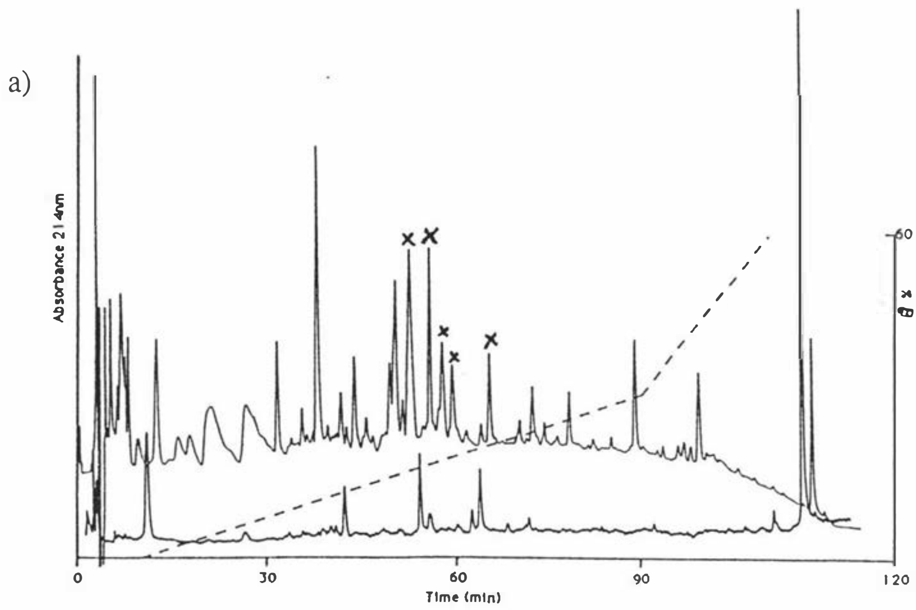


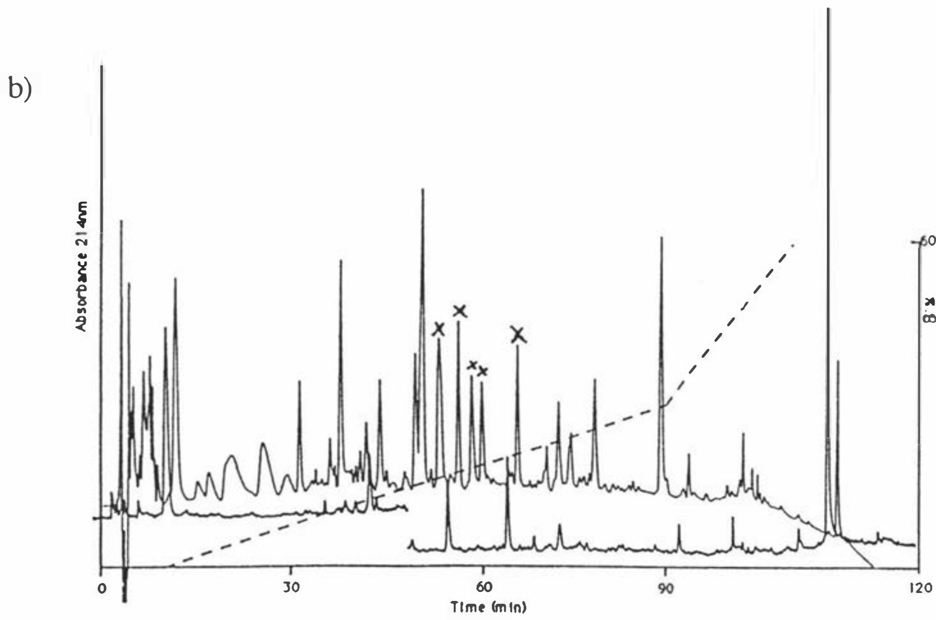
Figure 3.3.46. UV spectra of a) cloned 18 kDa protein. b) synthetic 18 kDa.which are nearly identical.

Figure 3.3.47. Tryptic digests maps of a) synthetic 18 kDa protein. b) recombinant 18 kDa protein. c) material that streaks. The differences between the synthetic purified 18 kDa protein and the recombinant protein is in the peak heights of some of the fractions. For the tryptic map c) the solvent was changed and was not able to have the formic acid concentration ~0.8% adjusted to give the correct separation as for the purified and recombinant protein. The peaks are however the same. The marked peaks were the fractions checked by mass spectrometry which showed that the masses of the peptides between the synthetic purified 18 kDa protein and the recombinant 18 kDa protein were the same.

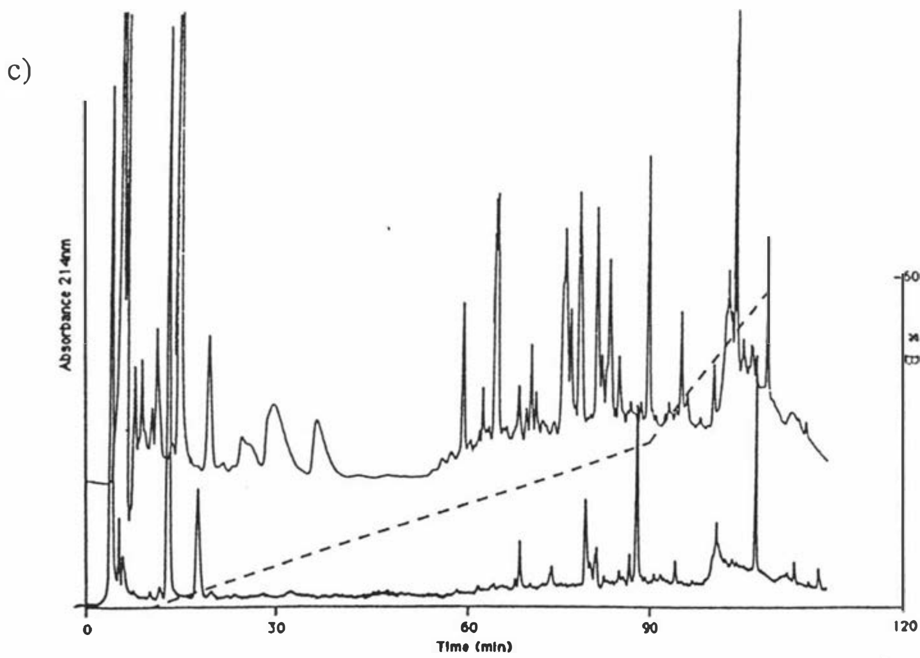
a)



b)



c)



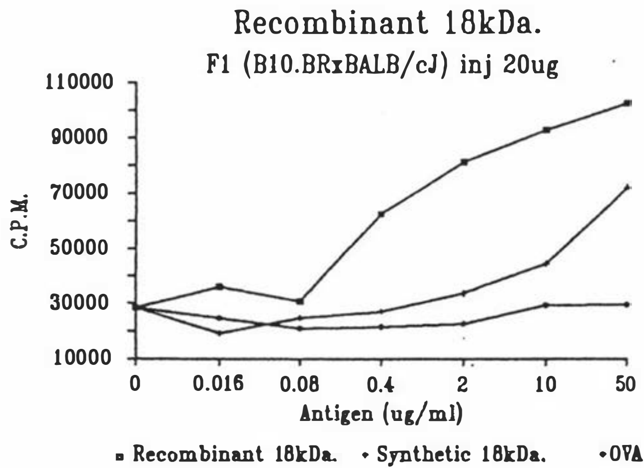
Attempts at purification of the synthetic protein with either ion exchange using a Pharmacia Mono Q HR 5/5 column, or hydrophobic interaction chromatography with a Pharmacia Phenyl Sepharose HR 5/5 column proved to be unsuccessful in providing any identifiable product, probably due to the large range of materials present in the synthetic protein with different electropotentials and behaviour. This large range of material could include incorrectly folded structure leading to aggregated products, correctly folded aggregated protein, breakdown products caused for example by cleavage of the Asp-Pro bond or a small percentage of side chain protecting groups still present on the protein chain.

3.3.3.9 Protein Assays

It is very difficult with the lack of a functional assay for the protein to assess the success of the synthesis in terms of it functionality. The best that could be done regarding this aspect was to compare the proliferation of T-cells as a result of injecting mice with the synthetic or the cloned 18kDa protein. As can be seen by the graphs of the T-cell proliferation (Figure 3.3.48) the synthetic product does appear to behave in the same manner as the recombinant 18 kDa protein. The proliferation of T-cells as a result of injecting with synthetic or recombinant 18 kDa protein is at its highest when cells were challenged in culture with the protein from the same source. This probably means that the synthetic 18 kDa protein is not being recognised in exactly the same manner as the recombinant protein.

Further proof of this is also shown in the results of antibody recognition of the recombinant and synthetic protein. The antibodies were raised from the peptides synthesised previously in Chapter 2 and used for western blots. In this technique the protein is electroeluted onto nitrocellulose from a SDS-PAGE gel then incubated with anti peptide antibodies. After developing, the protein that binds the antibody is visualised by staining with 4-chloronaphthol. As a check some of the nitrocellulose blots were stained to detect the total protein on the nitrocellulose (Figure 3.3.49). The total protein stain shows that the whole streak of the synthetic 18 kDa material (lane2) stains as protein as does the 90mer (lane 1). The recombinant 18 kDa band (lane 3) shows up as two bands as a result of the 18kDa protein and 15 kDa protein which does not have the C-terminal 131-148. The 15 kDa protein occurs in the cloning process.

The western blots for antibodies to 1-20 16-35, 76.95, RLP-23 (131-148) and the L-5 antibody were done so that the major antibody recognition areas (L5 and 76-95) can be checked as well as checking whether the N-terminal of the protein (1-20 and 16-35) is present and recognisable compared to the C-terminal of the protein (RLP-23). From the



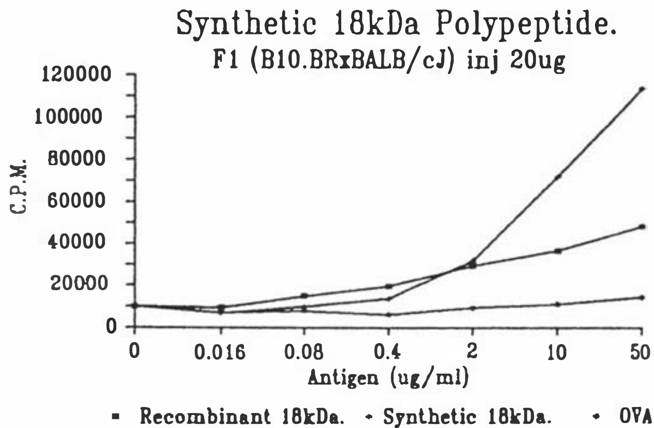


Figure 3.3.48. T-cell proliferation response to immunisation with synthetic 18 kDa protein compared to the proliferation response due to immunisation with recombinant protein. a) Mice immunised with recombinant 18 kDa and challenged *in vitro* with recombinant , synthetic and ovine serum albumin. b) Mice immunised with synthetic 18 kDa protein and challenged *in vitro* with recombinant synthetic and ovine serum albumin. The T-cell response differs between the recombinant and the synthetic protein which could be due to incorrect folding of the synthetic product or chemically modified amino acids.

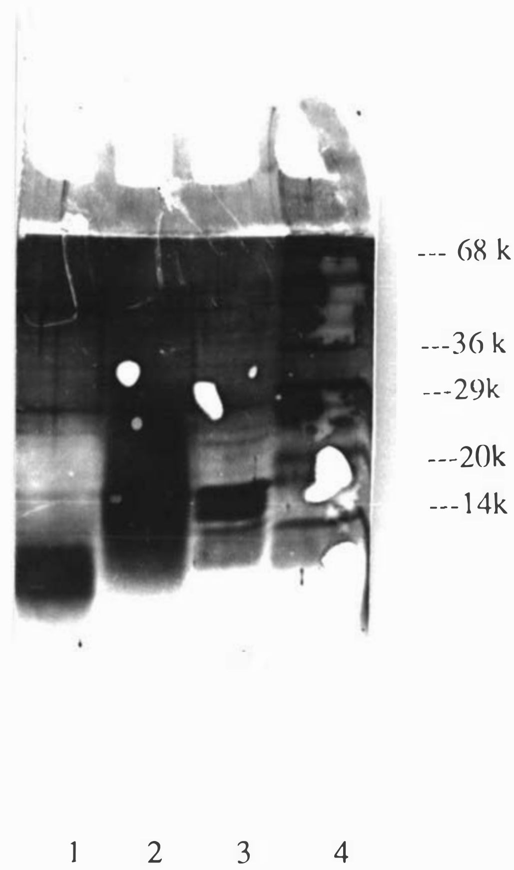
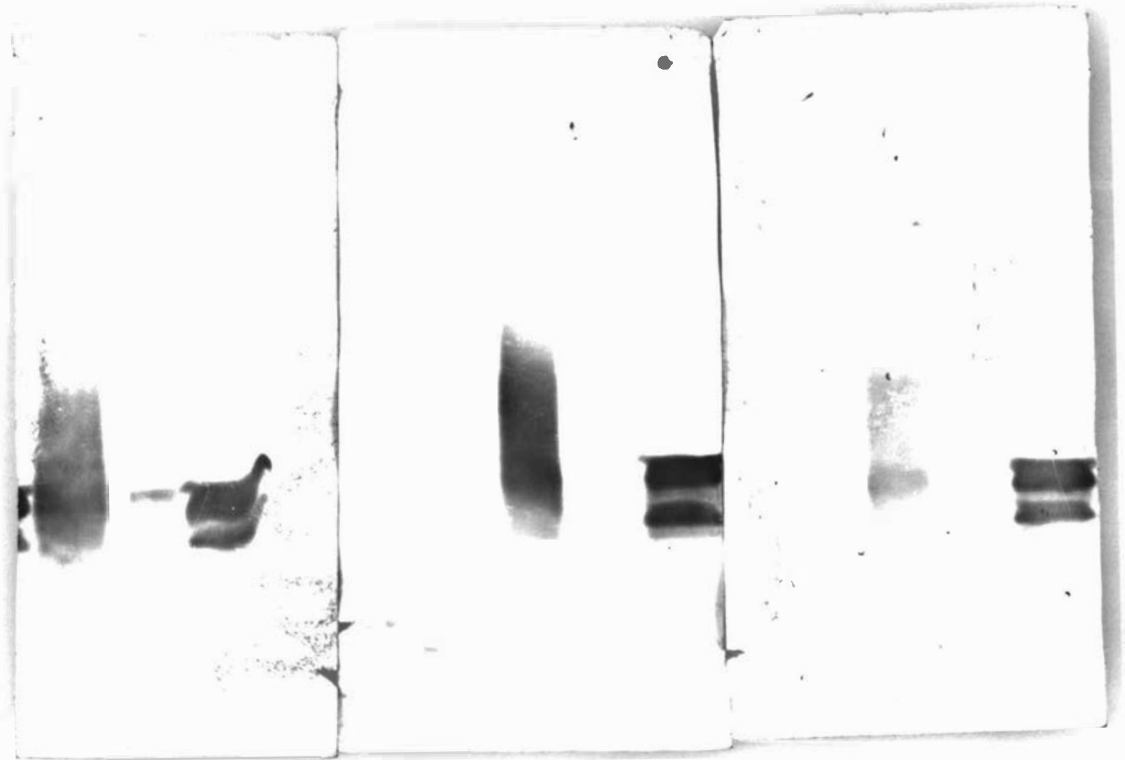
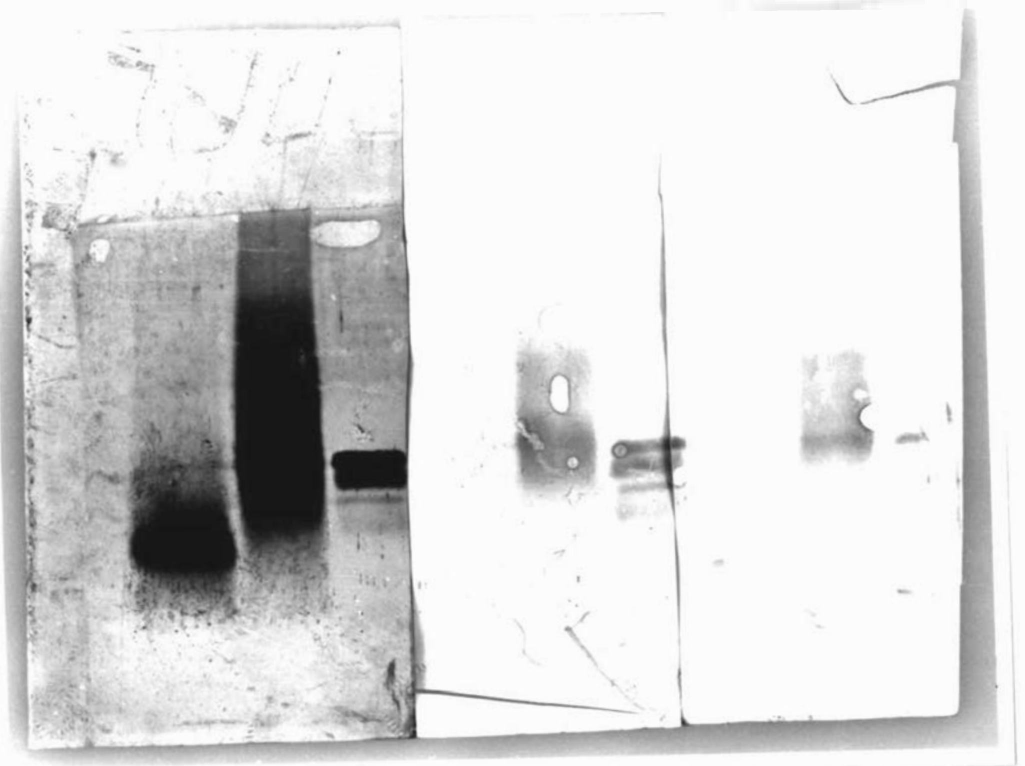


Figure 3.3.49. Western blot stained to detect protein[186]. Lane: 1 90 amino acid peptide. Lane 2: Synthetic 18 kDa protein. Lane 3: Recombinant 18 kDa protein. Lane 4: MW markers as indicated.

Figure 3.3.50. Western blots comparing the synthetic and the recombinant 18 kDa protein. Strip 1 binding L5 monoclonal antibody Strip 2 Binding anti peptide 1-20 antibody, Strip 3 Binding anti peptide 76-95 antibody. Lane 1 in all cases is synthetic 18 kDa protein. Lane 2 recombinant 18 kDa protein.

Figure 3.3.51. Western blots comparing the synthetic and the recombinant 18 kDa protein. Strip 1 binding antipeptide RLP-23 (131-148) antibody. Strip 2: Binding antipeptide 16-35 antibody. Strip 3 Binding antipeptide 1-20 antibody.


1
2
L5
1
2
α1-20
1
2
α76-95


1
2
3
α23
1
2
α16-35
1
2
α1-20

western blots (Figure 3.3.50 and 3.3.51), it can be seen that the entire synthetic protein material, the band concentrated at the 18 kDa mark and the streak on the SDS-PAGE gel is recognised, by the antibodies. The synthetic protein concentration is ten times as much as the recombinant protein suggesting that not all of the material present is protein in the correct form. From this it can be concluded that apart from the requirement of a higher concentration of synthetic product, the synthetic 18 kDa protein is recognised by the antibodies to the 18kDa protein. The fact that the entire synthetic product ie the band and the streak, is recognised biologically, suggests that whatever is causing the differentiation between the banded 18 kDa protein and the rest of the material does not interfere with the biological recognition of the protein. It can therefore be inferred that any protecting groups that maybe present on the threonines or any other amino acid side chains that could still be present, do not block totally the recognition of biologically active epitopes on the protein.

It might be more interesting on the other hand if epitopes requiring some degree of secondary structure requirement were also assayed, as the antibodies are raised against linear epitopes (apart from L5) and the presence of protecting groups may prevent the correct folding of the protein that enables recognition. If the streak is not recognised by these antibodies, then some idea of what is causing the streaking might be gained. The fact that the L5 antibody does recognise the streaked material as well as the band of the synthetic protein suggests that either the L5 antibody recognises a linear determinant on the protein with no secondary structural features playing a part in the antibody recognition sites of the protein or sufficient amino acids are available for the antibody to recognise and bind the synthetic material.

3.3.3.10. Conclusion

The 18 kDa protein has been synthesised correctly with a reasonably high yield of 49% on the resin with an average coupling step of 99.49%. The cleavage and purification of the protein is hampered by the formation of aggregates of the protein probably formed during the evaporation of the HF from the protein-resin. The HF cleavage needs to be done under unusual conditions where the HF vessel is warmed during the evaporation step or, a two step cleavage method such as that described above, to ensure the removal of the side chain benzyl protecting group on threonine. The warming of the HF solution to ambient temperature could however cause the peptide chain to hydrolyse if enough HF is present or affect some of the side chains of the amino acids such as glutamic acid, histidine and tryptophan.

More work specifically targeting the cause of the spreading of the protein on SDS-PAGE gels needs to be done. The streaking could be due to protecting groups still attached

to the protein chain or chemical modification of some of the amino acids. It is noted from the analytical techniques used here, that the synthetic protein consists of protein material that has a wide range of molecular weights, due at least in part to the deletion and premature termination of peptide chains as well as aggregated products.

By using SDS in the HF step the formation of 'aggregates' is prevented to a significant degree, but there are problems handling the protein during purification, due to the high concentration of SDS relative to the crude material. As is shown by the SDS-PAGE gels (Figure 3.3.26), the SDS does appear to cause precipitation of quite a large amount of material at the stacking gel front. Exchanging the SDS if at all possible during a gel filtration, with a volatile detergent may allow the protein to be handled in the final stages of purification by conventional chromatographic techniques such as HPLC and ion exchange. Ideally a combination of the SDS cleavage and the final purification scheme used in cleavage 6 would be extremely useful.

The concentration of SDS to protein was too high to allow the protein to be separated from its deletion and termination products on a gel filtration column. Considering that only a small amount of the SDS dissolves in HF [185] the concentration of SDS should be decreased to provide the critical amount of SDS needed to form micelles (1.8 M)[191].

It is of interest to note that the 90 amino acid protein was purified relatively easily probably due to the precipitation of the aggregated products during dialysis, to give a very tight band of 90 amino acid protein on an SDS gel compared to the diffuse type band of the 18 kDa protein. There was no sign of the precipitation of the 'aggregated' material during the dialysis of the 148 amino acid protein. It may be that in the case of the 18 kDa protein, this precipitate is causing the protein to smear on the SDS-PAGE gels. The formation of the aggregated product could occur during the synthesis of the protein on the resin or by the removal of the side chain protecting groups, under S_n2 cleavage conditions, once the majority of the protecting groups are removed and the DMS and HF have been evaporated from the protein resin. At this point the protein is in a highly concentrated liquid form.

The extra 50 amino acids added to the N-terminus appears to make a difference in the manner in which the protein behaves or binds to the aggregated material. As an example the presence of a tryptophan in the protein relaxin B chain causes the peptide of that chain to form a highly ordered β-sheet . The removal of this tryptophan from the C-terminus of the B chain peptide or replacement of the tryptophan with phenylalanine significantly improved the behaviour of the peptide. It wouldbe interesting to swap the tryptophan for another amino acid such as phenylalanine or naphthylalanine in the 18 kDa protein, to see if this is the reason why the protein's aggregated products are inseparable from the folded protein.

The change of tryptophan for these amino acids would give reasonably close steric properties for the folding of the protein chain. The changes of tryptophan with other amino acids has been done with other proteins in which the tryptophan appeared to significantly affect the properties of the protein isolated from the HF cleavage[161, 193]

The final purification of the protein from cleavage # 4 was of interest as it was shown that the protein could be separated from the aggregated products on the HPLC column which had not been observed previously. It could be that the slightly slower gradient run on the HPLC could have separated the protein from the aggregated material, but from the results of previous analytical HPLC this is unlikely. The separation of unfolded and correctly folded protein has been observed previously on HPLC columns (C4 column at 4°C using isopropanol as the organic modifier)[194]. However the conditions which the preparative HPLC was carried out under should not have provided the discrimination that did occur on the HPLC column (ie C-18 column at 25°C using acetonitrile as the organic modifier). It could well be that the acid-desalting step folowed by a gel filtration in 6M guanidine.HCl under neutral conditions would allow the protein to separate from its aggregated products as indicated by the gel filtration profile (Figure 3.3.42). If this is so, it could become a very useful tool in future purifications of the synthetic protein. The gradual removal of the guanidine.HCl from the protein on the preparative HPLC column could be the step that leads to some of the synthetic protein folding correctly as in the case of the 90 amino acid protein and the clean monomeric protein derived from cleavage 6.

A major drawback in using this procedure developed for the protein from cleavage #4, could be the amount of aggregated product that forms as a result of dissolving the crude protein after the HF cleavage in acetic acid, freezing and then lyophilisation. It may be better to reduce the formation of aggregated products, by dissolving the protein in acetic acid and be loading immediately onto a gel filtration column for desalting. The peak thus isolated should be concentrated, by ultrafiltration, with the resulting retenate being loaded onto another gel filtration column and eluted using 6M guanidine.HCl at pH7.

At the start of this work it was planned that as well as the 50 mers at least one other synthesis greater than 70 amino acids would be done before the total synthesis. In future investigations on the synthetic protein I would suggest that two further syntheses should be carried out, before another total synthesis of the total 148 amino acid protein. A completely separate synthesis of the 90 amino acid protein should be done and another synthesis to give a 120 amino acid fragment should also be done, to determine if the streaking or aggregation is caused by the presence of the amino acids in the N-terminal 50 or 28 residues respectively. It would also be interesting to see if replacement of the tryptophan makes any improvement in the stability and folding of the intact protein. It is felt that in line with the

results from other laboratories [132c,158,170, 171] that at least two more syntheses of the total protein need to be carried out before the chemical synthesis is sufficiently fine tuned to be able to provide biologically active pure synthetic 18 kDa protein. This is assuming that the aggregated products level can be reduced. These syntheses will also need alterations of the synthetic method, which may entail capping protocols and triple couplings particularly near the N-terminal of the protein to provide as near complete 99.9% coupling along the entire 148 amino acids as possible and prevent the multiple biologically active sequences of the protein with an infinite number of small alterations in the % of any one amino acid.

On the other hand due to the expense involved, particularly due to the size of the group of people involved in this laboratory, it may be best if the longer syntheses are left alone and the project changes to target the synthesis of the dominant epitopes or combinations of epitopes. This should be aimed at a synthetic molecule that encompasses the antigenic regions of the 18 kDa protein giving a much smaller synthetic peptide structure which will have the immunological recognition and biological effects of the full length protein. This could be accomplished either by peptides synthesised in a long chain with the epitopes separated by glycines or a Gly-Pro-Gly sequence that would confer a turn in the chain. The other alternative would be to synthesise the amino acids in such a manner as to provide the correct orientation and distances to provide models of the tertiary or secondary structure features of the protein for which a 3 dimensional model of the protein would prove invaluable. For example Tam[195] and Mutter[196] synthesise antigenic epitopes or peptides on an amino acid skeletal framework of lysines to give structures termed Multiple Antigen Peptide (MAP) and Template Assembled Synthetic Protein (TASP) respectively which can be used as carriers to present the peptides to the immune system. The synthesis of combinations of epitopes on such frameworks could provide the basis for a potential vaccine based on epitopes from the 18 kDa protein.

The synthesis of the 18 kDa protein has further illustrated that larger proteins can be synthesised using the advanced synthesis techniques such as those developed by Kent and Clarke-Lewis. The major problem comes in the cleavage and the purification of the synthetic protein. Most strategies in the literature rely on gel filtration and ion exchange chromatography. It appears from the results here that selecting the right conditions is critical in the success of separating the purified protein from aggregated products that associate strongly with the protein.

The SDS polyacrylamide gel of the 'purified' 18 kDa protein compares favourably to the gel of interleukin 1β synthesised by Lobel *et.al* [174]. From the picture published of the gels of the interleukin 1β protein it appears as though the same streaking behaviour that occurred for the 18 kDa protein also occurred for interleukin 1β which was purified by a

one step biotin avidin column. It should be noted that the gels published for the synthesis of interleukin 3 do not show this problem [132c]. The one major difference between interleukin 1β, the 18 kDa protein and interleukin 3 is that interleukin 3 does not have a tryptophan present. It could also be an advantage in the purification of proteins, despite their difficulty in forming the disulphide bond, for the protein to have Cys-Cys disulphide bonding. This would help the protein to fold into the correct tertiary structure. In the 18 kDa proteins case there is no Cys-Cys bridging and thus there is no restriction in the formation of the correct tertiary structure.

The total synthesis of the 18 kDa protein was initially planned to provide usable quantities of 18 kDa protein that were not available by recombinant methods at that time. The synthetic protein has been shown to be recognised by antibodies covering the protein similar to the recombinant protein. Whether the slight variation in the T-cell response is due to the difference in recognition sites on the synthetic protein has yet to be investigated. As yet no study has been made with the 18 kDa protein to see if it induces immunity in leprosy patients so no conclusion can be made on whether any fragments of the 18 kDa protein can be used to confer immunity in potential leprosy sufferers as a potential vaccine. Shortly larger amounts of the overlapping 20mers will be distributed world wide for preliminary *in vitro* studies in this area. Progress has been made in understanding how the 18 kDa protein is recognised by the immune system and how that response could differ in humans, offering an expla nation for the different response to *M.leprae* infection observed in humans..

The major drawback in the synthesis of this protein is the lack of a true functional assay for the protein such as those used for most of the previously synthesised proteins. The antibody recognition and T-cell proliferation assays are not discriminatory enough to determine whether the synthesis has been a complete success in providing a viable functional protein that could replace the recombinant material currently being used in the assays for this protein.

Appendix.

APPENDIX

A.1 INTRODUCTION

Circular Dichroism (CD) is used to predict the secondary structure of peptides and proteins[197]. The basic principle relies on the response of dissolved proteins to polarised radiation with the response dependent on the structural symmetry of the molecule. CD results from the difference in absorbance in left and right polarised light. CD is usually expressed as molar elipticity which is equal to $\Delta\varepsilon$ 3300 where $\Delta\varepsilon$ ($\varepsilon_L - \varepsilon_R$) is the difference in the extinction coefficent between left and right polarised light.

In peptides and proteins there are three kinds of asymmetry that can give rise to optical activity: 1 The primary structure maybe asymmetric (the α carbon atoms of most amino acids have 4 different substituents), 2 A helical secondary structure can result in optical activity for electronic transitions in the main chain or in helical organised amino acid side chains, 3 The folded, tertary structure may lead to the induction of a symmetric or weakly asymMetric group into an asymmetric enviroment.

The three main types of organisation in proteins namely α helix, β sheet and random coil have distinctive CD spectra (figure A.1).

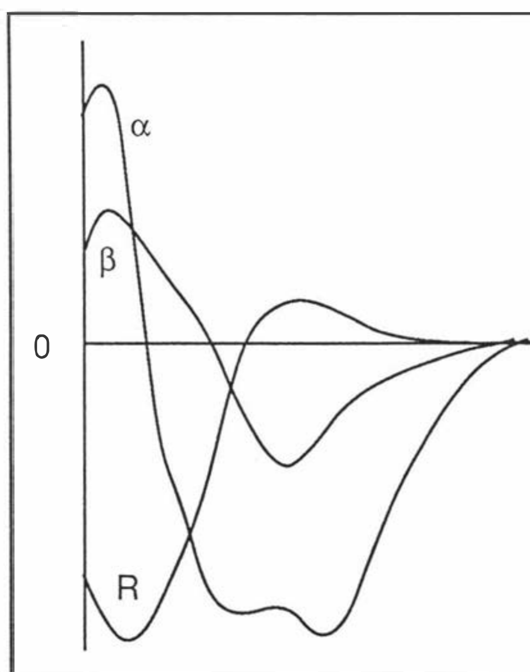


Figure A.1. Profiles of the respective forms of secondary structures of poly lysine in the R (random), α helical and β–sheet conformations[198]

The CD intensities at 207, 290, or 222 nm have been used to determine the percentage of helix in a protein. The CD predicted secondary structures agree in most cases with X-Ray diffraction results. The major advantage of using CD for secodary structure analysis lies in the speed of the method compared with X-ray crystallography or nuclear magnetic resonance analysis.

In this appendix the CD spectra of two peptides are investigated in detail while other peptides are described in less detail. The CD structures were studied in phosphate buffer at pH 7, 0.1% TFA pH2, trifluoroethanol and a 40 mM SDS solution. The size of the peptides, 20 amino acids, means that the peptide is unlikely to form a stable secondary structure in aqueous solution but will show random formation[199]. The trifluoroethanol and SDS micelles serve as model enviroments to induce conformation in peptides. As the 18 kDa protein is a membrane bound peptide, the SDS which acts as a membrane-mimic should be able to stabilise an amphipathic structure such as an α helix.

The aim of this area of work was to analyse the key immunological areas of the 18 kDa protein through analysis of the CD spectra of the 20 mer peptides and the larger 50 and 90 mer peptide fragments, as well as the cloned and synthetic 18 kDa protein. CD has been used previously to study T-cell antigenic regions from whale myoglobin which fold as amphipathic helices[200]. It was hoped that from this study that the CD spectra would also correlate the T-cell stimulatory peptides from the 18 kDa protein with amphipathic helical regions.

A.1.2 EXPERIMENTAL.

Circular Dichroism spectra were recorded using a Jovian Mark V spectropolarimeter equipped with a Sillex (Apple 2 Clone) computer as a data processor. A cylindrical fused quartz cell of 1 mm pathlength was employed. The data are expressed in terms of θ_T the total molar ellipticity (deg x cm^2 x $dmol^{-1}$). The number of moles was not accurately determined (see reason below). Trifluoroethanol (TFE) was purchased from Aldrich (Germany), and SDS from BDH Chemicals (UK). All water was of Milli Q quality. All spectra are recorded over 190 or 200-250 nm at a sensitivity setting of 10^{-6}. Most spectra are an average of three cycles and base line corrected by subtraction of the solvent spectra for the peptide.

Spectra were analysed by the computer program CONTIN version November 1980 of Provencher[201] run on a micro Vax computer. This program takes the parameters of the CD spectra and calculates the % of α helix and β sheet formation from the CD spectra

parameters. β turns and random structure formation are calculated with a lesser degree of accuracy.

A.1.3 RESULTS AND DISCUSSION.

Two regions of peptide were chosen for the initial CD investigation of their secondary stucture. These two peptides were SLP-9 (61-80) and SLP-11 (91-110) which were predicted from the secondary structure prediction plots to have contrasting secondary structure composition (figure 2.3.13). Peptide 61-80 was predicted to consist mainly of β turnswith some β sheet content while peptide 91-110 was predicted to contain a 3 turn α helix.

Peptide 61-80 at pH 7 showed little propensity for formation of any secondary structural element apart from random coil (figure A.2). The TFE peptide solution showed a slight change in the spectra with slight indication of secondary stucture formation (figure A.3). The SDS peptide solution showed a significant change in the CD spectra (figure A.4) with apparent secondary structure formation. From the CONTIN program the peptide appears to be mainly β sheet. The peptide composition in the SDS solution was 12% helix, 45% β sheet, 23% β sheet and the remaider 19%.

Peptide 91-110 showed a greater propensity to form secondary stucture in solution. At pH7 the peptide appeared to form random structures while at pH2 there was some indication of secondary structure formation (figure A.5). In TFE and SDS solutions the peptide adopted a secondary structure consistent with the formation of an α helix (figure A.6). When the computer program CONTIN was applied to the values collected from the CD spectra for the TFE solution, the secondary structural content was found to be 32% α helix, 40% β sheet, and 28% β turn. At pH 2 with 50% TFE, the % of α helix goes up to 49% with β turn dropping to 11% and the β sheet content remaining the same. This effect of enhancing the CD spectra at pH 2 probably due to the breaking of the H bonds which would cause the peptide to self aggregate giving β sheets.

Peptide 91-115 and peptide 1-20 in 80% TFE (figure A.7) also showed a propensity to form secondary structure similar to peptide 91-110 though not to the same degree as with peptide 91-110 which was analysed in neat TFE. Applying the CONTIN program to peptide 1-20 the peptide consisisted of 10% α helix 58% β sheet and 27% β turn.

Figure A.2. CD spectra of peptide 61-80 a) PO_4 pH 7. b) 0.1% TFA pH2. c) Neat TFE. d) 0.1% TFA solution from b with 50% TFE. e) 40 mM SDS. The peptide showed no propensity to form secondary structures as predicted from theT-cell assay and the secondary structure prediction program.

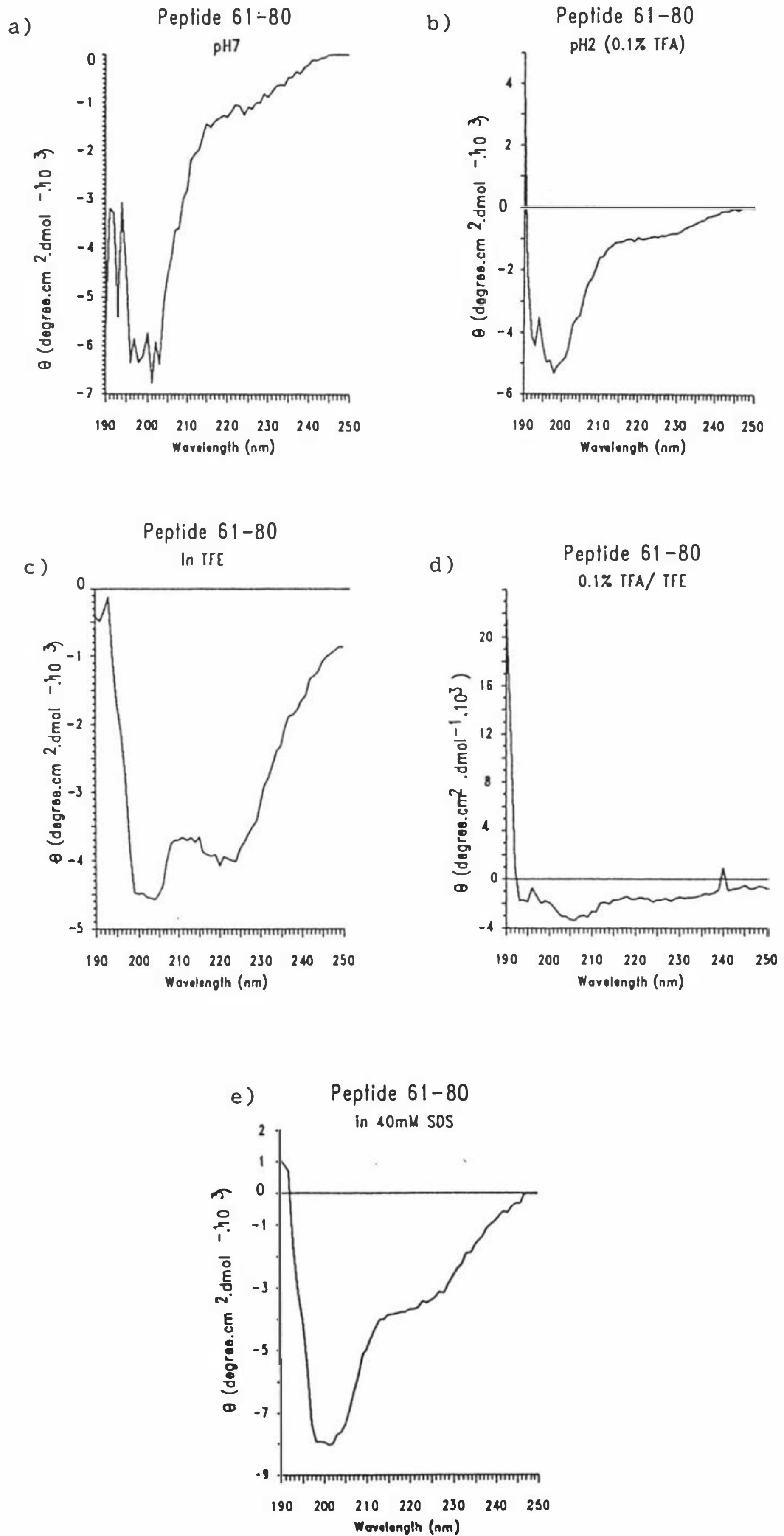
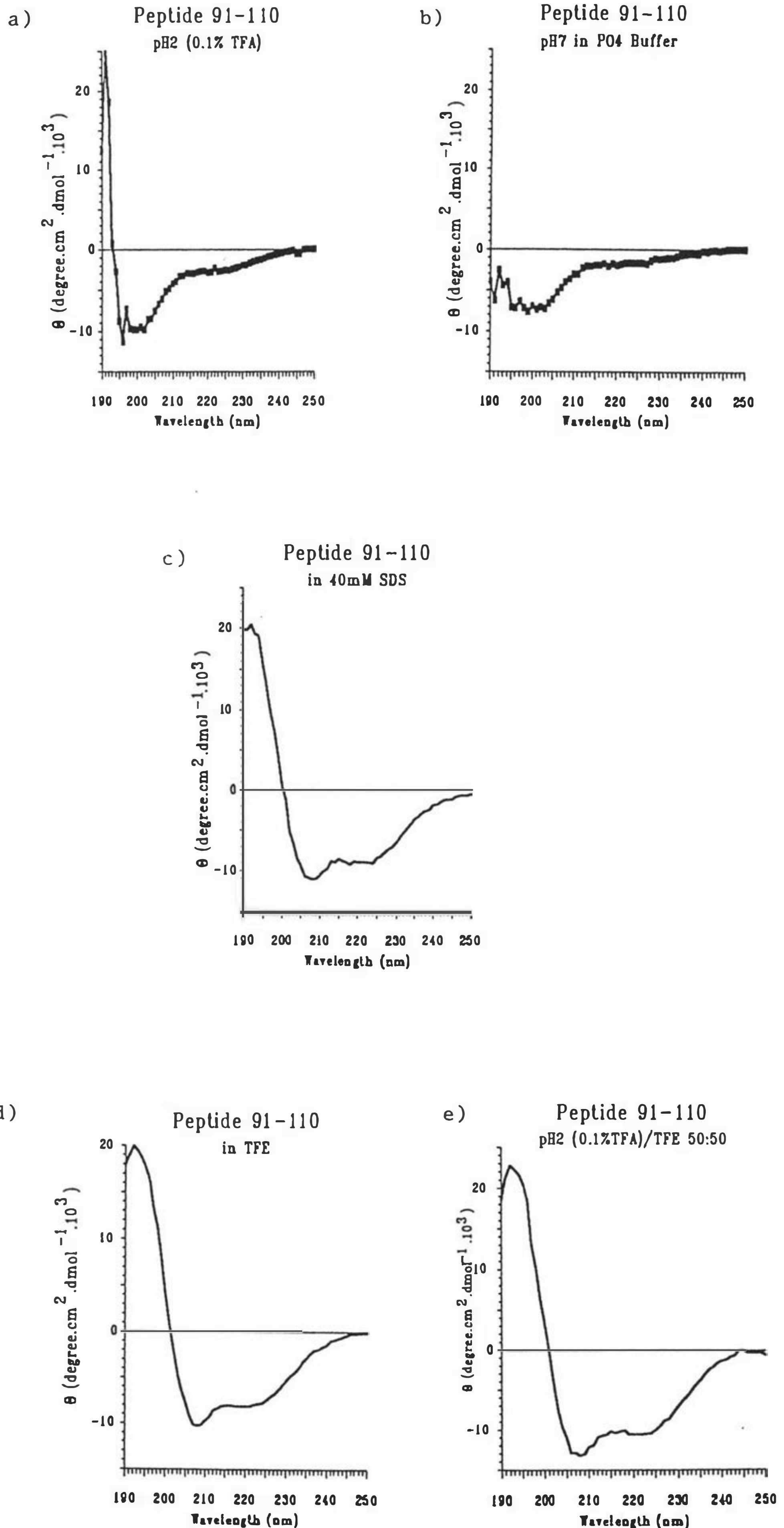

a)
Peptide 61-80
pH7
θ (degree.cm 2.dmol -1.10 3)
Wavelength (nm)
b)
Peptide 61-80
pH2 (0.1% TFA)
θ (degree.cm 2.dmol -1.10 3)
Wavelength (nm)
c)
Peptide 61-80
In TFE
θ (degree.cm 2.dmol -1.10 3)
Wavelength (nm)
d)
Peptide 61-80
0.1% TFA/ TFE
θ (degree.cm2 .dmol-1.103)
Wavelength (nm)
e)
Peptide 61-80
in 40mM SDS
θ (degree.cm 2.dmol -1.10 3)
Wavelength (nm)

Figure A.3. CD spectra of peptide 91-110. a) 0.1% TFA pH2, b) PO_4 pH7, c) 40 mM SDS, d) neat TFE, e) 0.1% TFA solution from a) with 50% TFE. The peptide form stable secondary structure in SDS and TFE solutions.


a)
Peptide 91-110
pH2 (0.1% TFA)
b)
Peptide 91-110
pH7 in PO4 Buffer
c)
Peptide 91-110
in 40mM SDS
d)
Peptide 91-110
in TFE
e)
Peptide 91-110
pH2 (0.1%TFA)/TFE 50:50
θ (degree.cm 2 .dmol -1.10 3)
Wavelength (nm)

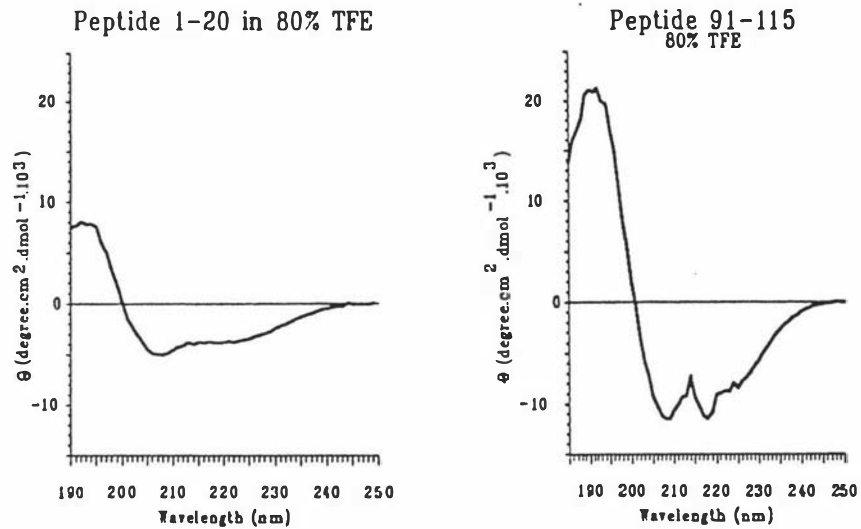


Figure A.4. a) CD spectra of peptide 1-20 in 80% TFE. Peptide would not dissolve in neat TFE. b) CDspectra of peptide 91-115 in 80% TFE which formed a stable secondary structure similar to peptide 91-110.

CD spectra were also run on two 50 mer peptides 101-148 and 50-100 (figure A.8). These two peptides showed only random structure in phosphate solution at pH 7 as expected from the SDS-PAGE gels which showed considerable aggregated material. The 18 kDa recombinant protein also showed (figure A.9) only random structure formation in phosphate buffer. However on the addition of TFE to give a 50% solution the 18 kDa protein did give an indication of formation of secondary structure.

All of the above sprectra were carried out on an instrument which was used on an irregular basis, kept at the Dairy Reseach Institute in Palmerston North. The instrument was not able to be calibrated correctly with the voltage settings for the instrument at the lower end of the wavlength scale around 200 nm which was incorrectly set. The crystal used to polarise the light had also just been replaced. Therefore any structural characteristics are at best only indications of the true secondary structure content of the peptides.

Running the standard d-10-camphorsulfonic acid in water it was discovered that the ratio of $\theta_{192.5nm}/\theta_{290.5nm}$ was 1.813 instead of the recommended 2.00[202]. Despite numerous attempts at correcting this ratio this was the best result that could be obtained. The CD spectra of a standard protein myoglobin was run and the values transferred to the CONTIN program. The secondary structure of myoglobin was found to be 55% α helix,

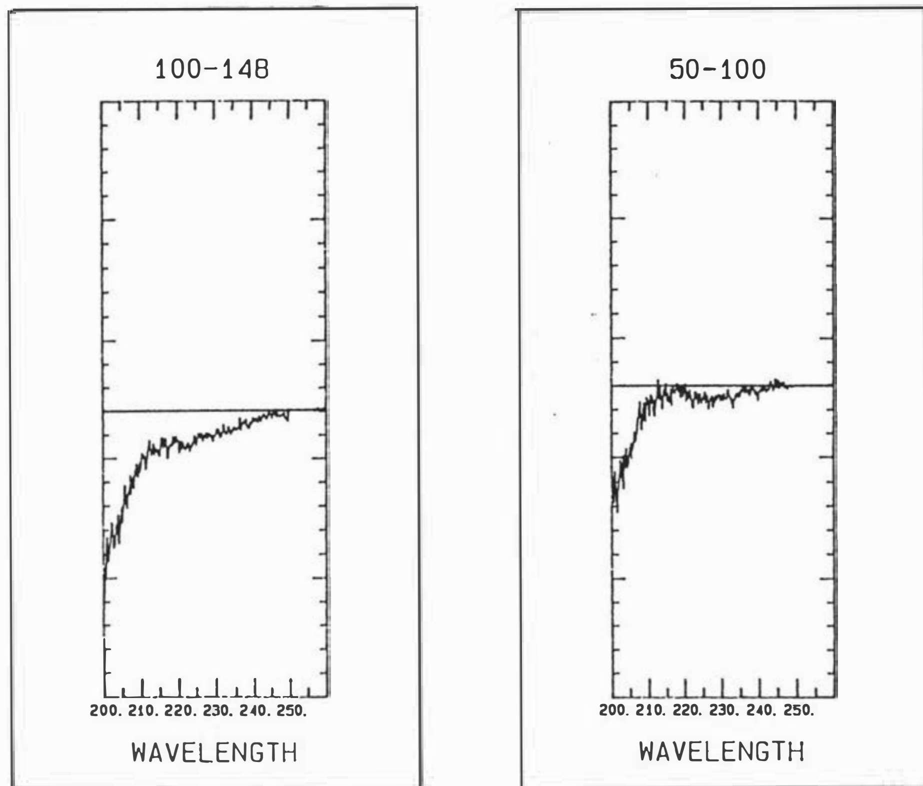


Figure A.5. CD spectra of peptides 101-148 and 50-100 both in PO_4 pH7 showing random structure only.

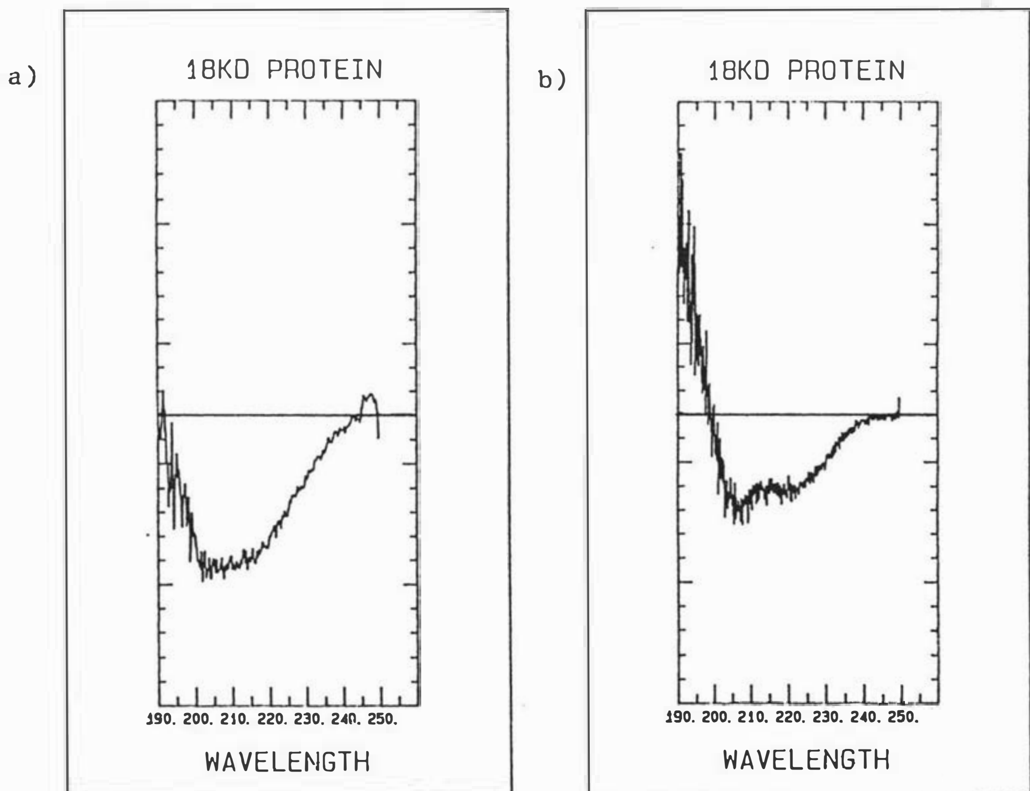


Figure A.6. a) CD spectra of recombinant 18 kDa protein in PO_4 at pH7 showing random structure formation. b) CD spectra of recombinant 18 kDa protein in PO_4/TFE 50:50 vol:vol.

11% β sheet , 13% β turn with the remainder 22%, compared to the literature values of 74% α helix, 0.02% β sheet, 0.05% β turn with the remainder 19%. Despite this it was decided to proceed and run the CD spectra for as many peptides as possible to be able to get at least an overall preliminary picture of the secondary structure content of the 18 kDa protein.

Unfortunately before the study could proceed beyond the peptides described above the voltage readings shifted and were not able to be readjusted. This gave considerable random noise in the region 190-220 nm making any further reasonably accurate measurements impossible. Attempts by the Dairy Reseach Institute to rectify the problems with the instrument has proved futile.

A.1.4. CONCLUSION.

As even minor differences between CD spectropolarimeters can produce large differences in the % secondary structure content of proteins[203], the values measured and contained in this work are not to be used as an absolute indication of the % secondary structure of the peptides.

In the peptides and the protein investigated, it can be concluded that in the presence of enviroments that induce conformation of peptides, the peptides do form secondary structures that could mimic the conformation of the protein in its native state. Peptide 91-110 does appear to have induced helix formation while the peptide 61-80 shows little sign of the formation of the helices predicted in the computer secondary structure prediction programs. It is hoped that this preliminary study can be continued at a later date once the problems with the instrument are overcome.

References.

REFERENCES

1 Maurice, J., *New Scientist* Feb 4th (1989) p 48-52.

2 *Tropical Diseases: Progress in International Research, 1987-1988.* Ninth Programme Report of the UNDP/World Bank/WHO Special Programme for Reseach and Training in Tropical Diseases. WHO Geneva (1989) p 93-100.

3 Hansen, G.H.A., *Norsk Magazin for Laegervindenskaben.* (1874) Suppl 4: 1-88.

4 Shepard, C.C., *Journal of Experimental Medicine.* 112 445-454 (1960).

5 Johnson, P.A.S., *International Journal of Leprosy.* 55 535-547 (1987).

6 Kichheimner, W.F. and Storrs, E.E., *International Journal of Leprosy.* 39 693-702 (1971).

7 Wolf, R.H., Gormus, B.J., Martin, L.N. and Baskin, G.B., *Science.* 227 529-531 (1985).

8 Thangaraj, R.H. and Yawalker, S.J., in *Leprosy for Medical Practitioners and Paramedical Workers.* Ciba Geigy, Basle Switzerland. (1986).

9 Baohong, J.. *Leprosy Review* 56 265-278 (1985).

10 Lauras, C.C.G., Grosset, J.H., Constant Desportes, M. and Brucher, G., *International Journal of Leprosy.* 52 101-102 (1984).

11 Warndorff-van Diepen, T.. *International Journal of Leprosy.* 50 139-142 (1982).

12 Moudigil, K.D., Mukherjee, R., Itty, M.B. and Talwar, G.P., in *Progress in Vaccinology. Vol 2.* Talwar G.P.,Ed, Springer-Verlag N.Y.. (1989) Pg 347-354.

13 Young, D.B., Fohn M,.J. and Khanolkar, S.R., *Leprosy Review.* 56 193-198 (1985).

14 Guelpa-Lauras, C.C.G., Perani, E.G., Giroir, A.M. and Grosset, J.H.. *International Journal of Leprosy.* 55 70-77 (1987).

15 Seydel, J.K., Rosefeld, M. and Satish, M., *Leprosy Review.* 57 Suppl 3 235-253 (1986).

16 Rees, R.J.W., *British Medical Bulletin.* 44 560-564 (1988).

17 Mitchison D.A., Ellard G.A. and Grosset J.. *British Medical Bulletin.* 44 757-774 (1988).

18 Convit, J., Aranazu, M. and Ulrich, M., *International Journal of Leprosy.* 50 415-424 (1982).

19 Fine, P.E.M., *Reviews of Infectous Diseaeses.* 11 Suppl 2 S353-S359 (1989).

20 Shinnick, T.M., Sweetser, D., Thole, J.,van Embden, L. and Young, R.A., *Infection and Immunity.* 55 1932-1935 (1987).

21 Mustafa, A.S., in *Progress in Vaccinology Vol 2.* Talawar, G.P., Ed, Springer-verlag N.Y. (1989) p 323-334.

22 Deo, M.G., Bapat, C.V. and Bhalerao, V., *International Journal of Leprosy.* 51 540-549 (1983).

23 Jacobs, W.R., Tuckman, M. and Bloom, B.R., *Nature.* 327 532-535 (1987).

24 Singh, N.B., Lowe, A.C.R.E., Rees, R.J.W. and Colston, M.J., *Infection and Immunity.* 57 653-655 (1989).

25 Lamb, F.I., Singh, N.B. and Colston, M.J., *The Journal of Immunology.* 144 1922-1925 (1990)

26 Bloom, B.R., *The Journal of Immunology.* 137 i-ix (1986)

27 Bloom, B.R. and Mehra, V.. *Immunological Reviews.* 80 5-30 (1984)

28 a) Hunter, S.W. and Brennan, P.J., *Journal of Bacteriology* , 147 728-735 (1981).
b) Hunter, S.W., Fujiwara, T. and Brennan, P.J., *The Journal of Biological Chemistry.* 257 15072-15078 (1982).

29 Fuliwara, T., Hunter, S.W. and Cho, S.N.. *Infection and Immunity.* 43 245-252 (1984).

30 Levis, W.R., Meeker, H.C., Schuller-Levis, G., Sersen, E., Brennan, P.J. and Fried, P., *The Journal of Biological Chemistry,* 156 763-769 (1987).

31 Molloy, A., Gaudernack, G., Levis, W.R., Cohn,Z.A and Kaplan, G., *Proceedings of the National Acadamy of Science U.S.A.* 87 973-977 (1990).

32 Mehra, V., Brennan, P.J., Rada, E., Convit, J. and Bloom, B.R., *Nature.* 308 194-196 (1984).

33 Rojas-Espinosa, O., *Journal of Leukocyte Biology.* 43 468-470 (1988).

34 Chan, J., Fujiwaka, T., Brennan, P., McNeil, M., Turlo, S.J., Sibille, J.C., Snapper, M., Aisen, P. and Bloom, B.R., *Proceedings of the National Acadamy of Science U.S.A.* 86 2453-2457 (1989).

35 Engers, H.D., *et.al Infection and Immunity.* 48 603-605 (1985).
Engers, H.D., *et.al Infection and Immunity.* 51 718-720 (1986).

36 Young, R.A., Mehra, V., Sweetser, D., Buchanan, T., Clarke-Curtis, T., Davis, R.W.and Bloom B.R., *Nature* 316 450-452 (1985) .

37 Rumschalg, H.S., Shinnick, T.M. and Cohen, M.L., *Journal of Clinical Microbiology* 26 2200-2202 (1988).

38 Young, D., Lathigra, R., Hendrix, R., Sweetser, D. and Young, R.A., *Proceedings of the National Acadamy of Science U.S.A* . 85 4267-4270 (1988).

39 a) Young, D., Latigra, R., Hendrix, R., Sweetser, D. and Young, R.A., *Proceedings of the National Acadamy of Science U.S.A* . 85 4267-4270 (1988).
b)Young, R.A., *Annual Reviews of Immunology.* 8 401-420 (1990).

40 Britton, W.J., Hellquist, L., Basten, A. and Raison, R.L., *Journal of Immunology.* 135 171-4177 (1985).

41 Britton, W.J., Hellquist, L., Basten, A. and Inglis, A.S., *Journal of Experimental Medicine.* 164 695-708

42 Brown, W.B., Larrabee, W.B. and Kim, P.S., *Leprosy Reviews.* 57 157-162 (1986).

43 Meeker, H.C., Williams, D.L., Anderson, D.C., Gilis, T.P., Schiller-Levis, G. and Levis, W.R., *Infection and Immunity.* 57 3689-3694 (1989).

44 Anderson, D.C., Van Schooten, W.C.A., Barry, M.E., Jansson, A.A.M., Buchanan, T.M. and deVries, R.R.P., *Science.* 242 259-261 (1988).

45 Cox, J.H., Ivanyi, J., Young, D.B., Lamb, J.R., Syrell, A.D. and Francis, M.J., *European Journal of Immunology.* 18 2015-2019 (1988).

46 Klaster, P.R., Van Rens, M.M. and Eggeltte, T.A., *Clinical and Experimental Immunology.* 56 537-544 (1984).

47 Van Schooter, W.C.A., Ottenhoff, T.H.M., Klaster, P.R., Thole, J., De Vries, R.R.P. and Kolk, A.H.J., *European Journal of Immunology.* 18 849-854 (1988).

48 Hartskeel, R.A., De Witt, M.Y.L. and Klaster, P.R., *Journal of General Microbiology.* 135 2357-2364 (1989).

49 Abou-Zeid, C., Ratliff, T.L., Wiker, H.G., Harboe, M., Bennedson, J. and Rook, G.A.W., *Infection and Immunity.* 56 3046-3051 (1988).

50 Young, D.B., Fohn, M.F., Khandkar, S.K. and Buchanan, T.M., *Clinical and Experimental Immunology.* 60 546-552 (1985).

51 Mustafa, A.S., Gill, H.K., Nerland, A., Britton, W.J., Mehra, V., Bloom, B.R. Young, R.A. and Godal, T., *Nature* 319 63-68 (1986).

52 Booth, R.J., Harris, D.P., Love, J.M. and Watson, J.D., *The Journal of Immunology.* 140 597-601 (1988).

53 Nerland, A.H., Mustafa, A.S., Sweetser, D., Godal, T. and Young, R.A., *Journal of Bacterology.* 170 5919-5921 (1988).

54 Doherty, T.M., Booth, R.J., Love, S.G., Gibson, J.J., Harding, D.R.K. and Watson, J.D., *The Journal of Immunology.* 142 1691-1695 (1989).

55 Harris, D.P., Backstrom, B.J., Booth, R.J., Love, S.G., Harding, D.R. and Watson, J.D., *The Journal of Immunology.* 143 2006-2012 (1989).

56 Oftung, F., Shinnick, T.M., Mustafa, A.S., Lundin, K.E.A., Godal, T. and Nerland A.H., *The Journal of Immunology.* 144 1478-1483 (1990)

57 Hunter, S.W., McNeil, M., Modlin, R.L., Mehra, V., Bloom, B.R. and Brennan, P.J., *The Journal Of Immunology.* 142 2864-2872 (1989).

58 Mehra, V., Bloom, B.R., Torigian, V.K., Mandich, D., Ghreichel, M., Young, S.M.M., Salgame, P., Convit, J., Hunter,S.W., McNeil, M., Brennan, P.J., Rea, T.H. and Modlin, R.L., *The Journal Of Immunology.* 142 2873-2878 (1989).

59 a) Willams, D.L., Gillis,, T.P., Booth R.J., Looker, D. and Watson, J.D., *Journal of Infectious Diseases,* 162 193-200 (1990).
b) Hackel, C., Hovards, S., Portaels, F., van Elsen, A., Herzog, A. and Bollen, A., *Molecular and Cellular Probes* 4 205-210 (1990).

60. Fisher, E., *Ber dtsch Chem Ges* 39 530 (1906)

61. Curtis, T., *J. Prakt. Chem.* 24 239 (1881).

62. Brittle, J.L., Houghton, R.A., Alexander, H., Shinnick, T.M., Sutcliffe, J.G., Lerner, R.A., Rowlands, D.J. and Brown F., *Nature* 298 30-33 (1982)

63. Gerin, J.L., Alexander, H., Shih, J.W., Purcell, R.H., Dapolito, T., Engle, R., Green, N., Sucliffe, J.G., Shinnick, T.M. and Lerner, R.A., *Proceedings of the National Acadamy of Science. U.S.A.* 80 2365-2369 (1983).

64. Jacob, C.O., Sela, M., and Arnon, R., *Proceedings of the National Acadamy of Science. U.S.A.* 80 7611-7615 (1983).

65. Audibert, F., Jolivet, M., Chedid, L., Arnon, R. and Sela, M., *Proceedings of the National Acadamy of Science. U.S.A.* 79 5042-46 (1982).

66. Dale, J.B.and Beachey, E.H., *Journal of Experimental Medicine.* 156 1165-1176 (1982).

67. Wilson, I.A., Niman, H.L., Houghton, R.A., Cherenson, A.R., Connolly, M.L. and Lerner, R.A., *Cell* 37 767-778 (1984).

68. Zavala, F., Masuda, A., Garves, P.M., Nussenzweig, V. and Nussenzweig, R.S., *Journal of Immunology* 135 2790-2793 (1985).

69. Emini, E.A., Jameson, B.A. and Wimmer, E., *Nature* 304 699-703 (1983).

70. Kennedy, R.C., Henkel, R.D., Paoletti, D., Allen, J.S., Lee, T.H., Essex, M. and Dreesman, E.R., *Science* 231 15556-1559 (1986).

71. Franks, F., in *Characterisation of Proteins*. Ed Franks, F., The Humana Press. Clifton, N.J.. (1988).

72. Barany, G., Kneib-Cordonier, N. and Mullen, D.G., *International Journal of Peptide and Protein Research.* 30 705-739 (1987)

73. Beacham, J., Dupuis, G., Finn, F.M. Storey, H.T., Yanaihara, C., Yanaihara, N. and Hofmann, K., *Journal of the American Chemical Society* 93 5526-5539 (1971)

74. Kai, Y., Nakashita, M., Suwa, K., Sano, A., Ikeda, Y., Ueki, Y. and Ono, K., in *Peptide Chemistry 1988* . Ueki M., Ed, Protein Reseach Foundation. Japan.(1989) p 91-96.

75. Morimoto, H., Fuji, N., Otaka, A., Kuramochi, K. and Yajima, H., in *Peptide Chemistry 1988* . Ueki M., Ed, Protein Reseach Foundation. Japan. (1989) p 211-214.

76. Merrifield, R.B., *Journal of the American Chemical Society* 85 2149-2154 (1963).

77. Cautrecacus, P., *Journal of Biological Chemistry* 245 3059-3065 (1971)

78. Axin, R., Porath, J. and Ernback, S., *Nature* 214 1302-1304 (1967)

79. Caruthers, M.H., *Science* 230 281-285 (1985).

80 Yajima, H., Fujii, N., Ogawa, H. and Kawatani, H., *Journal of the Chemical Society. Chemical Communications,* 107-108 (1973).

81. Merrifield, R.B., and Littau, V., in *Peptides 1968* Bricas E., Ed. North Holland, Publishers, Amsterdam 1968. p 179-182

82. Sarin, V.K., Kent, S.B.H. and Merrifield, R.B., *Journal of the American Chemical Society* 102 5463-5470 (1980).

83 Merrifield, R.B. and Barany in *The Peptides: Analysis and Synthesis and Biology,* Gross, E. and Meienhofer, J.,Ed, Academic Press. NY. 1981 Vol 2, p 1-235.

84 Yamashiro, D., Blake, J. and Li, C.H., *Tetrahedron Letters,* 1469-1472 (1976).

85 a) Tam, J., *Peptides Structure and Function* Ed Deber,C., Hurby, V. and Koppel, K., Pierce Chemical Company Rockford Ill, (1985) p 423-425.
b) Tam, J., *International Journal of Peptide and Protein Research,* 29 421-431 (1987)

86 Toniolo, C., Bonora, G., Moretto, V. and Bodanszky, M., *International Journal of Peptide and Protein Research.* 25 425-430 (1985).

87. Mitchell, A.R., Erickson, B.W., Ryabtseu, M.N., Hodges, R.S. and Merrifield, R.B., *Journal American Chemical Society* 98 7357-7362 (1976).

88. Feinberg, R.S. and Merrifield, R.B.. *Tetrahedron* 30 3209-3212 (1974).

89. Mitchell, A.R., Kent, S.B.H, Engelhard, M. and Merrifield, R.B., *Journal of Organic Chemistry* 43 2845-2852 (1987).

90. Belleau, B. and Malek, G.. *Journal of the American Chemical Society* 90 1651-1652 (1968)

91. Hudson, D.. *Journal of Organic Chemistry.* 53 617-624. (1988).

92. Clarke-Lewis, I. and Kent, S.B.H.. in *Receptor Biochemistry and Methodology: The use of HPLC in Protein Purification and Characterisation.* Kerlavage, A.R. Series Ed Venter, J.C. and Harrison, L.C., Ed, Alan R. Liss, NY.

93. Merrifield, R.B., Singer, J. and Chait, B.T.. *Analytical Biochemistry.* _ 399-414 (1988)

94. Mojsov, S., Mitchell, A.R. and Merrifield, R.B.. *Journal of Organic Chemistry.* 45 555-560 (1980).

95. Koenig, W. and Geiger, R., in *Peptides 1969.* Scoffene, E., Ed, Amsterdam North Holland Publishers. (1971) p 17-22.

96. Bodanszky, M and Bodanzsky, A., in *The Practice of Peptide Synthesis*, Springer-Verlag, Berlin, (1984).

97. Castro, B., Dormoy, J.R., Evin, G. and Selve, C., *Tetrahedron Letters* 1219-1222 (1975).

98. Feiberg, R.S. and Merrifield, R.B., *Journal of the American Chemical Society,* 97, 3485-3496 (1975).

99. Birr C., Lochinger W., Stahnke G. and Lang P.. *Justus Liebiegs Annals of Chemistry.* 763 162-172. (1972).

100. Sandberg, B.E. and Ragnarsson, U., *International Journal of Peptide and Protein Research.* 7 503-4 (1975).

101. Matsueda, G.R. and Stewart, J.M., *Peptides. Chemistry Structure and Biology.* Walter, R. and Meienhofer,J., Ed Ann Arbor. M.I. (1975) p 333-339.

102. Wang, S.S., Chen, S.T., Wang, K.T. and Merrifield, R.B.. *International Journal of Peptide and Protein Research.* 30 662-667 (1987).

103. Rosen, O., Rubinrautis, S. and Fridkin, M.. *International Journal of Peptide and Protein Reseacrh* 35 545-549 (1990)

104. Barany, G. and Merrifield, R.B., *Journal of the American Chemical Society.* 99 7363-7365 (1977).

105. Barlos, K, Papaionnou, D. and Theodropoulos, D., *Journal of Organic Chemistry,* 47 1324-1326 (1982).

106. Barany, G. and Albericio, F., *Journal of the American Chemical Society.* 107 4936-4942 (1985).

107. Carpino, L.A. and Han, C.C., *Journal of the American Chemical Society,* 92 5748-5749 (1970).

108. Atherton, E., Fox, H., Harkiss, D. and Shepard, R.C.,*Journal of the Chemical Society. Chemical Communications.* 539-540 (1978).

109. Fields, G.B. and Noble, R.L., *International Journal of Peptide and Protein Research* 35 161-214 (1990).

110. Arshady, R., Atherton, E., Gait, M.J., Lee, K. and Shepard, R.L., *Journal of the Chemical Society. Chemical Communications.* 423-425 (1979)

111. Atherton, E., Brown, E. and Shepard, R.L., *Journal of the Chemical Society. Chemical Communications.* 1151-1152 (1981).

112. Small, P.W. and Sherrington, D.C., *Journal of the Chemical Society. Chemical Communications.* 1589-1591 (1989).

113. a) Kent, S.B.H., Hood, L.E., Beilan, H., Bridghan, J., Marriot, M., Meister, S., and Geiser, T., in *Peptide Chemistry 1984, Proceedings of the Japanese Peptide Symposium.* Izumiya N., Ed, Protein Research Foundation. Osaka. p 217-222.
b) Kent, S.B.H., Hood, L.E., Beilan, H. and Meister, S., *Peptides 1984: Proceedings of the 18th European Peptide Symposium, Sweden, 1984.* Ragnarsson, U., Ed,. Almquist and Wicksell, Stockholm, 1984. p 185-188

114. Applied Biosystems Model 430A Peptide Synthesiser Users Manual Version 1.3B (1988), (Part No. 900066)

115. Houghten, R.A., *Proceedings of the National Acadamy of Science. U.S.A.* 82 5131-5135 (1985).

116. Geysen, H.M., Meleon, R.H. and Bartling, S.J., *Proceedings of the National Acadamy of Science. U.S.A.* 81 3998-4002 (1984).

117. Kent, S.B.H. and Merrifield, R.B., *International Journal of Peptide and Protein Research.* 22 57-65 (1983).

118. Hancock, W.S., Prescot, D.J., Vagelos, P.R. and Marshall, G.R., *Journal of Organic Chemistry.* 38 774-781 (1973).

119. Kaiser, E., Colescott, R.L., Bossinger, C.D. and Cook, P.I., *Analytical Biochemistry.* 34 595-598 (1973).

120. Sarin, V.K., Kent, S.B.H., Tam, J.P. and Merrifield, R.B., *Analytical Biochemistry.* 117 147-153 (1981).

121. a) Gisen, B.F., *Analytical Chimie Acta.* 58 248-249 (1978).
b) Hancock, W.S., Battersby, J.E. and Harding, D.R.K., *Analytical Biochemisrty* 69 497-503 (1975)

122. Westall, F.C., Scotchler, J. and Robinson, A.B., *Journal of Organic Chemistry* 37 3363-3365 (1972).

123. Niall, H.D., Tregear, G.W. and Jacobs, J., in *Chemistry and Biology of Peptides.* Ed. Meienhofer, J. Ann Arbor. M.I. (1972) p 695-699.

124. a) Matsueda, G.R., Harber, E. and Margolies, M.N., *Biochemistry* 20 2571-80 (1981).
b) Hewick, R.M., Hunkapiller, M.W., Hood, L.E. and Dreyer, W.J., *The Journal of Biological Chemistry.* 256 7990-7997 (1981)

125. Tam, J.P., Heath, W.F. and Merrifield, R.B., *Journal of the American Chemical Society.* 105 6442-6455 (1982).

126. Tam, J.P., Heath, W.F., and Merrifield, R.B., *Journal of the American Chemical Society* 108 5242-5251 (1986).

127. Gehrke, C.W., Wall, L.L., Absheer, J.S., Kaiser, F.E. and Zumwalt, R.W.. *Journal of the Association of Official Analytical Chemists.* 68 811-821 (1985).

128. Male, D., Champion, B. and Cooke, A., in *Advanced Immunology.* Grower Medical Publishers. London (1987)

129. Delisi, C. and Berzofsky, J.A., *Proceedings of the National Acadamy of Sciences. USA* 82 7048-7052 (1985).

130. Doherty, J.M., Backstrom, B.T., Prestidge, R.L., Love, S.G., Harding, D.R.K. and Watson, J.D., *Journal of Immunology,* (in press) (1991)

131. Backstrom, B.T., Harris, D.P., Doherty, M.T., Prestidge, R.L., Love, S.G., Harding, D.R. and Watson, J.D., *Journal of Immunology,* (in press). (1991).

132. a) The vessel was designed by Steven Clarke and machined by Frank Ostrander, both of the Biology Division of the California Institute of Technology, L.A..
b) Kent, S.B.H. and Clarke-Lewis, I., in *Synthetic Peptides in Biology and Medicine,* Alitalo, P., Partanen, P., and Vaheri, A., Ed Elsevier, Amsterdam 1985 p 25-57.
c) Clarke-Lewis, I., Abersold, R.A., Ziltner, H., Schrader, J.W., Hood, L.E. and Kent, S.B.H., *Science.* 231 134-139 (1986).
d) Kent, S.B.H., Parker, K.F., Schiller, D.L., Woo, D.D.L., Clarke-Lewis, I. and Chait, B.T., in *Peptides Chemistry and Biology, Proceedings of the Tenth American Peptide Symposium.* , Marshall, G.R., Ed ESCOM, Lieden 1988 p 173-178
e) Kent, S.B.H., *Annual Review of Biochemistry.* 57 957-984 (1988).

133. Goren, H.J. and Fridkin, M., *International Journal of Peptide and Protein Research.* 11 1-8 (1978).

134. *Introduction to Cleavage Techniques: Stratagies in Peptide Synthesis.* Applied Biosystems Inc. (1990).

135. *Solid Phase Peptide Synthesis.* Stewart, J.M. and Young, J.D. Second Ed Pierce Chemical Company, Rockford Illinois (1984).

136. Penke, B., Ferenczi, R. and Kovas, K., *Analytical Biochemistry.* 60 45-50 (1974).

137. Matsueda, G.R., *International Journal of Peptide and Protein Research.* 20 26-34 (1982).

138. Smith, P.K., Krohn, R.I., Hemanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M. and Klenk, D.C., *Analytical Biochemistry.* 150 76-85 (1985).

139. Gibson, J.J. Personal Communication.

140. Cullwel, A. and Sparrow, J.T.. ABI User Bulletin #18 April (1987).

141. Aimoto, S. and Shimonishi, Y., *Bulletin of the Chemical Society of Japan.* 49 317-320 (1976).

142. Hopp, T.P. and Woods, K.R., *Proceedings of the National Acadamy of Science. U.S.A.* 78 3824-3828 (1981).

143. Kyte, J. and Doolittle, R.F., *Journal of Molecular Biology.* 157 105-132 (1982).

144. Garnier, J., Osguthorpe, D.J. and Robinson, B., *Journal of Molecular Biology* 120 97-120 (1978).

145. Knighton, D., PhD Thesis. Massey University, NZ (1983)

146. Chou, P.Y. and Fasman, E.D., *Biochemistry* 13 211-222 (1974)

147. Margalitt, H., Spouge, J.L., Cornette, J.L., Cease, K., Delisi, C. and Berzofsky, J.A., *The Journal of Immunology.* 138 2213-2229 (1987).

148. Mosmann, T.R., Cherwinski, H., Bond, M.W. Giedlin, M.A. and Coffaman, R.L.. *The Journal of Immunology.* 136 2348-2357 (1986).

149. Fiorentino, D.F. and Mossmann, T.R. *Journal of Experimental Medicine.* 170 2081-2095 (1989)

150. Milton, R.C.de L., Milton, S.C.F. and Adams, P.A., *Journal of the American Chemical Society.* 112 6039-6046 (1990)

151. Kent, S.B.H., Personal communication and Proceedings New Zealand Institute of Chemistry Conference, Wellington,NZ.(1990).

152. Mar glin, A. and Merrifield, R.B., *Journal of the American Chemical Society.* 88 5051-5052 (1966).

153. Gutte, B. and Merrifield, R.B., *Journal of Biological Chemistry.* 246 1922-1941 (1971).

154 Li, C.H. and Yamashiro, D., *Journal of the American Chemical Society.* 92 7608-7609 (1970).

155. Noda, K., Terada, S., Mitsuyasu, M., Waki, M., Kato, T.and Izumiya, N., *Naturwissenschaften* 58 147-148 (1971)

156. Tan, N.H. and Kaiser, E.T., *Journal of Organic Chemistry* 41 2787-2793 (1976).

157. Aoyagi, H., Yonezawa,H., Takahashi, N., Kato, T., Izumiya, H.and Yang, C.C., *Biochemica BiophysicaActa* 263 823-826 (1972).

158. Barstow, L.E., Cornelius, D.A., Hurby, V.J., Shimoda, T.,Rupley, T.A., Sharp, J.J., Robinson, A.B. and Kamen, K.D., in *Chemistry and Biology of Peptides.* Meienhofer, J., Ed, Ann Arbor Science Publishers, Michigan (1972) p 231-233

159. Sharp, J.J., Robinson, A.B. and Kamen, A.D., *Journal of the American Chemical Society.* 95 6097-6108 (1973).

160. Hancock, W.S., Prescott, D.J., Marshall, G.R. and Vagelos, P.R., *Journal of Biological Chemistry* 247 6224-6233 (1972).

161. Sano, S. in *Oxidation Reduction Enzymes.* Akeson, A. and Ehrenberg, A., Ed, Pergamon Oxford (1972) p 35-45.

162. Harding, D.R.K., Battersby, J.E., Husbands, D.R. and Hancock, W.S., *Journal of the American Chemical Society.* 98 2664-2665 (1976).

163. Yamashiro, D and Li, C.H., *Journal of the American Chemical Society.* 100 5174-5179 (1978).

164. Fairwell, T., Hospattanker, A.V., Ronan, R., Brewer, H.B., Chang, J.K., Shimizu, M. and Zitzner, L., *Biochemistry.* 22 2691-2697 (1983).

165. Heath, W. and Merrifield, R.B., *Proceedings of the National Acadamy of Sciences U.S.A..* 83 6367-6371 (1986).

166. Tam, J., Sheikh, M.A., Solomon, D.S. and Ossowski, L., *Proceedings of the National Acadamy of Sciences U.S.A..* 83 8082-8086 (1986).

167. Fairwell, T., Hospattanker, A.V., Brewer, H.B., Khan, S.A., *Proceedings of the National Acadamy of Sciences U.S.A..* 84 4796-4800 (1987).

168. Nokihara, K. and Semba, T., *Journal of the American Chemical Society.* 110 7847-7854 (1988).

169. Clarke-Lewis, I., Lopez, A.F., To, L.B., Vadas, M.A., Schrader. J.W., Hood, L.E., Kent, S.B.H., *Journal of Immunology.* 141 881-889 (1988).

170. Schneider, J. and Kent, S.B.H., *Cell* 54 363-368 (1988).

171. Nutt, R.F., Brady, S.F., Darke, P.L., Ciccarone, T.M., Cotton, C.D., Nutt, E.M., Rodkey, J.A., Bennet, C.D., Waxman, L.H., Sigal, I.S., Anderson, P.S. and Verber, D.F., *Proceedings of the National Acadamy of Sciences U.S.A..* 85 7129-7133 (1988).

172. Ramage, R., Green, J. and Ogunjobi, O.M., *Tetrahedron Letters* 30 2149-2152 (1989).

173. Briand, J.P., Van Dorsselaer, A., Raboy, B. and Muller, S., *Peptide Research* 2 381-388 (1989).

174. a) Lobl, T.J., Diebel, M.R. and Yem, A.W., *Analytical Biochemistry,* 170 502-511 (1988).
b) Diebel, M.R., Lobl, T.J., Yem, A.W., Tracey, D.E. and Paselay, J.W., in *Peptides Chemistry and Biology Proceedings of the 10th American Peptide Symposium,* Marshall, G.R., Ed, ESCOM Science Publishers, Lieden p 245-248.

175. Niall, H.D., *Nature (New Biology)* 230 90-91 (1971).

176. Yajima, H., Kiso,Y., Okada, Y. and Watanabe,H., *Journal of the Chemical Society. Chemical Communications.* 106-107 (1974)

177. Ludwick, A.G., Jelinski, L.W., Live, D., Kintanar, A. and Dumais. J.J., *Journal of the American Chemical Society.* 108 6493-6496 (1986).

178. Kent, S.B.H., in *Peptides Structure and Function* . Ed Hurby, V.J. and Rich, D.H., Pierce Chemical Company, Rockford Illinois (1984) p 99-102.

179. Hahn, K.W., Klis, W.A. and Stewart J.M., *Science* 248 1544-1547 (1990)

180. Laemmli, U.K., *Nature* 277 680- (1970).

181. Ornstein, L., *Annual NY Acadamy of Science.* 121 321 (1964).

182. Davis, B.J., *Annual NY Acadamy of Science.* 121 404 (1964)

183. Poll, D.J. and Harding, D.R.K. *Journal of Chromatography* 469 231-239 (1989).

184. Yajima, H., Fuji, n., Funakoshi, S., Watanabe, T., Murayama, E. and Akira, O., *Tetrahedron* 44 805-819 (1988).

185. Tomich, J.M., Carson, L.W., Kanes, K.J., Vogelaar, N.J., Emerling, M.R. and Richards, J.H., *Analytical Biochemistry* 174 197-203 (1988).

186. Glenny, J.,*Analytical Biochemistry* 156 315-319 (1986).

187. Landon, M., *Methods in Enzymology* 47 145-149 (1977).

188. Marcus, F., *International Journal of Peptide and Protein Research.* 25 542-546 (1985).

189. Applied Biosystems Protein Sequencer User Bullitin Number 13 August 16 (1985).

190. Pestidge, R.L..Personal communication

191. Masher,E., Lundah, P., *Journal of Chromatography* 476 147-158 (1989).

192. Woldawer, A., Miller, M., Jaskolski, M., Sathyanarayana,B.K., Baldwin, E., Weber, I,T., Selk, L.M., Clawson, L., Schneider,J. and Kent, S.B.H., *Science* 245 616-621 (1989).

193. Tregear, G.W., Du,Y.C., Wang, K., Southwell,C., Jones, P., John, M., Gorman, J., Kemp, B. and Niall, H.D., in *Biology of Relaxin and its Role in the Human.*, Bigazzi, M., Greenwood, F.L. and Gasparri, F., Ed, Elsevier, NY, NY, (1983), p 42-55.

194. Lin, S. and Karger, B.L., *Journal of Chromatography,* 499 89-102 (1990)

195. Tam, J.P., *Proceedings of the National Acadamy of Sciences U.S.A.* 85 5409-5413 (1988).

196. Mutter, M., Hersperger, R., Gubernator, K. and Muller, K., *Proteins: Structure Function and Genetics* 5 13-21 (1989)

197. Yang, J.T., Chen, G.C. and Jirgrnsons, B., in *Handbook of Biochemistry and Molecular Biology,* Fasman, G.D.,Ed, 3rd ed Vol 3 p 3-13.(1986)

198. Greenfield, N. and Fasman, G.D., *Biochemstry* 8 4108-4116 (1969).

199. Shoemaker, K.R., Kim, P.S., Brems, D.N., Marquesse, S.J., York, E.J., Chaiken, I.M., Stewart, J.M. and Baldwin, R.L., *Proceedings of the National Acadamy of Science. U.S.A.* 82 2349-2353 (1985)

200. Lark, L.R., Berzofsky, J.A. and Gierasch, L.M., *Peptide Research* 2 314-321 (1989).

201 Provincher, S.W., in CONTIN Users Manual Heildburg Germany (1980)

202. Yang,J.T., Wu, C.S.C. and Martinez, H.M., *Methods in Enzymology* 130 208-269 (1986).

203. Henessey, J.P. and Johnson, W.C., *Analytical Biochemistry* 125 177-188 (1982).

The Complete Amino Acid Sequence Of *M.Leprae* 18 kDa Protein

1 10 20
Met Leu Met Arg Thr Asp Pro Phe Arg Glu Leu Asp Arg Phe Ala Glu Gln Val Leu Gly

21 30 40
Thr Ser Ala Arg Pro Ala Val Met Pro Met Asp Ala Trp Arg Glu Gly Glu Glu Phe Val

41 50 60
Val Glu Phe Asp Leu Pro Gly Ile Lys Ala Asp Ser Leu Asp Ile Asp Ile Glu Arg Asn

61 70 80
Val Val Thr Val Arg Ala Glu Arg Pro Gly Val Asp Pro Asp Arg Glu Met Leu Ala Ala

81 90 100
Glu Arg Pro Arg Gly Val Phe Asn Arg Gln Leu Val Leu Gly Glu Asn Leu Asp Thr Glu

101 110 120
Arg Ile Leu Ala Ser Tyr Gln Glu Gly Val Leu Lys Leu Ser Ile Pro Val Ala Glu Arg

121 130 140
Ala Lys Pro Arg Lys Ile Ser Val Asp Arg Gly Asn Asn Gly His Gln Thr Ile Asn Lys

141 148
Thr Ala His Glu Ile Ile Asp Ala

Tos Bzl OBzl Tos OBzl OBzl Tos OBzl

Met Leu Met Arg Thr Asp Pro Phe Arg Glu Leu Asp Arg Phe Ala Glu Gln Val Leu Gly

Bzl Bzl Tos OBzl CHO Tos OBzl OBzl OBzl

Thr Ser Ala Arg Pro Ala Val Met Pro Met Asp Ala Trp Arg Glu Gly Glu Glu Phe Val

OBzl OBzl Cl-Z OBzl Bzl OBzl OBzl OBzl Tos

Val Glu Phe Asp Leu Pro Gly Ile Lys Ala Asp Ser Leu Asp Ile Asp Ile Glu Arg Asn

Bzl Tos OBzl Tos OBzl OBzl Tos OBzl

Val Val Thr Val Arg Ala Glu Arg Pro Gly Val Asp Pro Asp Arg Glu Met Leu Ala Ala

OBzl Tos Tos Tos OBzl OBzl Bzl OBzl

Glu Arg Pro Arg Gly Val Phe Asn Arg Gln Leu Val Leu Gly Glu Asn Leu Asp Thr Glu

Tos Bzl Br-Z OBzl Cl-Z Bzl OBzl Tos

Arg Ile Leu Ala Ser Tyr Gln Glu Gly Val Leu Lys Leu Ser Ile Pro Val Ala Glu Arg

Cl-Z Tos Cl-Z Bzl OBzl Tos DNP Bzl Cl-Z

Ala Lys Pro Arg Lys Ile Ser Val Asp Arg Gly Asn Asn Gly His Gln Thr Ile Asn Lys

Bzl DNP OBzl OBzl

Thr Ala His Glu Ile Ile Asp Ala

Protecting Groups used for the Syntheis of the 18kDa Protein

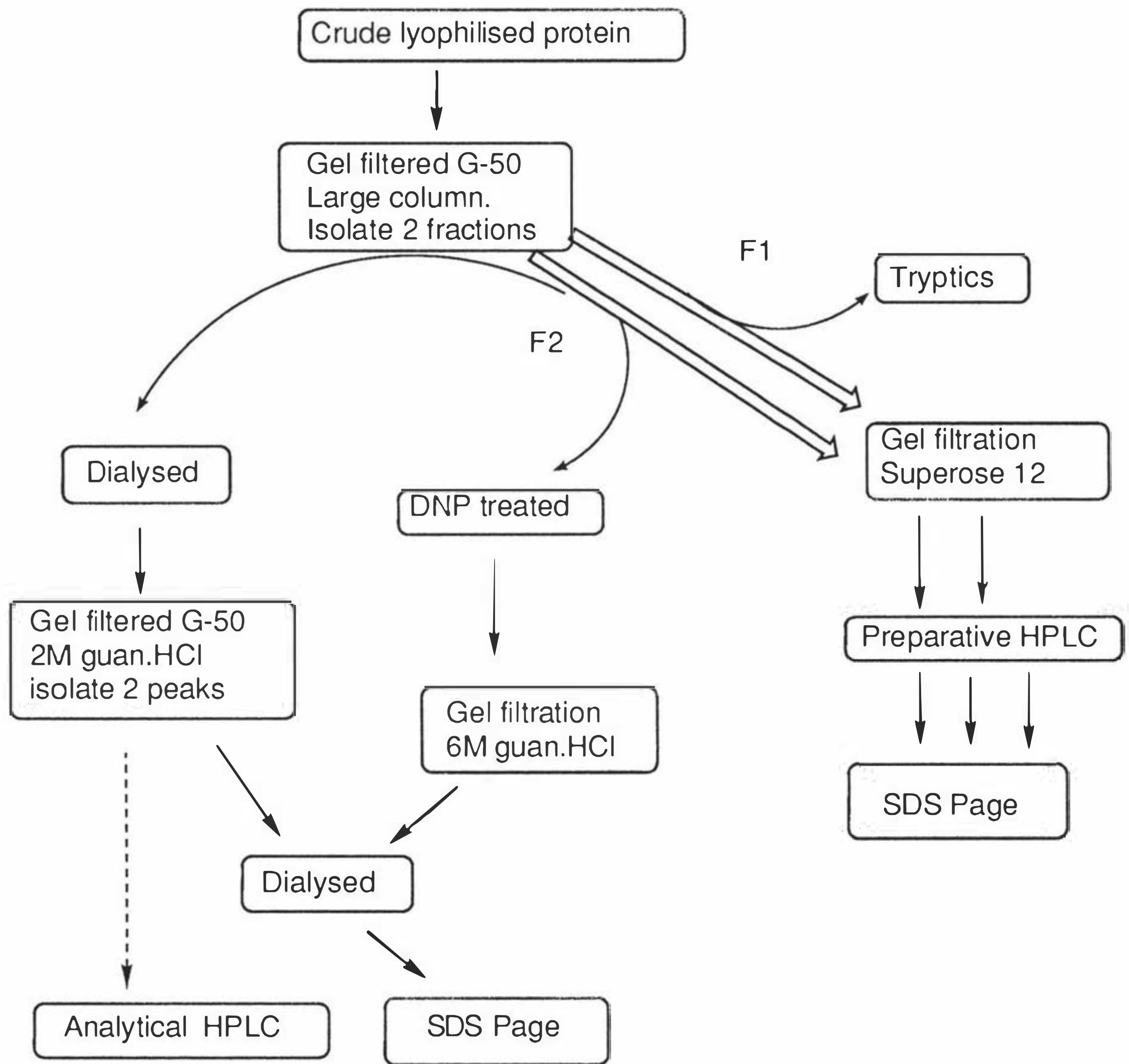


Figure P.1 Flow chart showing the purification scheme for the synthetic protein isolated from cleavage 4. ⟹ the bulk of the material in these two fractions was purified using the method developed from the 90 amino acid fragment.

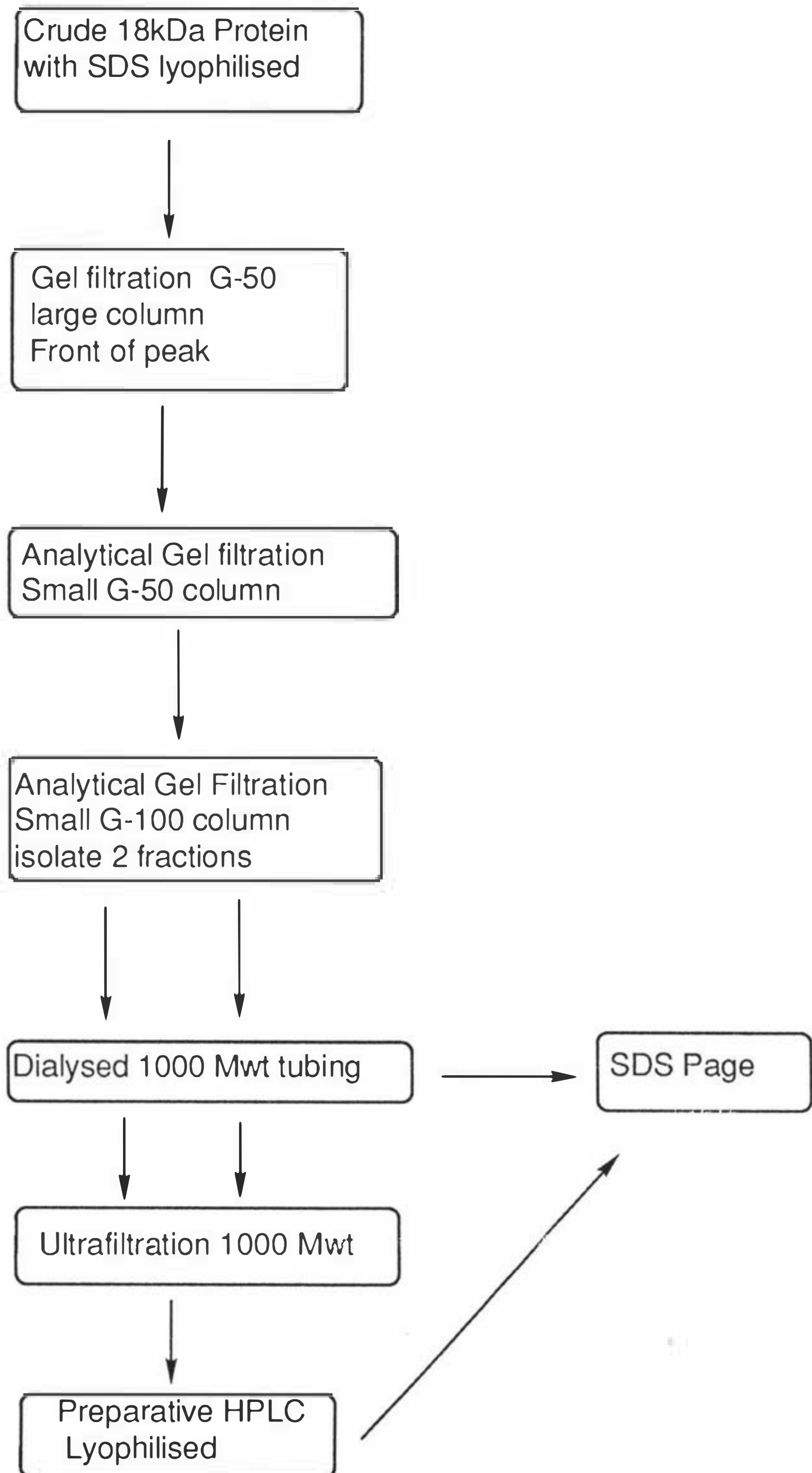


Figure P.2 Flow chart showing the purification scheme for the SDS cleavage, number 5.

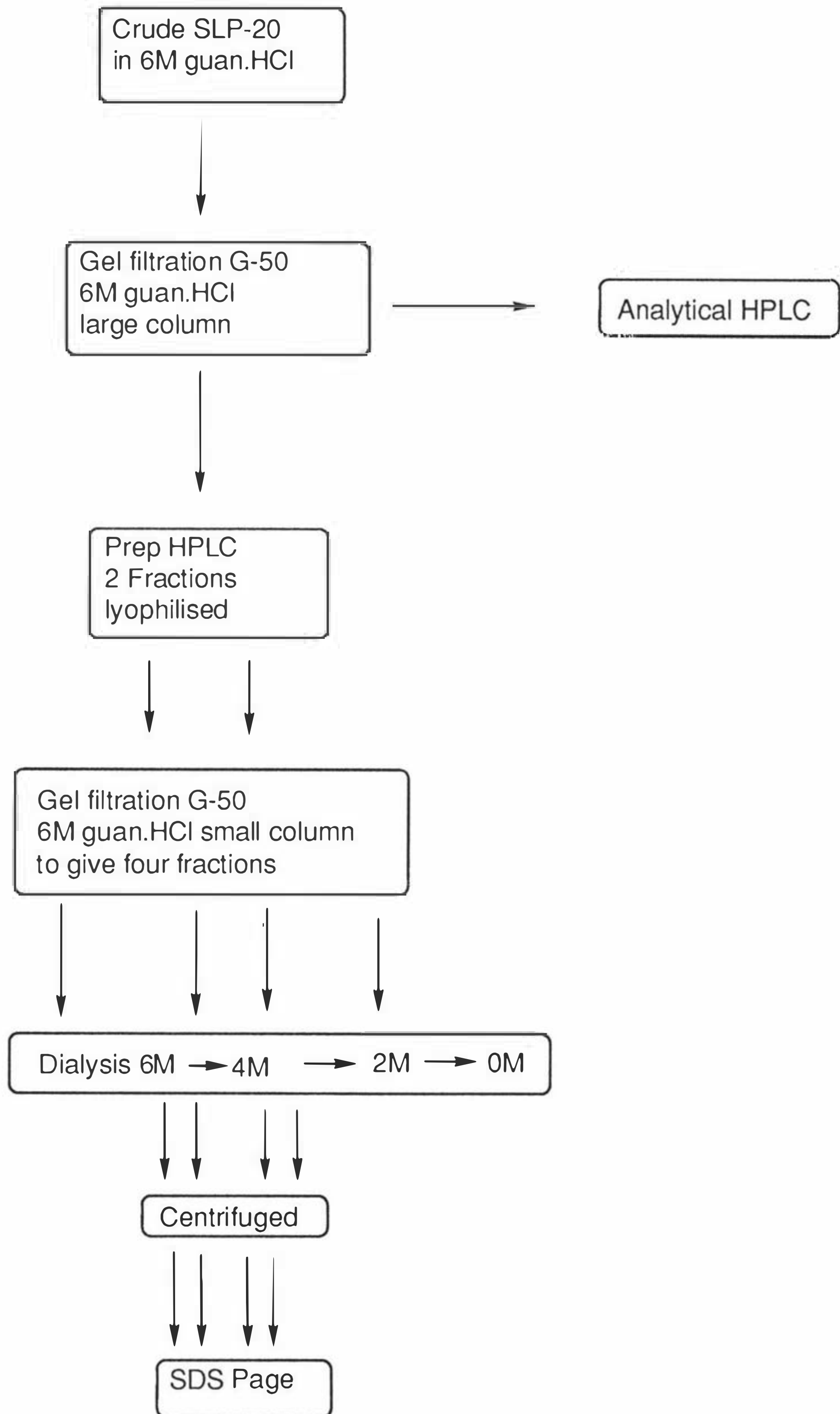


Figure P.3 Flow chart for the purification of the 90 amino acid fragment fom the protein synthesis

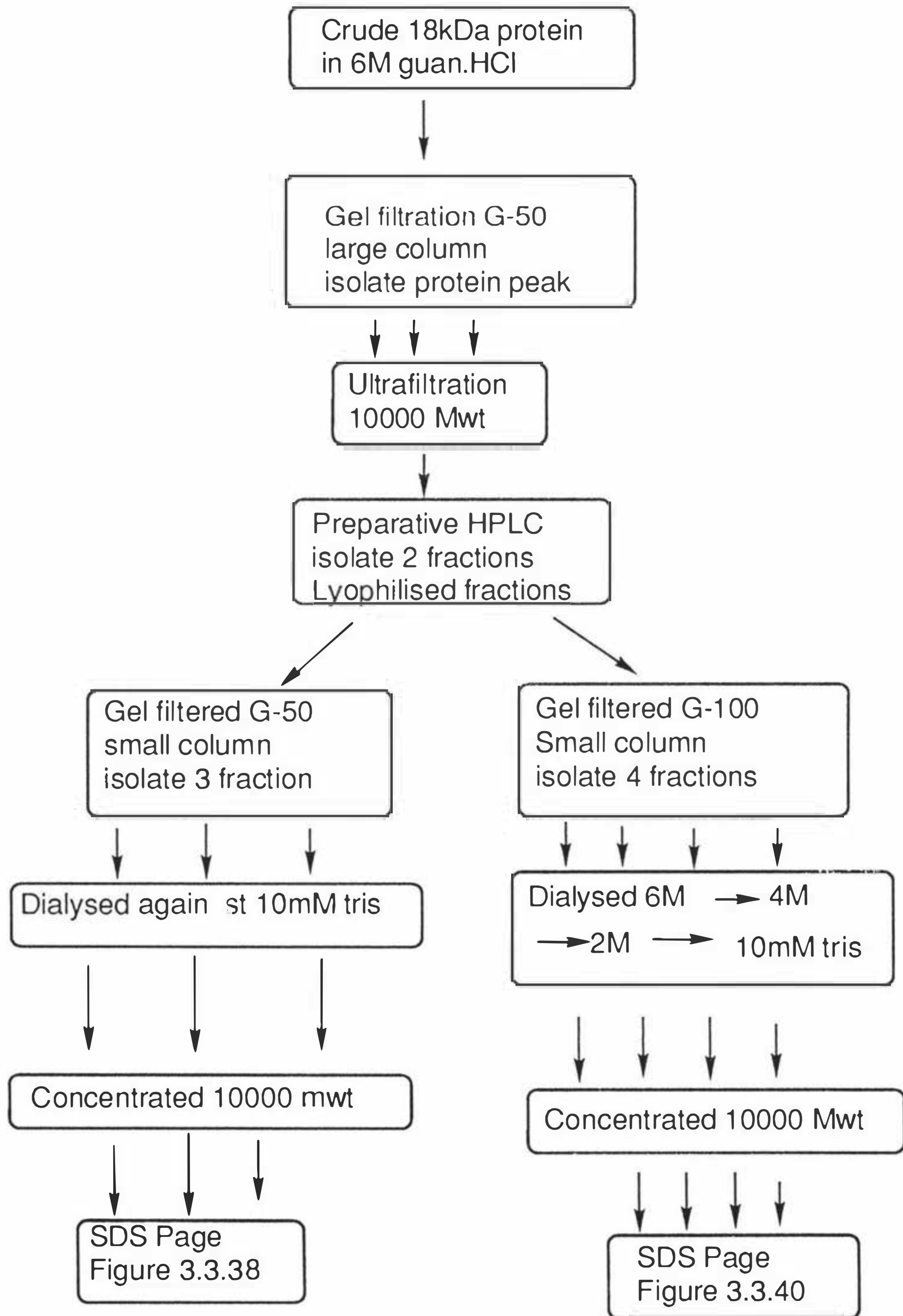


Figure P.4 Flowchart for the purification of the synthetic 18 kDa protein.