THE CHEMICAL SYNTHESIS AND CHARACTERIZATION OF BIOMEMBRANE POLYPEPTIDES

By

BALASUBRAMANIYAM RAMESH B.Sc., (Hons.)

Department of Protein and Molecular Biology,
Royal Free Hospital School of Medicine,
Rowland Hill St.,
London NW3 PF.

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy for the University of London, 1995.
ABSTRACT

Fourier transform infrared (FTIR) and circular dichroism (CD) spectroscopy have been applied to the characterization of chemically synthesized polypeptides corresponding to transmembrane proteins.

A new disaggregation process was introduced and applied to these hydrophobic polypeptides to enable them to be incorporated into membrane mimetic environments without exhibiting the aggregated form. This is illustrated by a comparison of the FTIR spectrum of the natural Pf1 coat protein with that of synthetic Pf1 coat protein when incorporated into lipid vesicles and detergent micelles, before and after disaggregation. The natural Pf1 coat protein and the disaggregated synthetic protein exhibit an α-helical arrangement, whilst the aggregated synthetic protein exhibits a β-sheet type structure. The existence of these aggregated forms of protein can be shown to make spectral analysis difficult and unreliable.

The characterization of synthetic polypeptides corresponding to sequences from (a) ion-selective pore H5, (b) the proposed voltage "sensor" S4 and (c) the region including the transmembrane domains S4-S5-H5-S6 of the voltage-sensitive potassium channel protein was carried out. The FTIR spectra of the disaggregated monomeric, extended, tetrameric version of H5, and the S4-S5-H5-S6 synthetic polypeptides exhibit predominantly an α-helical structure when incorporated into membrane mimetic environments. CD spectroscopy also shows an α-helical conformation of the disaggregated monomeric polypeptide H5 in lipid micelles. When incorporated into lipid planar bilayers the monomeric and the tetrameric H5 polypeptides exhibit ion-channel activity.

The FTIR spectra of the synthetic S4 polypeptide were examined in aqueous buffer, trifluoroethanol solvent and lipid vesicles. The polypeptide in trifluoroethanol and lipid vesicles adopts predominantly an α-helical conformation, whilst in aqueous solution adopts a random coil arrangement. CD spectroscopy also shows an α-helical conformation in lipid micelles and a random coil structure in aqueous buffer. The coil to helix transition observed for the S4 polypeptide when transferred from an aqueous to a lipid environment indicates a high degree of conformational flexibility in response to its environment. This property may have important implications for its proposed role as a voltage-"sensor" during channel activation in response to an electric potential across the membrane.
ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Dennis Chapman for his continued support, kindness and encouragement throughout my studies.

I would also like to express my sincere thanks to the Dean, Professor A.J. Zuckerman for his continued interest throughout my studies.

Special thanks go to Dr. Parvez I. Haris for his invaluable advice and help in the interpretation of FTIR spectra and above all for the moral support throughout the project.

I would also like to express my thanks to Dr. Fred Kurzer who constantly refreshed my knowledge of chemistry.

Also thanks to Dr. K.S. Srai, Dr. B.M. Charalambous and Professor P. Fatt for valuable advice and discussion.

I like to also express my special thanks to Christine Hall for her kindness and help throughout the studies.

Finally I am forever indebted to my family for their patience and encouragement during the course of this study.
# TABLE OF CONTENTS

ABSTRACT 2  
ACKNOWLEDGEMENTS 3  
INDEX TO FIGURES 13  
INDEX TO TABLES 22  
ABBREVIATIONS USED 23  

**CHAPTER ONE: BIOLOGICAL MOLECULES**  25  
1.1 INTRODUCTION 26  
1.2 PROTEINS 27  
1.3 AMINO ACID AND PROTEIN CHEMISTRY 29  
  1.3.1 Amino Acids 29  
  1.3.2 Side-Chain Interactions 43  
    1.3.2.1 Hydrogen Bonds and Salt Bridges 43  
    1.3.2.2 Hydrophobic Interactions 44  
    1.3.2.3 Disulphide Bonds, Thioesters and Thioethers 46  
1.4 THE AMIDE BOND 50  
1.5 PROTEIN STRUCTURE 54  
  1.5.1 Protein Secondary Structure 55  
    1.5.1.1 Helices 56  
    1.5.1.2 $\beta$-Structures 61  
    1.5.1.3 Turns 64  
  1.5.2 Tertiary Structure 68
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5.3 Quaternary Structure</td>
<td>68</td>
</tr>
<tr>
<td>1.6 CHEMICAL SYNTHESIS OF PROTEINS</td>
<td>69</td>
</tr>
<tr>
<td>1.7 SOLID PHASE PRINCIPLE</td>
<td>70</td>
</tr>
<tr>
<td>1.8 SOLID SUPPORTS</td>
<td>72</td>
</tr>
<tr>
<td>1.9 PROTECTING GROUPS</td>
<td>73</td>
</tr>
<tr>
<td>1.9.1 Amino Protecting Groups</td>
<td>74</td>
</tr>
<tr>
<td>1.9.2 Side-Chain Protecting Groups</td>
<td>76</td>
</tr>
<tr>
<td>1.10 ATTACHMENT TO SOLID SUPPORTS</td>
<td>76</td>
</tr>
<tr>
<td>1.11 FORMATION OF THE PEPTIDE BOND</td>
<td>78</td>
</tr>
<tr>
<td>1.11.1 Coupling Reagents</td>
<td>79</td>
</tr>
<tr>
<td>1.11.2 Preformed Symmetrical Anhydrides</td>
<td>82</td>
</tr>
<tr>
<td>1.11.3 Active Esters</td>
<td>82</td>
</tr>
<tr>
<td>1.12 THE SYNTHESIS CYCLE</td>
<td>83</td>
</tr>
<tr>
<td>1.13 DEPROTECTION AND DETACHMENT FROM SOLID SUPPORT</td>
<td>84</td>
</tr>
<tr>
<td>1.14 BIOLOGICAL MEMBRANES</td>
<td>85</td>
</tr>
<tr>
<td>1.15 LIPIDS</td>
<td>90</td>
</tr>
<tr>
<td>1.15.1 Fatty Acids</td>
<td>93</td>
</tr>
<tr>
<td>1.15.2 Phospholipids</td>
<td>96</td>
</tr>
<tr>
<td>1.15.2.1 Phosphoglycerides</td>
<td>96</td>
</tr>
<tr>
<td>1.15.2.2 Choline Phosphoglycerides</td>
<td>97</td>
</tr>
<tr>
<td>1.15.2.3 Ethanolamine Phosphoglycerides</td>
<td>99</td>
</tr>
<tr>
<td>1.15.2.4 Other Phosphoglycerides</td>
<td>99</td>
</tr>
<tr>
<td>1.15.2.5 Phosphosphingosides</td>
<td>100</td>
</tr>
<tr>
<td>1.15.2.6 Properties of Phospholipids</td>
<td>100</td>
</tr>
<tr>
<td>1.16 GLYCOSPHINGOSIDES AND STEROIDS</td>
<td>101</td>
</tr>
</tbody>
</table>
CHAPTER THREE: FTIR SPECTROSCOPIC STUDIES OF CHEMICALLY SYNTHESIZED POLYPEPTIDES CORRESPONDING TO BACTERIOPHAGE PF1 COAT PROTEIN AND ITS EFFICIENT DISAGGREGATION AND RECONSTITUTION INTO MEMBRANE MIMETIC ENVIRONMENTS

3.1 INTRODUCTION

3.2 MATERIALS AND METHODS

3.2.1 Preparation of MBHA resin incorporating the acid labile HMPA linker to produce the resin

3.2.2 Esterification of the first Fmoc-amino acid using pre-formed symmetrical anhydrides


3.2.5 Detachment and deprotection of the synthetic Pf1 coat protein and its 18-residue N-terminal polypeptide

3.2.6 Purification of the synthetic Pf1 coat protein and its 18-residue N-terminal polypeptide

3.2.7 Amino acid compositional and molecular mass analysis

3.2.8 Reconstitution of the synthetic Pf1 coat protein and its N-terminal polypeptide in lipid vesicles and detergent micelles

3.2.9 FTIR spectroscopy

3.2.10 ATR-FTIR spectroscopy
3.3 RESULTS

3.3.1 Chemical and physical analysis of the synthetic Pf1 coat protein and its 18-residue N-terminal polypeptide 176

3.3.2 Studies of the synthetic Pf1 coat protein in DMPC vesicles before disaggregation in $^2$H$_2$O buffer 183

3.3.3 Studies of the synthetic Pf1 coat protein in DMPC vesicles after disaggregation in $^2$H$_2$O buffer 189

3.3.4 Studies of amide hydrogen exchange in synthetic Pf1 coat protein in DMPC vesicles after disaggregation in H$_2$O buffer 192

3.3.5 Studies of the intramembrane orientation of the synthetic Pf1 coat protein in DMPC films 196

3.3.6 Studies of the synthetic Pf1 coat protein in SDS micelles before disaggregation in $^2$H$_2$O buffer 200

3.3.7 Studies of the synthetic Pf1 coat protein in SDS micelles after disaggregation in $^2$H$_2$O buffer 203

3.3.8 Studies of the synthetic 18-residue polypeptide corresponding to the N-terminal region of the Pf1 coat protein before and after disaggregation in $^2$H$_2$O buffer 203

3.3.9 Studies of the synthetic 18-residue polypeptide corresponding to the N-terminal region of the Pf1 coat protein in DMPC vesicles before and after disaggregation in $^2$H$_2$O buffer 209

3.4 DISCUSSION 212

3.5 SUMMARY 217
CHAPTER FOUR: STRUCTURE OF SYNTHETIC POLYPEPTIDES 
CORRESPONDING TO THE PORE-FORMING AND TRANSMEMBRANE 
DOMAINS OF A VOLTAGE-GATED POTASSIUM CHANNEL 

PROTEIN IN MODEL LIPID MEMBRANES AND MICELLES 218 

4.1 INTRODUCTION 219 

4.2 MATERIALS AND METHODS 228 


4.2.4 Assembly of the 119-residue polypeptide incorporating the transmembrane domains S4-S5-H5-S6 230 

4.2.5 Detachment and deprotection of the monomeric, extended and tetrameric H5 and the 119-residue polypeptide incorporating the S4-S5-H5-S6 transmembrane domains 233 

4.2.6 Purification of the synthetic monomeric, extended, tetrameric H5 and the 119-residue polypeptide incorporating the S4-S5-H5-S6 transmembrane domains 234 

4.2.7 Amino acid compositional and molecular mass analysis 235 

4.2.8 Reconstitution of the synthetic monomeric, extended tetrameric H5 and the 119-residue polypeptide incorporating the S4-S5-H5-S6 transmembrane domains in lipid vesicles and micelles 236 

4.2.9 FTIR spectroscopy 236
4.2.10 CD spectroscopy

4.2.11 Planar bilayer studies

4.2.12 Molecular modelling

4.3 RESULTS

4.3.1 Chemical and physical analysis of the synthetic monomeric, extended, tetrameric H5 and the 119-residue polypeptide incorporating the transmembrane S4-S5-H5-S6 domains

4.3.2 Structure of the synthetic monomeric H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer

4.3.3 Structure of the synthetic monomeric H5 polypeptide in LPC micelles in $^2$H$_2$O buffer

4.3.4 Structure of the synthetic monomeric H5 polypeptide in DMPG vesicles in $^2$H$_2$O buffer

4.3.5 Structure of the synthetic extended H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer

4.3.6 Structure of the synthetic tetrameric H5 polypeptide in DMPC micelles in $^2$H$_2$O buffer

4.3.7 Structure of the synthetic 119-residue polypeptide incorporating the transmembrane S4-S5-H5-S6 domains in DMPC vesicles in $^2$H$_2$O buffer

4.3.8 Planar bilayer measurements

4.3.9 Molecular modelling and analysis of amphipathicity

4.4 DISCUSSION

4.5 SUMMARY
CHAPTER FIVE: CONFORMATIONAL ANALYSIS OF A SYNTHETIC POLYPEPTIDE CORRESPONDING TO THE VOLTAGE-SENSING S4 DOMAIN OF A VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN

5.1 INTRODUCTION
5.2 MATERIALS AND METHODS
  5.2.1 Assembly of the S4 peptidyl-resin sequence: Fmoc-I-L-R-V-I-R-L-V-R-V-F-R-I-F-K-L-S-R-H-S-Rink-MBHA-resin
  5.2.2 Detachment and deprotection of the synthetic S4 polypeptide
  5.2.3 Purification of the synthetic S4 polypeptide
  5.2.4 Amino acid compositional and molecular mass analysis
  5.2.5 Reconstitution of the synthetic S4 polypeptide in 98% TFE, DMPC vesicles and LPC micelles in $^2\text{H}_2\text{O}$ buffer
  5.2.6 FTIR and CD spectroscopy
5.3 RESULTS
  5.3.1 Chemical and physical analysis of the synthetic S4 polypeptide
  5.3.2 Conformation of the synthetic S4 polypeptide in 98% TFE in $^2\text{H}_2\text{O}$ buffer
  5.3.3 Conformation of the synthetic S4 polypeptide in LPC micelles in $^2\text{H}_2\text{O}$ buffer
  5.3.4 Conformation of the synthetic S4 polypeptide in DMPC vesicles in $^2\text{H}_2\text{O}$ buffer
  5.3.5 Conformation of the synthetic S4 polypeptide in $^2\text{H}_2\text{O}$ buffer
5.4 DISCUSSION
INDEX TO FIGURES

Figure 1.1: (a) Structure of the unionized form of an amino acid 31
(b) Structure of the zwitterion form of an amino acid

Figure 1.2: Stereoisomers of the α-amino acids 32

Figure 1.3: Some amino acids modified by post-translation 41

Figure 1.4: Disulphide bond formation on oxidation with two molecules of cysteine, and the reverse reaction under appropriate conditions 47

Figure 1.5: Peptides containing disulphide bonds 49

Figure 1.6: Torsional angles of the peptide backbone 51

Figure 1.7: Amide bonds 52

Figure 1.8: (a) The α-helix 59
(b) The hydrogen bonding arrangement in the helix

Figure 1.9: (a) Parallel β-sheet structure 62
(b) Antiparallel β-sheet structure

Figure 1.10: Illustration of a type I β-turn in a protein 65

Figure 1.11: Reaction scheme for solid-phase peptide synthesis 71

Figure 1.12: Protecting group strategies in SPPS 75
(a) Merrifield protection scheme
(b) Orthogonal protection scheme

Figure 1.13: Widely used "permanent" side-chain protecting groups compatible with Fmoc chemistry 77

Figure 1.14: Coupling reaction in SPPS 80

Figure 1.15: In situ coupling reagents and additives suitable for SPPS 81

Figure 1.16: Membrane structure 88
Figure 1.17: The amphipathic nature of phospholipids and cholesterol

Figure 1.18: Structure of a triglyceride

Figure 1.19: Structural arrangement of the carbon chains in fatty acids:
(a) saturated trans form
(b) unsaturated trans form
(c) unsaturated cis form

Figure 1.20: Structure of some phosphoglycerides

Figure 1.21: Structures adopted by phospholipids in aqueous solution

Figure 1.22: A Galactocerebroside

Figure 2.1: The electromagnetic spectrum and interactions with matter

Figure 2.2: Conventional X-ray diffraction

Figure 2.3: Schematic diagram for the most important processes involved in light scattering by molecules:
(a) Rayleigh scattering
(b) Production of the Stokes-Raman line
(c) Production of the anti-Stokes-Raman line

Figure 2.4: Vibrational modes of CH₂ groups of lipid acyl chains:
(a) Stretching vibrations
(b) Bending vibrations

Figure 2.5: FTIR transmittance spectrum of H₂O and of ²H₂O recorded in a calcium fluoride cell fitted with a 6μm tin spacer

Figure 2.6: Schematic representation of the Michelson interferometer optical system used in an FTIR spectrometer

Figure 2.7: Output of the Michelson interferometer as a function of mirror displacement:
(a) a monochromatic source
(b) a polychromatic source
Figure 2.8: Illustration of the process of second derivative of a Lorentzian band:
(a) Synthetic Lorentzian with $v = 5 \text{cm}^{-1}$
(b) Fourier transform of (a)
(c) Parabolic weighting function
(d) Product of (b) and (c)
(e) Fourier transform of (d) yielding the second derivative of (a)

Figure 2.9: Illustration of the mechanism of deconvolution:
(a) multiplication of the decaying cosine train by an increasing exponential function.
(b) multiplication of the finite cosine train by an apodization function

Figure 2.10: Internal reflection spectroscopy
(a) Conditions for single internal reflection
(b) Multiple internal reflection

Figure 2.11: Amide I, II and III vibrational modes of a peptide

Figure 3.1: Amino acid sequence of Pf1 coat protein

Figure 3.2: Model for the structural changes that occur in Pf1 coat protein viral assembly

Figure 3.3: The expanding cylinder of the virus shown relative to the assembling coat protein

Figure 3.4: Model of Pf1 coat protein in lipid bilayers showing the amino terminal amphipathic helix parallel to the plane of the bilayer

Figure 3.5: The amphipathic synthetic peptide corresponding to the N-terminus sequence of Pf1 coat protein

Figure 3.6: Preparative scale reverse-phase HPLC chromatogram of crude synthetic Pf1 coat protein

Figure 3.7: Preparative scale reverse-phase HPLC chromatogram of crude synthetic peptide corresponding to the N-terminus sequence of Pf1 coat protein
Figure 3.8: Analytical reverse-phase HPLC of synthetic Pf1 coat protein after purification

Figure 3.9: Analytical reverse-phase HPLC of synthetic peptide corresponding to the N-terminus sequence of Pf1 coat protein after purification

Figure 3.10: Laser-desorption mass spectra of synthetic Pf1 coat protein

Figure 3.11: Laser-desorption mass spectra of synthetic 18-residue N-terminal region of Pf1 coat protein

Figure 3.12: The FTIR spectra of PF1 coat protein from 1800 cm\(^{-1}\) to 1500 cm\(^{-1}\) of synthetic before disaggregation and natural forms in DMPC vesicles in \(^2\)H\(_2\)O buffer at 20°C

Figure 3.13: Second-derivative spectra of Pf1 coat protein of synthetic before disaggregation and natural forms in DMPC vesicles in \(^2\)H\(_2\)O buffer at 20°C

Figure 3.14: The FTIR absorption spectra of PF1 coat protein of synthetic after disaggregation and natural forms in DMPC vesicles in \(^2\)H\(_2\)O buffer at 20°C

Figure 3.15: Second-derivative spectra of PF1 coat protein of synthetic after disaggregation and natural forms in DMPC vesicles in \(^2\)H\(_2\)O buffer at 20°C

Figure 3.16: The FTIR absorption spectra of synthetic Pf1 coat protein in DMPC vesicles in H\(_2\)O and \(^2\)H\(_2\)O buffer at 30°C

Figure 3.17: Second-derivative spectra of synthetic Pf1 coat protein in DMPC vesicles in H\(_2\)O and \(^2\)H\(_2\)O buffer at 30°C

Figure 3.18: Internal reflectance infrared spectra of Pf1 coat protein of synthetic after disaggregation and natural forms in DMPC films with beams polarized perpendicular to the plane of the film
Figure 3.19: Internal reflectance infrared spectra of Pf1 coat protein of synthetic after disaggregation and natural forms in DMPC films with beams polarized parallel to the plane of the film

Figure 3.20: Internal reflectance infrared spectra of DMPC films with reconstituted synthetic Pf1 coat protein after disaggregation with beams perpendicular and parallel to the plane of the film

Figure 3.21: The FTIR absorption spectra of PF1 coat protein of the synthetic before disaggregation and natural forms in SDS micelles in $^2$H$_2$O buffer at 20°C

Figure 3.22: Second-derivative spectra of PF1 coat protein of synthetic before disaggregation and natural forms in SDS micelles in $^2$H$_2$O buffer at 20°C

Figure 3.23: The FTIR absorption spectra of Pf1 coat protein of the synthetic after disaggregation and natural forms in SDS micelles in $^2$H$_2$O buffer at 20°C

Figure 3.24: Second-derivative spectra of PF1 coat protein of synthetic after disaggregation and natural forms in SDS micelles in $^2$H$_2$O buffer at 20°C

Figure 3.25: The FTIR absorption spectra of the synthetic 18-residue N-terminal peptide of PF1 coat protein before and after disaggregation in $^2$H$_2$O buffer at 20°C

Figure 3.26: Second-derivative spectra of the synthetic 18-residue N-terminal peptide of Pf1 coat protein before and after disaggregation in $^2$H$_2$O buffer at 20°C

Figure 3.27: The FTIR absorption spectra of the synthetic 18-residue N-terminal peptide of Pf1 coat protein before and after disaggregation in DMPC vesicles in $^2$H$_2$O buffer at 20°C
Figure 3.28: Second-derivative spectra of the synthetic 18-residue N-terminal peptide of Pf1 coat protein before and after disaggregation in DMPC vesicles in $^2$H$_2$O buffer at 20°C

Figure 4.1: Proposed transmembrane topology and subunit number of Na$^+$, Ca$^{2+}$ and K$^+$ channel protein

Figure 4.2: Transmembrane folding model for a voltage-activated K$^+$ channel

Figure 4.3: The amino acid sequence of the pore-forming region of the ShA K$^+$ channel above a postulated membrane folding structure

Figure 4.4: Amino acid sequences of the synthetic:
(a) H5 polypeptide
(b) Extended H5 polypeptide
(c) Tetrameric H5 polypeptide
(d) S4-S5-H5-S6 polypeptide

Figure 4.5: Schematic representation of solid phase Fragment condensation synthesis of the 119-residue S4-S5-H5-S6 polypeptide

Figure 4.6: Analytical reverse-phase HPLC of the synthetic H5 polypeptide after purification

Figure 4.7: Analytical reverse-phase HPLC of the synthetic extended H5 polypeptide after purification

Figure 4.8: Analytical reverse-phase HPLC of the synthetic tetrameric H5 polypeptide after purification

Figure 4.9: Analytical reverse-phase HPLC of the 118-residue S4-S5-H5-S6 polypeptide after purification

Figure 4.10: Laser-desorption mass spectra of H5 polypeptide

Figure 4.11: Laser-desorption mass spectra of Extended H5 polypeptide

Figure 4.12: Laser desorption mass spectra of Tetrameric H5 polypeptide
Figure 4.13: The FTIR absorption spectrum of synthetic monomeric H5 polypeptide in DMPC vesicles in D2O buffer at 30°C

Figure 4.14: The second-derivative spectrum of synthetic monomeric H5 polypeptide in DMPC vesicles in D2O buffer at 30°C

Figure 4.15: The FTIR absorption spectrum of synthetic monomeric H5 polypeptide in LPC micelles in D2O buffer at 30°C

Figure 4.16: The second-derivative spectrum of synthetic monomeric polypeptide in LPC micelles in D2O buffer at 30°C

Figure 4.17: CD spectrum of synthetic H5 polypeptide in LPC micelles in D2O buffer

Figure 4.18: The FTIR absorption spectrum of synthetic monomeric H5 polypeptide in DMPG vesicles in D2O buffer at 30°C

Figure 4.19: The second-derivative spectrum of synthetic monomeric H5 polypeptide in DMPG vesicles in D2O buffer at 30°C

Figure 4.20: The FTIR absorption spectrum of synthetic extended H5 polypeptide in DMPC vesicles in D2O at 30°C

Figure 4.21: The second-derivative spectrum of synthetic H5 extended polypeptide in DMPC vesicles in D2O buffer at 30°C

Figure 4.22: The FTIR absorption spectrum of synthetic tetrameric H5 polypeptide in DMPC vesicles in D2O buffer at 30°C

Figure 4.23: The second-derivative spectrum of synthetic tetrameric H5 polypeptide in DMPC vesicles in D2O buffer at 30°C

Figure 4.24: The FTIR absorption spectrum of synthetic 119-residue S4-S5-H5-S6 polypeptide in DMPC vesicles in D2O at 30°C
Figure 4.25: The second-derivative spectrum of synthetic 119-residue S4-S5-H5-S6 polypeptide in DMPC vesicles in $^2$H$_2$O at 30°C

Figure 4.26: Channel formation by monomeric and tetrameric H5 polypeptide

Figure 4.27: Modelling the H5 helix by simulated annealing

Figure 5.1: Putative transmembrane folding scheme of voltage dependent K$^+$ channels. The S4 sequence of Shaker K$^+$ channel is shown below the folding scheme

Figure 5.2: Sequence of the synthetic S4 polypeptide

Figure 5.3: Preparative reverse-phase HPLC of the synthetic S4 polypeptide

Figure 5.4: Analytical reverse-phase HPLC of the synthetic S4 polypeptide

Figure 5.5: Laser-desorption mass spectra of the synthetic S4 polypeptide

Figure 5.6: FTIR absorption spectrum of synthetic S4 polypeptide in 98% trifluoroethanol in $^2$H$_2$O at 30°C

Figure 5.7: Second-derivative spectrum of synthetic S4 polypeptide in 98% trifluoroethanol in $^2$H$_2$O at 30°C

Figure 5.8: FTIR absorption spectrum of synthetic S4 polypeptide in LPC micelles in $^2$H$_2$O buffer at 30°C

Figure 5.9: Second-derivative spectrum of synthetic S4 polypeptide in LPC micelles in $^2$H$_2$O buffer at 30°C

Figure 5.10: CD spectra of synthetic S4 polypeptide in aqueous solution and LPC micelles in $^2$H$_2$O buffer
Figure 5.11: FTIR absorption spectrum of synthetic S4 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30°C

Figure 5.12: Second-derivative spectrum of synthetic S4 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30°C

Figure 5.13: FTIR absorption spectrum of synthetic S4 polypeptide in $^2$H$_2$O buffer at 30°C

Figure 5.14: Second-derivative spectrum of synthetic S4 polypeptide in $^2$H$_2$O buffer at 30°C

Figure 5.15: FTIR absorption spectrum of synthetic S4 polypeptide in LPC micelles (dashed trace); in DMPC vesicles (red trace); in aqueous solution (bold trace) in $^2$H$_2$O buffer at 30°C

Figure 5.16: Second-derivative spectrum of synthetic S4 polypeptide in LPC micelles (dashed trace); in DMPC vesicles (red trace); in aqueous solution (bold trace) in $^2$H$_2$O buffer at 30°C
# INDEX TO TABLES

<p>| Table 1.1: | Amino Acid Structures And Properties | 33 |
| Table 1.2: | Torsional Angles For Helices | 58 |
| Table 1.3: | Torsional Angles For β-Structure | 63 |
| Table 1.4: | Torsional Angles For Turns | 67 |
| Table 1.5: | Chemical Composition Of Cell Membranes | 86 |
| Table 1.6: | Influence Of Headgroup, Chain Length And Unsaturation Upon Tₘ And Enthalpy Changes | 107 |
| Table 2.1: | Characteristic IR Absorption Frequencies Of Proteins | 152 |
| Table 2.2: | Characteristic IR Absorption Frequencies Of Lipids | 156 |
| Table 3.1: | Amino Acid Analysis Of Synthetic Pf1 Coat Protein | 181 |
| Table 3.2: | Amino Acid Analysis Of The 18-residue Synthetic Peptide corresponding To The N-terminal Sequence Of Pf1 Coat Protein | 182 |
| Table 4.1: | Amino Acid Analysis Of Synthetic H5 Polypeptide | 244 |
| Table 4.2: | Amino Acid Analysis Of Synthetic Extended H5 Polypeptide | 245 |
| Table 4.3: | Amino Acid Analysis Of Tetrameric H5 Polypeptide | 246 |
| Table 4.4: | Amino Acid Analysis Of Synthetic S4-S5-H5-S6 Polypeptide | 247 |
| Table 5.1: | Amino Acid Analysis Of Synthetic S4 Polypeptide | 292 |</p>
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac₂O</td>
<td>Acetic anhydride</td>
</tr>
<tr>
<td>ACN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>ATR</td>
<td>Attenuated total reflection</td>
</tr>
<tr>
<td>BES</td>
<td>N,N-bis-(2-Hydroxyethyl)-2-aminoethanesulphonic acid</td>
</tr>
<tr>
<td>BMC</td>
<td>Butylmethylcarbodiimide</td>
</tr>
<tr>
<td>Boc</td>
<td>t-Butoxycarbonyl</td>
</tr>
<tr>
<td>BOP</td>
<td>Benzotriazol-yl-oxy-tris-(dimethylamino)-phosphoniumhexafluorophosphate</td>
</tr>
<tr>
<td>CaF₂</td>
<td>Calcium fluoride</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>DCCI</td>
<td>Dicyclohexylcarbodiimide</td>
</tr>
<tr>
<td>DCM</td>
<td>Dichloromethane</td>
</tr>
<tr>
<td>DCU</td>
<td>Dicyclohexylurea</td>
</tr>
<tr>
<td>DIPCIDI</td>
<td>Diisopropylcarbodiimide</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethylformamide</td>
</tr>
<tr>
<td>DMPC</td>
<td>1-α-Dimyristoyl phosphatidylcholine</td>
</tr>
<tr>
<td>DMPG</td>
<td>1-α-Dimyristoyl Phosphotidyl-DL-glycerol</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-Dimethylaminopyridine</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DPC</td>
<td>Dodecylphosphocholine</td>
</tr>
<tr>
<td>DSC</td>
<td>Differential scanning calorimetry</td>
</tr>
<tr>
<td>diPhyPc</td>
<td>1,2-Diphytanoyl-3-sn-phosphatidylcholine</td>
</tr>
<tr>
<td>EDT</td>
<td>Ethanedithiol</td>
</tr>
<tr>
<td>Fmoc</td>
<td>9-Fluorenylethoxycarbonyl</td>
</tr>
<tr>
<td>FTIR</td>
<td>Fourier transform infrared</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine triphosphate</td>
</tr>
<tr>
<td>HBTU</td>
<td>2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate</td>
</tr>
<tr>
<td>HF</td>
<td>Hydrogen fluoride</td>
</tr>
<tr>
<td>hGh</td>
<td>Human growth hormone</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HOBt</td>
<td>N-Hydroxybenzotriazole</td>
</tr>
<tr>
<td>HMPB</td>
<td>4-(4-Hydroxymethyl-3-methoxyphenoxy)-butyric acid</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>LPC</td>
<td>Lysophosphatidylcholine</td>
</tr>
<tr>
<td>MBHA</td>
<td>Methylbenzyhydrylamine</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>MMP</td>
<td>N-Methylmorpholine</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>ODHBt</td>
<td>3-hydroxy-2-3-dihydro-4-oxo-benzotriazine</td>
</tr>
<tr>
<td>OPfp</td>
<td>Pentafluorophenyl ester</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PSA</td>
<td>Preformed symmetrical anhydride</td>
</tr>
<tr>
<td>PyBOP</td>
<td>Benzotriazole-1-yl-oxy-tris-pyrrolidino-Phosphonium hexafluorophosphate</td>
</tr>
<tr>
<td>PyBroP</td>
<td>Bromo-tris-pyrrolidino-phosphonium hexafluorophosphate</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecylsulphate</td>
</tr>
<tr>
<td>SPPS</td>
<td>Solid phase peptide synthesis</td>
</tr>
<tr>
<td>TBTU</td>
<td>2-(1H-Benzotriazole-1-yl)-1,3,3-tetramethyluronium tetrafluoroborate</td>
</tr>
<tr>
<td>TNTU</td>
<td>2-(5-Norbornene-2,3-dicarboximido)-1,1,3,3-tetramethyluronium tetrafluoroborate</td>
</tr>
<tr>
<td>TSTU</td>
<td>O-(N-Succinimidyl)-1,1,3,3-tetramethyluronium tetrafluoroborate</td>
</tr>
<tr>
<td>$T_m$</td>
<td>Temperature of the main phase transition in phospholipids</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>TFE</td>
<td>Trifluoroethanol</td>
</tr>
<tr>
<td>TFMSA</td>
<td>Trifluoromethanesulfonic acid</td>
</tr>
<tr>
<td>TMSOTf</td>
<td>Trimethylsilyl trifluoroacetate</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra-violet</td>
</tr>
</tbody>
</table>
CHAPTER ONE

BIOLOGICAL MOLECULES
1.1 INTRODUCTION

Among the numerous biological molecules in the living cell, three have attracted the greatest attention: (1) protein (2) RNA and (3) DNA. They are macromolecules, large molecules that are linear polymers assembled from simple subunits, or monomers. It was the proteins that attracted tremendous attention and resources and the reasons were clear. In some specialised tissues, the production and accumulation of large amount of only one kind of protein facilitated the isolation and purification for further study. However during the past decade two major technologies have provided enormous progress in understanding biological systems. One is the exploitation of recombinant DNA technology and the second is advances in solid phase peptide synthesis. The ever increasing number of primary sequence of proteins that are being elucidated by the former is not being matched by the number of three-dimensional crystal structures. With the advent of powerful microcomputers, accompanied by speed and convenience of data acquisition and analysis, there is now an increased trend towards the complementary use of biochemical and biophysical techniques in understanding the relationship between structure and function of proteins and of biological membranes. With this combination, detailed mechanisms can sometimes be proposed for biological action. It is still a difficult task to predict the structure of a protein from its amino acid sequence or the subtle conformational effects caused by the substitution of one amino acid into another. The studies presented in this thesis describes the interaction of synthetic polypeptides corresponding to sequences of biomembrane proteins on reconstitution into model membranes.
Many biophysical techniques have been used to either determine the structure of the entire molecule or individual domains of biological molecules. They include X-ray crystallography, NMR, CD, Raman spectroscopy and infrared spectroscopy. The two major techniques that have provided detailed structural information at atomic level on a number of soluble proteins are X-ray crystallography and NMR spectroscopy. In spite of these remarkable achievements, our understanding of the detailed structure and, therefore of the function of membrane proteins lags far behind our knowledge of soluble proteins. Since then three notable membrane protein complexes, the bacterial photosynthetic reaction centres of *Rhodopseudomonas Viridis* (Disenhofer et al. 1985), *Rhodobacter sphaeroids* (Chang et al. 1985) and the matrix porin of *Rhodobacter capsulatus* (Weiss et al. 1990) have been determined at high resolution by X-ray crystallography. This has paved the way for the application of other biophysical techniques in the study of proteins and their incorporation into biological membranes. An understanding of the resolution and limitations of the biophysical techniques applied to biomolecules can only be gained with a knowledge of the properties of these molecules, therefore chapter one begins with the introduction of the fundamental properties and structure of proteins, its chemical synthesis, and its association with lipids in biomembranes.

1.2 PROTEINS

Proteins are the working molecules of biological processes encoded by genes. The proteins in turn give rise to structure and, by virtue of their selective interaction with other molecules create genes and all the various
machinery of life. Through all the functional diversity of proteins there exists a common theme: for the most part, proteins function by selectively binding to molecules, creating aggregated copies of the same protein to form bulk structures such as fibre, a sheet or a tubule. Other proteins have an affinity for molecules different from themselves. In immune response, for example, specific immunoglobulins bind to foreign substances on bacteria, viruses and cells from other species. Proteins thus can perform the role of immunosurveillance by distinguishing between self and nonself. In transportation, haemoglobin and myoglobin binds to oxygen and then releases it in distant tissues. Whereas in storage, transferrin ferry iron in plasma to the liver where it is stored as ferritin. As regulators of genetic expression, they bind to specific patterns of nucleotide bases in DNA. Receptor proteins embedded in the cell membrane recognise messenger molecules (such as hormones and neurotransmiters), which may themselves be proteins that have a specific affinity for the receptors. In a wide range of chemical reaction catalysis is carried out by specific macromolecules called enzymes. Virtually all the activities of proteins can be defined in terms of such exquisite specificity. In fact, each protein has a unique structure, and as a consequence of that structure, it has a unique shape and function in any given set of environmental conditions. The physical and chemical properties of proteins are determined by the nature of the constituent amino acid side chains and by the polyamide peptide backbone itself. Twenty protein amino acids are coded for by DNA, which are translationally incorporated into proteins. In order to comprehend protein structure in detail, the composition and properties of individual amino acids...
1.3 AMINO ACID AND PROTEIN CHEMISTRY

The nature of the constituent amino acid side chains and the polyamide peptide backbone dictates the physical and the chemical properties of proteins and peptides. There are twenty protein amino acids which are coded for by DNA, which are translationally incorporated into proteins. Some amino acids can undergo modification in the protein post-translationally to yield new amino acids. There are also some polypeptides which are not necessarily assembled via the ribosomal machinery but by enzyme interaction, especially in the lower eukaryotes. These polypeptides are sometimes found to contain unusual amino acids.

1.3.1 AMINO ACIDS

As their name suggests, amino acids are organic acids which contain an amino group. Except for proline, the biologically important amino acids all belong to the group called α-amino acids. That is, they are all organic acids containing an unsubstituted amino group (−NH₂), a hydrogen atom, and a R group bonded to a carbon atom known as α-carbon (Fig. 1.1). The tetrahedral arrangement of the four different substituent groups on the α-carbon atom confers optical activity on amino acids. The mirror images of each other are referred to as D and L isomers (Fig 1.2). Living organisms can differentiate between these two forms, and most can synthesize or metabolise only the L amino acids. The differences among amino acids, then, stem from differences...
in their side chains, namely, in shape, size and polarity. The structures of the 20 primary protein amino acids (those coded for by DNA) are depicted in Table 1.1 along with their three-letter abbreviations, one-letter codes, and a general grouping by physical properties. The amino acids can generally be grouped into hydrophobic and hydrophilic residues. The hydrophobic residues include those with alipathic side-chains, such as alanine, valine, isoleucine, leucine and methionine, and those containing aromatic side chains, such as phenylalanine, tyrosine and tryptophan. The hydrophilic residues can be categorised conveniently into three groups: (1) amino acids with neutral, polar side chains, such as serine, threonine and asparagine and (2) and those with acidic side chains, such as aspartic acid and glutamic acid, and (3) those with basic side chains such as histidine, lysine and arginine.

These classifications however are not entirely exclusive, alanine for instance with its small alipathic side-chain and glycine can sometimes be located in the hydrophilic regions of proteins and peptides. The arginine and lysine on the other hand with their long alkyl chain can confer those residues an overall hydrophobic character with just the terminal charged group being hydrophilic. The two amino acids, cysteine and proline have unusual properties that sets them apart. The thiol moiety of cysteine can be oxidatively linked to another cysteine thiol to form a disulphide (-S-S-) linkage. This oxidized dimer of cysteine is known as cystine which is a prominent feature by which polypeptide chains are covalently linked together to stabilise secondary and tertiary structure or hold two peptide chains together.
Figure 1.1. (a) Structure of unionized and (b) zwitterion forms of an amino acid.
Figure 1.2. Stereoisomers of the $\alpha$-amino acids.

L-Amino acid  
(a)

D-Amino acid  
(b)
Table 1.1. Amino acid Structures and Properties

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Planar representation</th>
<th>Side chain conformation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neutral, Hydrophobic, Aliphatic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycine</td>
<td><img src="image" alt="Glycine Planar Representation" /></td>
<td><img src="image" alt="Glycine Side Chain Conformation" /></td>
</tr>
<tr>
<td></td>
<td>Gly (G)</td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td><img src="image" alt="Alanine Planar Representation" /></td>
<td><img src="image" alt="Alanine Side Chain Conformation" /></td>
</tr>
<tr>
<td></td>
<td>Ala (A)</td>
<td></td>
</tr>
<tr>
<td>Valine</td>
<td><img src="image" alt="Valine Planar Representation" /></td>
<td><img src="image" alt="Valine Side Chain Conformation" /></td>
</tr>
<tr>
<td></td>
<td>Val (V)</td>
<td></td>
</tr>
<tr>
<td>Leucine</td>
<td><img src="image" alt="Leucine Planar Representation" /></td>
<td><img src="image" alt="Leucine Side Chain Conformation" /></td>
</tr>
<tr>
<td></td>
<td>Leu (L)</td>
<td></td>
</tr>
</tbody>
</table>
Table 1.1. Amino acid Structures and Properties  (continued)

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Planar representation</th>
<th>Side chain conformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isoleucine</td>
<td><img src="image" alt="Isoleucine Structure" /></td>
<td><img src="image" alt="Isoleucine Conformation" /></td>
</tr>
<tr>
<td>Met</td>
<td><img src="image" alt="Met Structure" /></td>
<td><img src="image" alt="Met Conformation" /></td>
</tr>
<tr>
<td>Methionine</td>
<td><img src="image" alt="Methionine Structure" /></td>
<td><img src="image" alt="Methionine Conformation" /></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td><img src="image" alt="Phenylalanine Structure" /></td>
<td><img src="image" alt="Phenylalanine Conformation" /></td>
</tr>
</tbody>
</table>

**Neutral, Hydrophobic, Aromatic**
<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Planar representation</th>
<th>Side chain conformation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neutral, Hydrophobic, Aromatic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td><img src="image1" alt="Tyrosine structure" /></td>
<td><img src="image2" alt="Tyrosine side chain" /></td>
</tr>
<tr>
<td>Tyr</td>
<td><img src="image3" alt="Tyr planar" /></td>
<td><img src="image4" alt="Tyr side chain" /></td>
</tr>
<tr>
<td>Y</td>
<td><img src="image5" alt="Y planar" /></td>
<td><img src="image6" alt="Y side chain" /></td>
</tr>
<tr>
<td>Tryptophan</td>
<td><img src="image7" alt="Tryptophan structure" /></td>
<td><img src="image8" alt="Tryptophan side chain" /></td>
</tr>
<tr>
<td>Trp</td>
<td><img src="image9" alt="Trp planar" /></td>
<td><img src="image10" alt="Trp side chain" /></td>
</tr>
<tr>
<td>W</td>
<td><img src="image11" alt="W planar" /></td>
<td><img src="image12" alt="W side chain" /></td>
</tr>
</tbody>
</table>

| **Neutral, Hydrophilic** |
| Serine | ![Serine structure](image13) | ![Serine side chain](image14) |
| Ser | ![Ser planar](image15) | ![Ser side chain](image16) |
| S | ![S planar](image17) | ![S side chain](image18) |
Table 1.1. Amino acid Structures and Properties (continued)

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Planar representation</th>
<th>Side chain conformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threonine</td>
<td><img src="image" alt="Threonine Structure" /></td>
<td><img src="image" alt="Threonine Side Chain" /></td>
</tr>
<tr>
<td>Asparagine</td>
<td><img src="image" alt="Asparagine Structure" /></td>
<td><img src="image" alt="Asparagine Side Chain" /></td>
</tr>
<tr>
<td>Glutamine</td>
<td><img src="image" alt="Glutamine Structure" /></td>
<td><img src="image" alt="Glutamine Side Chain" /></td>
</tr>
</tbody>
</table>
Table 1.1. Amino acid Structures and Properties (continued)

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Planar representation</th>
<th>Side chain conformation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acidic, Hydrophilic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspartic acid</td>
<td><img src="image" alt="Aspartic acid structure" /></td>
<td><img src="image" alt="Aspartic acid side chain" /></td>
</tr>
<tr>
<td>Asp</td>
<td>Asp</td>
<td>D</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td><img src="image" alt="Glutamic acid structure" /></td>
<td><img src="image" alt="Glutamic acid side chain" /></td>
</tr>
<tr>
<td>Glu</td>
<td>Glu</td>
<td>E</td>
</tr>
<tr>
<td><strong>Basic, Hydrophilic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histidine</td>
<td><img src="image" alt="Histidine structure" /></td>
<td><img src="image" alt="Histidine side chain" /></td>
</tr>
<tr>
<td>His</td>
<td>His</td>
<td>H</td>
</tr>
</tbody>
</table>
Table 1.1. Amino acid Structures and Properties (continued)

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Planar representation</th>
<th>Side chain conformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td><img src="image" alt="Lysine structure" /></td>
<td><img src="image" alt="Lysine side chain" /></td>
</tr>
<tr>
<td>Arginine</td>
<td><img src="image" alt="Arginine structure" /></td>
<td><img src="image" alt="Arginine side chain" /></td>
</tr>
<tr>
<td>Cysteine</td>
<td><img src="image" alt="Cysteine structure" /></td>
<td><img src="image" alt="Cysteine side chain" /></td>
</tr>
</tbody>
</table>

**Thiol - Containing**
<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Planar representation</th>
<th>Side chain conformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proline</td>
<td><img src="image1.png" alt="Proline Planar" /></td>
<td><img src="image2.png" alt="Proline Side Chain" /></td>
</tr>
<tr>
<td>Pro</td>
<td><img src="image3.png" alt="Pro Planar" /></td>
<td><img src="image4.png" alt="Pro Side Chain" /></td>
</tr>
</tbody>
</table>

Table 1.1. Amino acid Structures and Properties (continued)
Normal aerobic environments, the stable disulphide form is the most prevalent, although free thiols are also present in some proteins, where they often function as ligands for metal chelation, or as nucleophiles in proteolytic enzymes such as in papain or as carbonyl activators in acyl transferases. The secondary amino acid proline is an α-imino acid and its cyclic structure of the secondary amine confers conformational effects on peptide and protein backbone. It often plays an important role in stabilising or influencing the secondary structure of proteins and is often present at hairpin turns in globular proteins and it is a major component of the structural protein collagen.

Some amino acids present in proteins are enzymatically modified post-translationally to give rise to new amino acids (Fig. 1.3). For example, collagen containing proline can be hydroxylated to yield trans-4-hydroxyproline (Hyp). Insufficient hydroxylation of proline in collagen results in a deficiency disease called scurvy. Another special amino acid is γ-carboxyglutamic (Gla) produced by carboxylation of glutamic acid by a vitamin K-dependent carboxylase.
Pyroglutamic Acid

γ- Carboxyglutamic Acid

Tyrosine O-Sulphate

Tyrosine O-Phosphate

Serine O-Phosphate

trans-4-Hydroxyproline

Figure 1.3, Some amino acids modified by post-translation
Defective carboxylation of glutamate in proteins results in faulty blood coagulation. The reversible phosphorylation of hydroxyl groups of tyrosine, serine and threonine by kinases and phosphatases play a vital role in regulating biological activity in target proteins. The hydroxyl of tyrosine in peptide hormones such as gastrin and cholecystokinin may also be sulphated. Glutamine, which is chemically unstable at the amino terminus of a protein or peptide in aqueous solution can undergo a non-enzymatic transformation. The conversion to pyroglutamic acid (pyrrolidone carboxylic acid) occurs by spontaneous cyclization of glutamine in these environments. The internal glutamine residue in a protein or peptide when exposed by proteolysis of an X-Gln bond (where X represents any amino acid) or during protein sequencing promotes this transformation.

Naturally occurring amino acids that are not found in proteins also exists in some numbers. The γ-aminobutyric acid amongst them play an important role as neurotransmitters. Others like ornithine are intermediates in metabolic pathways or like the dihydrophenylalanine, are precursors to amino acid-derived products such as catecholamines. The prokaryotes and the lower eukaryotes, especially in sponges, algae, yeasts and fungi have the largest number of non-protein amino acids, although peptide with unusual amino acids have also been found in chordates, such as tunicates as well. These amino acids are assembled into peptides by enzymes without the aid of ribosomes or RNA. So far 700 of these non-protein amino acids are known, and variations in their structure are immense (Hunt 1985). In the majority of cases these amino acids are usually the products of secondary metabolism, and their function in these
organisms are obscure. Some of these unusual amino acids when present in proteins can impart unusual biological properties. Such biomolecules have been the hallmark of biochemical tools such as antibiotics, immunomodulators or antineoplastic agents that have therapeutic application.

1.3.2 Side-Chain Interactions

The side-chains of amino acids can interact with each other, with the amide peptide backbone, with the surrounding solvent, and via non-covalent interactions. In covalent interactions cysteine thiols promote disulphide bond formation.

1.3.2.1 Hydrogen Bonds and Salt Bridges

Efficient solvation tend to be present in proteins with polar side-chains. Both acidic (Asp and Glu) and basic (Lys, Arg, and His) residues generally are located on surfaces of proteins where the charged groups of the side-chains are exposed to the surrounding solvent while the alkyl part of the Lys and Arg side chains is usually buried. Charged side-chain residues that become buried internally due to folding normally are involved in salt bridges, where acidic and basic side-chains can either form ionic bond with each other or linked via a water molecule (Baker et al., 1984).

Ser, Thr, Asn, Gln and Tyr which all possess non-ionic polar side-chains are also extensively hydrogen bonded, either to the bulk solvent or to the peptide backbone, other side-chain groups or to associated water molecules (Thanki et al., 1990). The side-chains of Ser, Thr, and Asn often in helices
participate in forming hydrogen bonds with the carbonyl oxygen of the third or fourth residue earlier in the sequence which may contribute to the stabilization of helical segments (Gray et al 1984). The formation of low energy intramolecular hydrogen-bonded aggregates or salt bridges in short peptides may play a vital role in association with macromolecular targets or receptor, although side-chain hydration usually occurs through the bulk solvent. The criteria of solvation and dissolution is an important factor when working with short peptides which is often overlooked in structure function studies.

1.3.2 2 Hydrophobic Interactions

Unlike the hydrophilic residues which tend to be well solvated, the hydrophobic residues have an aversion to aqueous environment. This phenomena is largely entropic, highlighting the unfavourable free energy involved in forming a water-hydrocarbon interface, where the side-chain would have to project into the aqueous phase (Tanford et al., 1973). It has been well documented that soluble proteins contain highly hydrophobic regions in the interiors and that proteins fold in a particular fashion so as to exclude as much as possible hydrophobic side-chains on the protein surface. It has been postulated that the pathway to protein folding proceeds through the formation of hydrophobic clusters which then rapidly reorganizes the peptide chain into various favourable low-energy secondary structures such as helices and \( \beta \)-structures (Rose et.al 1980). The influence of the hydrophobic effect to protein stability has been investigated by susceptibility to denaturation of a number of proteins in which single residues in the protein sequence were substituted by
site directed mutagenesis (Kellis et al., 1988). It was evident that the elimination of even a single methyl group (Ile to Val substitution) destabilised the protein by 1.1 kcal/mol., illustrating the importance of combined effect of hydrophobic effect on overall protein structure and stability. Despite the entropy-driven process of distributing non polar residues away from aqueous environment, also specific interactions of hydrophobic residues do occur. These interactions are dipole-induced dipole interactions called the van der Waals forces or London interactions (Burley and Petsko 1988). Compared to salt bridges and hydrogen bonds of hydrophilic residues, they are weak bonds with possible important role in local secondary structure and protein interactions. This particular kind of protein-protein interaction is found in proteins that play an important role in DNA binding. The aromatic rings in proteins also interact in this fashion but not typified by the parallel stacking of base pairs in DNA helix. This is due to electron-rich n-cloud found in aromatic residues arranged in parallel to and above and below the plane of the ring possessing positively polarized hydrogen atoms in the plane of the ring (Burley and Petsko 1988). The interactions of the edge of one ring with the face of the other in a perpendicular mode favours the interaction between the positively polarized hydrogens and the negatively polarized n-cloud. Peptides on the other hand due to their length do not favour the hydrophobic residues to arrange themselves in such a manner so as to be totally shielded from the solvent. This situation creates poor solvation and often contributes to solubility problems compared to proteins. For these reasons organic solvents are sometimes employed to achieve partial solubility which can be undesirable for biological
structural studies. Hydrophobic peptides are often found to be extensively aggregated with increasing concentration. Solubility of peptides can sometimes be improved by increasing the length, thus inducing stable secondary structures where the hydrophobic residues can be segregated.

1.3.2.3 Disulphide Bonds, Thioesters, and Thioethers

Cysteine residues can form a number of covalent linkages with other amino acid side-chains in contrast to the side-chain interactions of other residues which are non-covalent. The prevalent form is the disulphide bond formation under oxidative condition involving two cysteine thiol groups (Fig. 1.4). The sulphur-sulphur bond in a disulphide is rigid and is not freely rotatable, just like the amide bond. With a torsional angles in the region of either $+90^\circ$ or $-90^\circ$, it can exist in one of two rotamers. The complete disulphide moiety, $\text{CH}_2\text{-S}\text{-S}\text{-CH}_2$, can rotate as a unit by simultaneous rotation about the side-chain angles $\chi_1(\text{NH-CH}_2\text{-S})$ and $\chi_2(\text{CH}_2\text{-S})$. 
Figure 1.4. Disulphide bond formation on oxidation with two molecules of cysteine, and the reverse reaction under appropriate conditions.
Disulphide bonds play an important role in stabilizing tertiary structures in proteins. In biologically active peptides such as vasopressin, oxytocin, and somatostatin (Fig. 1.5) which all contain disulphide bridges, it maintains a stable conformation, without which, the peptide being small cannot maintain a stable conformation. The reduced form or substitution with a different residue can result in a complete or partial loss of biological activity. In a majority of cases local flexibility exist in disulphide containing peptides like vasopressin and it does not exhibit any form of predominant structure. The disulphides maintain and impose constraint to peptides to adopt conformation that otherwise would exist in a linear form. Cysteine residues are capable of forming thioester and thioether linkages. Activated acyl transferases and thiol proteases such as papain all possess thioesters in their acyl groups. In the binding sites of the complement C3b and \( \alpha_2 \)-macroglobulin, where the sequence Cys-Gly-Glu-Glu include a thioester involving the Cys'thiol and the Glu\(^4\) side-chain carboxylate (Sothrup-Jensen et al., 1980). In the peptide antibiotics such as the subtilin and nisin, cysteine in the form of alipathic thioether linkage is evident. The mushroom toxins, amanitin and phalloidin contain aromatic thioether linkages such as Cys-Trp (Wieland, 1968). Recently an unusual thioether Cys-Tyr conjugate was located in the enzyme galactose oxidase (Ito et al., 1991).
Figure 1.5 Peptides containing disulphide bonds
1.4 THE AMIDE BOND

In a protein or peptide, the amino acids are joined together through the α-carbonyl and α-amino group by an amide linkage. This is known as a peptide bond. The polyamide peptide backbone plays an important role in contributing to the overall protein and peptide structure. The amide bond is not adequately represented by the formula -CO-NH-, since this fails to show that there is a considerable resistance to rotation about the C-N bond. It is therefore appropriate to depict the amide bond as a planar unit with the carbonyl carbon, oxygen, nitrogen, and amide hydrogen all lying in the same plane, with extensive delocalization of the lone pair of electrons on the nitrogen atom imparting a substantial double-bond character to the C-N bond (Fig.1.6). This structure gives the amide bond several characteristics that are important in protein and peptide structure. Because of the partial double-bond character of the carbon-nitrogen bond, no free rotation occurs about this bond (the barrier to rotation is about 80kJ mol⁻¹ (Crieghton, 1979). The torsional angle of that bond, $\omega$ is defined by the peptide backbone atoms Cα-C(O)-N-Cα. The partial double-bond character confers two rotational isomers for the peptide bond: trans ($\omega = 180^\circ$) and cis ($\omega = 0^\circ$) (Fig.1.7). The trans configuration is generally the more stable and is often found on peptide bonds not involving proline. In amide bonds where, a proline residue is present, the energy involved in the trans X-Pro bond is high, and both the difference in energy between cis and trans isomers and the barrier to rotation is lowered.
Figure 1.6. Torsional angles of the peptide backbone.

$\phi = \psi = 180^\circ$ corresponds to fully extended polypeptide chain.
(a) Resonance forms of amide bond

trans amide bond
$\omega = 180^\circ$

cis amide bond
$\omega = 0$

(b) trans and cis amide bonds

Figure 1.7 Amide bonds
The cis-trans isomerism about the X-Pro bond is often located in peptides containing prolines. Using NMR the chemical shifts of some hydrogens, especially those on the proline δ-carbon are often different and can be detected. The rate of isomerization and the equilibrium ratio of cis to trans isomers are highly sequence dependant. The cis configurations has been shown empirically to increase in X-Pro peptides where X is a bulky, hydrophobic amino acid (Harrison et al., 1990). The X-ray crystal structure of several proteins has revealed specific cis X-Pro peptide bonds and only now it is being appreciated that it might play an important role in proteins (Frommel and Preissner, 1990). In protein refolding studies it has been postulated that proline isomerization might be involved in the transition (Kim and Baldwin, 1982). Recently, information has been gained regarding the immunosuppressive peptide cyclosporin A where the binding protein cyclophilin catalyzes the peptidyl proline to cis-trans isomerization (Fisher et al., 1989, Takahashi et al., 1989). Some unrelated immunosuppressive macrolides, such as FK506 which have a different mode of action from cyclosporin A, also have cellular binding proteins (immunophilins) which have proline isomerase activity (Harding et al. 1989). It is not well understood how proline isomerase activity is related to immunosuppressive activity, but the two activities can be inhibited by using synthetic analogues of cyclosporin or FK506 (Bierer et al. 1990).

The amide bond in its resonance forms imparts one other important characteristic in protein and peptide structure. The polar nature of the amide bond confers a significant dipole moment, which makes the amide carbonyl oxygen a particularly excellent hydrogen acceptor. Hydrogen bonds involving
The peptide backbone play a vital role in stabilizing protein secondary structures. Peptide bonds have a propensity to be solvated, by either bulk solvent or specifically bound water molecules, or by internal hydrogen bonds. This is evident in the hydrophobic interior of soluble protein molecules. In regular secondary structures such as helices and \(\beta\)-sheets, the peptide bonds also participate in an extensive network of hydrogen bonds involving internal polar side-chains, bound water molecules, and backbone interactions.

The polar nature of the amide bond can impart a net dipole to regular structures containing peptide bonds, such as helices. The partial positive charge of the amino terminus radiate towards the partial negative charge of the carboxyl terminus to yield an overall dipole in a helix, which can be an important factor in protein tertiary structure as well as a contributor to catalytic activity in enzymes. The dipole can participate in stabilizing interactions with substrates or to modify the \(pK_a\), or nucleophilicity, of catalytically active residues (Knowles, 1991).

1.5 **PROTEIN STRUCTURE**

Protein structure is organized in several hierarchical levels and can be conveniently divided into 4 levels of complexity (Richardson, 1981). The most basic level is the primary structure which may be defined as the linear sequence of L-amino acids linked by peptide bonds. The information which determines this sequence or primary structure is readily obtainable by chemical sequence analysis of the isolated, purified protein or by translation of the corresponding DNA coding sequence. The fact that only L-amino acids are
found in proteins (with rare exceptions in some peptides) fixes the "configuration" of these polymers. Further structural organization relates to the disposition of the polypeptide backbone in three dimensions and the orientations of the side-chains in relation to the backbone, which are collectively known as "conformation". Protein secondary structure is the first level of polypeptide folding stabilized primarily by hydrogen bonds between peptide bonds. The outcome of this interaction results in regular structural elements such as α-helices, β-sheets, turns and in addition "random coil". Tertiary structure describes how the secondary structural elements of a single protein chain interact with each other to fold into individual protein subunits to form a multimeric complex.

1.5.1 Protein Secondary Structure

Several types of regular structures can be promoted by the way chains of amino acid fold that are stabilized by interchain or intrachain hydrogen bonds in the amide backbone. Continuous regions of protein sequence can promote helices and turns that are stabilized by intrachain hydrogen bonds. Two or more chains can form sheets which are separated by intervening sequences (regions of secondary structure as well) and are stabilized by interchain hydrogen bonds. Regions of protein sequence that do not resemble helices, sheets or turns are termed random coil implying that no regular repeating structure exists. This does not imply that the conformation of the protein is random in those regions, because in any individual protein, the random coil sequences are just as highly ordered and reproducible just as other regions of
secondary structure.

The peptide backbone conformation can therefore be described by three torsional angles, namely, $\phi$ which is the angle defined by C(O)-N-Cα-C(O), $\psi$, which is defined by N-Cα-C(O)-N, and the amide torsional angle $\omega$, as depicted in Figure 1.6. These angles are defined by protein chemistry convention with respect to the peptide backbone. When the two carbonyl carbons are trans to each other, the angle $\phi$ is $180^\circ$, and the angle $\psi$ is $180^\circ$ with the two amide nitrogens in the \textit{trans} configuration to each other (Fig 1.7). The usual chemical definition for the angle $\psi$ is $+180^\circ$. Backbone torsional angles of $180^\circ$ would incur to fully extended peptide chain, just like a fully extended carbon chain.

1.5.1.1 Helices

The helix is one of the common types of secondary structure found in proteins, first postulated (Pauling et al., 1951; Pauling and Corey, 1951) where amino acid residues are wrapped around a central axis in a regular pattern. Helices can be uniquely defined by their $\phi$ and $\psi$ angles in a given planar, \textit{trans} peptide bond. Several types of helices are found in protein structures, the $\alpha$-helix being the most prevalent. These helices with their characteristic torsional angles are listed (Table 1.2). Helices are further characterized by the number of residues per turn and the hydrogen-bonding pattern, specifically the number of atoms in the cyclic structure formed by the hydrogen bond. Since there are approximately 3.6 amino acids per turn of the $\alpha$-helix and a total of 13 atoms in the ring closed by the formation of the hydrogen bond, the $\alpha$-helix is sometimes referred to as the $3.6_{13}$ helix. The pitch of the $\alpha$-helix is 0.54nm (Fig
1.8), which corresponds to a translation of 0.15nm per amino acid. Other types of helical structures are possible, such as $3_{10}$ helix which has 3 residues per turn and 10 atoms in the cyclic hydrogen-bonded structure. The $3_{10}$ helices are less stable than the usual $\alpha$-helix because the hydrogen bonds are somewhat distorted and short lengths of this type of helix are found in lysozyme and carbonate dehydratase. The $\alpha_{II}$ helix is similar to the $\alpha$-helix in terms of the hydrogen bonding patterns, but the plane of the peptide groups are not parallel to the central axis as with the $\alpha$-helix but tilted with the N-H bonds pointing inwards towards the axis. This type of helix is not a common occurrence in proteins but has been proposed to exist in bacteriorhodopsin (Krimm and Bandekar, 1986). There is also a handedness to helical structure which is defined by the screw sense of the helix. Observing down the helical axis from the amino terminal end, the protein or peptide backbone can be traced in a clockwise sense in a right-handed helix, towards the carboxyl terminus.
<table>
<thead>
<tr>
<th>Helix Type</th>
<th>$\phi$</th>
<th>$\psi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$-Helix (right-handed)</td>
<td>$-57^\circ$</td>
<td>$-47^\circ$</td>
</tr>
<tr>
<td>$\alpha$-Helix (left-handed)</td>
<td>$57^\circ$</td>
<td>$47^\circ$</td>
</tr>
<tr>
<td>$3_{10}$ Helix</td>
<td>$-60^\circ$</td>
<td>$-30^\circ$</td>
</tr>
<tr>
<td>Collagen Helix</td>
<td>$-51^\circ$</td>
<td>$153^\circ$</td>
</tr>
<tr>
<td></td>
<td>$-76^\circ$</td>
<td>$127^\circ$</td>
</tr>
<tr>
<td></td>
<td>$-45^\circ$</td>
<td>$148^\circ$</td>
</tr>
<tr>
<td>Polyproline</td>
<td>$-78^\circ$</td>
<td>$149^\circ$</td>
</tr>
</tbody>
</table>
Figure 1.8 (a) shows the α-helix; the bold lines trace the path of the polypeptide chain. (b) shows the hydrogen bonding arrangement in the α-helix more clearly.
Right-handed helices is the most common form that is found in proteins. The stability of helical structures are maintained by the intrachain hydrogen bonds. In a helix, the carbonyl bonds and amide NH bonds lie parallel to the helix’s axis with the carbonyls projecting downwards in the direction of the amino terminus. In an α-helix, the carbonyl of any residue $i$ is hydrogen-bonded to the amide NH of the $i+4$ residue (Fig. 1.8). In the $3_{10}$ helix, the $i$th residue carbonyl is hydrogen-bonded to the NH of the $i+3$ residue.

The helix in general, can be viewed as a polyamide cylinder, with a strong dipole moment imparted by the uniform arrangement of the carbonyl and NH groups, with the positive end at the amino terminus and the negative end at the carboxyl terminus. The side-chains of the amino acids are arranged radially outward from the helix when viewed down the helix axis. The interactions of helices with other peptide chains or small molecules occurs predominantly through the side-chain interactions and the helix dipole itself. This is due to all the backbone amide groups being involved in intrachain bonds. The formation of these intrachain hydrogen bonds in the helix can be equated as being internally solvated, thus promoting the helix to be more readily accommodating in a non-polar environment, such as a lipid bilayer. The membrane-spanning regions of transmembrane proteins are generally thought to be helical, with the axis of the helix perpendicular to the plane of the membrane (Eisenberg, 1984). Alamethicin, a pore-forming peptide antibiotic possess a highly helical structure and also a high degree of helix potential exists in hydrophobic stretch of the leader sequences of proteins that are transported through membranes.
1.5.1.2 β-Structures

If the formation of helices can be thought of as a form of secondary structure resulting from one strand of a peptide chain, then formation of β-structure requires two strands. Each peptide chain is generally in an extended conformation with interchain hydrogen bond occurring between the carbonyl and amide NHs of every other residue in the backbone. Adjacent strands adopt two possible arrangements: parallel, in which each peptide strand runs in the same direction, and antiparallel, in which the peptide strands run in opposite directions. These two arrangements are depicted in Figure 1.9. The β-structure can be considered as forming a surface or sheet, despite possessing a slight twist to it due to the peptide backbone for not being fully extended (typical angles are given in Table 1.3). The side-chain project above and below the rough plane of the sheet, with every other side-chain on one surface. The potential of forming larger structures involving many strands are possible because of the interchain hydrogen bonds between two adjacent chains involving only every other residue. These forms of extended β-structures are common features in proteins. The twist of β-structure also promotes sheets to adopt cylindrical structures, known as β-barrels another common structural motif.
Figure 1.9. (a) and (b) show parallel and antiparallel $\beta$-sheet structures respectively.
<table>
<thead>
<tr>
<th>Torsional Angles For $\beta$-Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
</tr>
<tr>
<td>Parallel $\beta$-Sheet</td>
</tr>
<tr>
<td>Antiparallel $\beta$-Sheet</td>
</tr>
</tbody>
</table>
1.5.1.3 Turns

The turns linking segments of secondary structure have been proposed to be fundamental in dictating the folded structures of native proteins (Fasman, 1989; Skolnick and Kolinski, 1990; Sibanda et al., 1989). In proteins, strands of α-helix or β-structure do not extend indefinitely, but rather fold back on themselves. Of the various other elements of regular structure in proteins the most important is the turns, which are involved in altering the direction of the peptide chain (Richardson and Richardson, 1989). Depending on the number of residues involved in the regular structure, turns can be classified as β-turns involving four amino acids in the sequence, while γ-turns contain three amino acids. Hydrogen bonds extending across the turn, stabilize each structure, in effect holding the two strands together. In turn structures it is usual to assign the first residue as \( i \). The hydrogen bond in a β-turn is between the carbonyl of the \( i \) residue and the NH of the \( i + 3 \) residue, resembling the starting point of an antiparallel β-structure. In a γ-turn the hydrogen bond is between the carbonyl of the \( i \) residue and the NH of the \( i + 2 \) residue, giving the equivalent of a seven membered ring. Turns are sometimes defined by the number of atoms in the ring structure formed by the hydrogen bond (Fig. 1.10). Therefore a β-turn would be a \( C_{10} \) turn and a γ turn a \( C_{7} \) turn. Therefore a range of β-turns are possible. Amide bonds being planar, β and γ turns can be identified by the torsional angles of the internal residues \( i + 1 \) and, in the case of β-turns, \( i + 2 \). Torsional angles associated with various types of β-turns and γ-turns are listed (Table 1.4).
Figure 1.10. Diagrammatic illustration of a type I $\beta$-turn in a protein
It is important to note that the inverse form of each turn in which the sign of the internal torsional angles is reversed can also be found. It is interesting to note that the type III $\beta$-turns is equivalent to a $3_{10}$ helix structure. Turns display several important structural features. Intramolecular hydrogen bonds are found with the rough plane created by the peptide backbone in a turn. In a $\beta$-turn, the amide bond between the $i+1$ and $i+2$ residues lies perpendicular to this plane. As it is not part of the hydrogen-bonding structure of the turns, its hydrogen bonding capacity must be fulfilled elsewhere. $\beta$-turns are more often located on protein surface, where the $i+1 - i+2$ peptide bond can be solvated by bulk solvent. In the hydrophobic interior of protein, a turn is often solvated by a bound water molecule (Smith and Pease, 1980). The side-chains of the $i+1$ and $i+2$ residues, which form the corners of the turn are also exposed apart from the amide bond. The cyclic nature of small peptides such as vasopressin or somatostatin, where a turn has been postulated, the corner residues are thought to play an important role in its biological activity (Rivier et al., 1976).

In a turn structure side-chains will adopt either a pseudo-axial or pseudo-equatorial arrangement around the cyclic hydrogen-bonded turn structure. Substituting a corner residue in a turn with a glycine or altering the chirality, imposes a change from pseudo-axial to the stable pseudo-equatorial without affecting the peptide backbone.
Table 1.4 Torsional Angles For Turns

<table>
<thead>
<tr>
<th>Type</th>
<th>$\phi$</th>
<th>$\psi$</th>
<th>$\phi$</th>
<th>$\psi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type $I\beta$</td>
<td>-60°</td>
<td>-30°</td>
<td>90°</td>
<td>0°</td>
</tr>
<tr>
<td>Type $I'\beta$</td>
<td>60°</td>
<td>30°</td>
<td>90°</td>
<td>0°</td>
</tr>
<tr>
<td>Type $II\beta$</td>
<td>-60°</td>
<td>120°</td>
<td>80°</td>
<td>0°</td>
</tr>
<tr>
<td>Type $II'\beta$</td>
<td>60°</td>
<td>-120°</td>
<td>-80°</td>
<td>0°</td>
</tr>
<tr>
<td>Type $III\beta$</td>
<td>60°</td>
<td>30°</td>
<td>-60°</td>
<td>30°</td>
</tr>
<tr>
<td>$\gamma$ turn</td>
<td>70-80°</td>
<td>-60-70°</td>
<td>60°</td>
<td>30°</td>
</tr>
<tr>
<td>Inverse turn</td>
<td>-70-85°</td>
<td>60-70°</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1.5.2 Tertiary Structure

This second level of folding gives rise to the overall shape of the complete polypeptide chain. This brings the secondary structural elements together in space to give components like "β-barrels" and "helix bundles". These elements are stabilized by the following types of interactions: hydrogen bonds, ionic bonds, van der Waals and hydrophobic interactions. These interactions between the secondary structural elements can also stabilize the elements themselves, such that a particular sequence may only be a stable helix in the context of the tertiary structure.

1.5.3 Quaternary Structure

The assembly of folded molecules into higher order complexes, for example, dimers, tetramers, and polymers, is known as quaternary structure. Polymer in this instance means a regular array of polypeptide monomers such as microtubules, fibrin etc. as distinct from amorphous, irregular aggregates such as precipitates. The interactions which give rise to these higher order complexes are specific to the components and are stabilized by the same three types of bonds listed above. In some cases, the organization of the domains in a multi-domain polypeptide can be considered to be quaternary in structure.
1.6 CHEMICAL SYNTHESIS OF PROTEINS

Proteins and polypeptides with important structure-function interrelationships isolated from animals, plants and microorganisms have been subjects of great interest for many years. Amongst them are many of the membrane associated proteins, neuropeptides, toxins, growth factors, antibodies and hormones. The realisation that these naturally occurring biological molecules have peptidyl structures encouraged the development of techniques for the organochemical synthesis of proteins and polypeptides with predetermined sequence. The concept of peptides and polypeptides was first introduced by Emil Fischer and their synthesis in the early 1900’s. The isolation, structure and synthesis of an active peptide hormone, the octapeptide oxytocin (Vigneaud et al. 1953) was a major accomplishment which initiated the advancement of peptide synthesis.

The synthesis of the 39-residue porcine adrenocorticotropic hormone (Schwyzer and Sieber, 1963) and the 27-residue secretin peptide (Bodansky and Williams 1967; Bodansky et al. 1967) was a monumental task undertaken with the then available technology and chemistry. Although subsequent rapid developments in genetic engineering made it possible to produce desired proteins, this process can be expensive and is generally unsuitable for obtaining desirable quantities with suitable rapidity. The need for biological investigation demanded chemical synthesis not only of the frequently inaccessible encoded gene products but also numerous analogues to fulfil the exacting requirements of the following applications: (a) Physicochemical studies, (b) Structure-activity determination, (c) Pharmacological investigation,
In recent years attention has therefore been focused on improving chemical techniques for polypeptide synthesis which has paved the way for basic research and biological membrane-function studies. The introduction of the solid-phase technique (Merrifield, 1963) has made the availability of polypeptides accessible as reagents in biomedical and biophysical studies. The classical methods of solution synthesis with its inherent inefficiency could not cater for these demands and attention. The solid-phase method has by far been the most successful and rapid. The popularity of this technique has given rise to an ever increasing number of syntheses and also synthesis of longer (Clark-Lewis et al. 1986; Gutte and Merrifield, 1969; Nutt et al. 1988; Schneider and Kent, 1988) peptides and some longer than 100 residues in length are becoming common nowadays.

1.7 THE SOLID-PHASE PRINCIPLE

The concept of all solid-phase synthesis (Fig. 1.11) involves retaining the chemistry used in solution synthesis (protection groups) but adding a covalent linkage step (anchoring) that attaches the growing chain be it peptide, oligonucleotide (Gait, 1984) or other suitable oligomer to a stable insoluble polymeric support.
Figure 1.11. Reaction scheme for solid-phase peptide synthesis.
This remains attached to the solid support throughout the synthesis cycle (deprotection/coupling) and is separated from soluble reagents and solvents by simple filtration and washing without incurring manipulative losses. The repetitive steps are quick and simple and amenable to automation. Once chain elongation has been accomplished, it is necessary to detach (cleave) the fully protected peptide from the solid support under conditions that are mild to labile residues in the sequence. Finally complete deprotection of the residues accompanied by purification and characterization are carried out in free solution. The main advantage of a solid-phase approach is that reactions are driven to completion by using excess soluble reagents promoting high efficiency and yields which can be recovered by filtration and washing without manipulative losses.

1.8 SOLID SUPPORTS

The essential feature which distinguishes solid-phase synthesis from other techniques is the solid support. The resin support is quite often a polystyrene suspension polymer cross-linked with 1% of divinylbenzene as cross-linking agent (Sarin et al., 1980). The resulting spherical beads when dry have an average diameter of 50μm with a functional value of 0.2 to 1.0 mmol/g. These beads in solvents commonly used for peptide synthesis, namely DMF and DCM swell 2.5 to 6.2 fold their original volume. Furthermore, as peptide chains are extended the dry volume increases to accommodate the increased mass and, most crucially the swollen volumes continue to increase (Sarin et al., 1980). This means that the polystyrene matrix and the peptide
chains are well solvated during the chemical reactions. This also means that reactions also occur within the interior of the cross-linked polymer matrix. Polymer supports where the insoluble support and peptide backbone have similar characteristics have also been developed (Atherton and Sheppard, 1989). Numerous solid supports have been investigated and several have been suitable for peptide synthesis. These have ranged from polyacrylamides, polymethylmethacrylate, polysaccharides, silica, phenolic resins and porous glass, but only the polyacrylamides have seen popular use.

1.9 PROTECTING GROUPS

To ensure specific formation of peptide bond between two structurally similar amino acids it is necessary that the amino group of one and the carboxy group of the other is inhibited from reacting in the coupling procedure. The measure of a successful synthesis lies in the choice and optimization of protection chemistry. The protecting groups also have the properties of altering the zwitterionic or dipolar character of amino acids and peptides. Trifunctional amino acids also require additional side chain protection to avoid participation of these reactive groups during chain assembly. The possibility of damage is always encountered during the final deprotection/cleavage step, since these are usually the harshest conditions.

Protecting groups which are retained until complete assembly of the peptide chain are labelled as "permanent" while those which are removed at intermediate stages are classed "temporary". The choice of protecting group combinations very much depends on the synthetic strategy chosen. In the vast
majority of solid-phase syntheses, a temporary amino protection and permanent side-chain protected amino acids are added stepwise in the C → N direction. Alternatively the orthogonal protection scheme can be used where two or more classes of groups with differing chemical labilities can be removed in any order and in the presence of the other classes. This scheme make it possible for the use of milder conditions because selectivity can be achieved on the differences in chemistry rather than reaction rates.

1.9.1 Amino Protecting Groups

In the "Merrifield technique" the chemistry is based on differential labilities (Fig 1.12a). The acid-labile "temporary" N°-Boc group is stable to nucleophiles and alkali and their derivatives are easily cleaved by anhydrous inorganic or organic acids (Barany and Merrifield, 1979). Boc removal is usually carried out with 50% TFA in DCM for 30 minutes. In the orthogonal approach where a base-labile "temporary" N°9-Fmoc group (Fig 1.12b) is introduced onto amino acids, has found widespread use in the solid-phase peptide synthesis. Fmoc removal is usually achieved with 20% piperidine in DMF for 10 to 18 minutes (Atherton et.al., 1978).
Figure 1.12. Protecting group strategies in SPPS. (a) "Merrifield" protection scheme, (b) Orthogonal protection scheme
1.9.2 Side-chain Protecting Groups

On utilizing a suitable N\textsuperscript{α}-amino protection group, compatible permanent protection groups for the reactive amino acid side chains must be selected. The selection is usually made in the context of minimizing potential unwanted side reactions. Problems are normally encountered either during the coupling steps or at the final cleavage/deprotection step. In solid phase peptide synthesis it is desirable if the permanent protection groups (Fig 1.13) on the reactive side-chains of the amino and the carboxyl-terminal resin linkage are cleaved simultaneously in a single step on completion of the synthesis. The selection of a benzyl ester for the peptide resin linkage fulfils this purpose and is also suitable for protection of amino acid side chains. Thus the benzyl ether, thioether, ester and urethane derivatives have been used for the side chain hydroxyl, thiol, carbonyl and amino functions.

1.10 ATTACHMENT TO SOLID SUPPORTS

In solid-phase peptide synthesis the chain elongation is initiated in the carboxyl → amino direction where the carboxyl residue of the selected sequence is attached to the support either directly or by means of an appropriate "handle" (Fig. 1.11). Appropriate "handles" may be selected so that on final cleavage a free acid or amide at the carbonyl terminus can be obtained. In Boc chemistry, peptide acids are generated using substituted benzyl esters that are cleaved in strong acid and at the same time as other benzyl derivatives are removed.
<table>
<thead>
<tr>
<th>Side Chain Protecting Groups</th>
<th>Protected Amino Acid</th>
<th>Stability</th>
<th>Removal</th>
</tr>
</thead>
<tbody>
<tr>
<td>tert-Butyl</td>
<td>Asp/glu (OtBu)</td>
<td>Base</td>
<td>TFA</td>
</tr>
<tr>
<td></td>
<td>Ser/Thr/Tyr (tBu)</td>
<td>Base</td>
<td>TFA</td>
</tr>
<tr>
<td>tert-Butyloxycarbonyl</td>
<td>Lys (Boc)</td>
<td>Base</td>
<td>TFA</td>
</tr>
<tr>
<td></td>
<td>His (Boc)</td>
<td>Base</td>
<td>TFA</td>
</tr>
<tr>
<td>Triphenylmethyl</td>
<td>His (Trt)</td>
<td>Base</td>
<td>TFA</td>
</tr>
<tr>
<td></td>
<td>Cys (Trt)</td>
<td>Base</td>
<td>TFA</td>
</tr>
<tr>
<td></td>
<td>Asn/Gln (Trt)</td>
<td>Base</td>
<td>TFA</td>
</tr>
<tr>
<td>tert-Butoxymethyl</td>
<td>His (Bum)</td>
<td>Base</td>
<td>TFA</td>
</tr>
<tr>
<td>2,2,5,7,8-Pentamethylchroman-6-sulfonyl</td>
<td>Arg (Pmc)</td>
<td>Base</td>
<td>TFA</td>
</tr>
</tbody>
</table>

Figure 1.13. Widely used "permanent Side-Chain Protecting Groups compatible with Fmoc Chemistry"
In Fmoc chemistry, peptide acids are generated using the 4-alkoxybenzyl alcohol resin/4-hydroxymethoxyphenoxy linker (Wang, 1972; Lu et al., 1981; Sheppard and Williams, 1982; Colombo et al., 1983; Albericio and Barany, 1985) which is cleaved by 50% TFA in DCM in 1 to 2 hours. A variety of supports and linkers are also available which can be cleaved using very dilute acid to generate protected peptide segments retaining side-chain protection groups. For generating carboxyl terminal peptide amide, the attachment procedure is a direct coupling of an N\textsuperscript{\textdegree}-protected amino acid to a suitable benzylhydrylamine resin. The linkage agent 4-methoxbenzhydrylamine (Matsueada and Stewart, 1981; Gaehde and Matsueada 1981) is compatible with Boc chemistry to yield peptide amides.

The benzhydrylamide system has also been adapted to create the TFA-sensitive 4-\(2',4',6'-\text{dimethoxyphenyl-aminomethyl}\)phenoxy amide linker (Rink 1987) to generate peptide amide which are compatible with Fmoc chemistry. All the anchoring linkages that ultimately provide peptide acids are esters, and for peptide amides are benzhydrylamide derivatives.

### 1.11 FORMATION OF THE PEPTIDE BOND

Carboxylic acids and amines do not form amides spontaneously. Hence the formation of the peptide bond entails activation of one of the participating components. To date no practical method is available for the activation of the amine component, but for the carboxyl components there are practically four major kinds of coupling techniques that are suitable for the stepwise introduction of N\textsuperscript{\textdegree}-protected amino acids for solid-phase synthesis. Coupling reagents are
always used in excess to drive the reaction to completion.

### 1.11.1 Coupling Reagents

In solid-phase mode an efficient and unambiguous peptide-bond formation requires chemical activation of the carboxyl component of the N°-protected amino acid. The classical approach to activation of the carboxyl component of N°-protected amino acid \textit{in situ} has been DCCI (Rich and Singh, 1979; Merrifield et al., 1988). Reactions are carried out at room temperature but the principal limitation is the formation of the insoluble DCU during the activation and acylation reaction (Fig. 1.14). The related carbodiimides used in solid-phase peptide synthesis such as DIPCDI and BMC have found wider application because the resultant urea coproduct is more soluble. Recently new reagents for \textit{in situ} activation have become widely accepted because of ease, fast reactions even between sterically hindered amino acids with the suppression of side reactions. Protocols involving BOP, HBTU, and TBTU (Fig. 1.15) have appropriately become popular in automated peptide synthesis (Fournier et al., 1988; Fields et al., 1991).
Figure 1.14 Coupling reaction in SPPS
Figure 1.15. In situ coupling reagents and additives suitable for SPPS
1.11.2 Preformed Symmetrical Anhydrides

PSA (Heimer et al., 1981) has been used by some research groups, mainly in Boc chemistry because of their high reactivity. They are generated in situ from the corresponding N’-protected amino acid using 2 to 4 equivalents of protected amino acid and one equivalent of DCCI in DCM (Fig. 1.14). The urea formed is removed by filtration, the solvent is exchanged to DMF for optimal couplings. Fmoc-amino acids PSA have been well studied and are commonly employed in solid-phase synthesis. There are some drawbacks using Fmoc PSA when intermediates are formed during PSA formation which can undergo rearrangement. Also not all Fmoc amino acids are readily soluble in DCM and require significant DMF for solubilization which can slow down the rate of activation. PSA is stable and is commonly prepared prior to coupling.

1.11.3 Active Esters

To achieve an unequivocal peptide bond formation it is essential that the activating group should not participate as an acylating agent. The acid chlorides and acid azides fulfil this requirement but the activated intermediates are prone to side reactions and rearrangement. These and other side reactions, as a consequence of overreaction, can be suppressed by the use of moderately reactive intermediates. Thus the extensively studied aminolysis of esters is apparently a useful alternative to acylation with mixed anhydrides because the alcohol components in esters cannot acylate amines.

Simple alkyl esters of protected amino acids undergo aminolysis but are slow in peptide bond formation. The phenyl esters with electronegative
substituents are more reactive and may aminolyse at ratios equivalent to the anhydrides. More recently workers have concentrated on HOBt (Fields et al., 1989) and ODHBt esters of protected amino acids. These esters which are not isolated can be readily synthesized from BOP, PYBOP, TBTU and HBTU activating agents and react rapidly with negligible side reactions (Fig. 1.15). OPfp esters (Kisfaludy and Schon, 1983; Green and Berman, 1990) are also efficient and their chemical structures suppress side reactions but react slower than PSA. The addition of HOBt increases the reaction (Hudson, 1990). Unfortunately, not all amino acid OPfp's esters are suitable, for example, the Fmoc-Thr(tBu)-Opfp and Fmoc-Ser(tBu)-Opfp do not crystallize and are difficult to isolate. The ODHBt on the other hand provide good alternatives and react very fast (Cameron et al., 1988). The attractive property of ODHBt is the appearance of a bright yellow colour in the presence of a free N®-amino group on deprotection which is eliminated upon coupling and its detection can be utilized for monitoring. This monitoring technique has been applied throughout the synthesis in this study.

1.12 THE SYNTHESIS CYCLE

The synthesis cycle begins by the attachment of the first amino acid reversibly to the solid support and the removal of an N®-protection group. This constitutes the deprotection step which must proceed to completion, when the next N®-protected and side chain protected amino acid is added as an activated ester or preformed symmetrical anhydride. When this coupling step is complete, a further cycle commences from the deprotection step (Fig.1.11).
When the desired peptide chain is achieved, depending upon the type of linker group used, the peptide can be liberated with a free C-terminal, carboxyl C-terminal amide or fully protected peptide fragment for the synthesis of longer peptides.

1.13 DEPROTECTION AND DETACHMENT FROM SOLID SUPPORT

When a synthesis has been successfully completed, the solid support bears a fully protected peptide. The stage is set for the difficult task of having to simultaneously detach the peptide from the resin support and removing all the side-chain protecting groups of the amino acid residues to yield the desired peptide. This procedure has always been accompanied by deleterious side-reactions involving certain side chain protecting groups which are liberated as stable cations during TFA cleavage and deprotection. These unwanted reactions are exacerbated in Boc-based synthesis where strong acids such as HF, TFMSA or TMSOTf are used. It is usually customary to incorporate mixtures of scavengers to cleavage solutions to quench all the liberated reactive carbonium ions originating from protective groups and linkers on the resin (Lundt et al., 1978).

Once the peptide has been released from the resin support and simultaneously deprotected, it is then necessary to apply stringent purification procedures. This procedure involves the precipitation of the cleaved crude peptide using cold diethyl ether which also removes the scavengers which are present in the mixture. The precipitated peptide is then further subjected to chromatographic separations after which amino acid analysis, molecular mass
determination and capillary zone electrophoresis are applied to confirm the integrity of the synthesized product. However these criteria are sometimes difficult to apply to very hydrophobic polypeptides and alternative techniques for the analysis the must be found.

1.14 BIOLOGICAL MEMBRANES

Biological membranes are composed of mainly lipids, proteins and carbohydrates in somewhat variable proportions (Table 1.5). Since the sugar residues are always attached to either the protein or lipid components, or both, one can actually consider membranes to be composed of proteins and glycoproteins, and of lipids and glycolipids (Fig. 1.16). The basic framework of all biological membranes is a double layer of lipid molecules, an arrangement originally proposed by E. Gorter and F. Grendel of the University of Leiden in 1925. Membranes carry out a number of important functions in a living organism. They act as structural barriers that maintain the integrity of cells. They are selective permeability barriers for the passage of molecules into and out a cell or organelle by utilizing specific pumps and gates within. These transport mechanisms also regulate the molecular and ionic composition of the intracellular medium.
### Table 1.5 Chemical Composition of Cell Membranes

<table>
<thead>
<tr>
<th>Membrane</th>
<th>Protein %</th>
<th>Lipid %</th>
<th>Carbohydrate %</th>
<th>Weight fraction of protein</th>
<th>Ratio of protein to lipid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myelin</td>
<td>18</td>
<td>79</td>
<td>3</td>
<td>0.18</td>
<td>0.23</td>
</tr>
<tr>
<td>Plasma membranes (blood platelets)</td>
<td>33-42</td>
<td>51-58</td>
<td>7.5</td>
<td>0.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Mouse liver cells</td>
<td>46</td>
<td>54</td>
<td>2-4</td>
<td>0.46</td>
<td>0.85</td>
</tr>
<tr>
<td>Rat liver cells</td>
<td>58</td>
<td>42</td>
<td>5-10</td>
<td>0.58</td>
<td>1.4</td>
</tr>
<tr>
<td>Human erythrocyte</td>
<td>49</td>
<td>43</td>
<td>8</td>
<td>0.49</td>
<td>1.1</td>
</tr>
<tr>
<td>Amoeba</td>
<td>54</td>
<td>42</td>
<td>4</td>
<td>0.54</td>
<td>1.3</td>
</tr>
<tr>
<td>L cells</td>
<td>60</td>
<td>40</td>
<td>5-10</td>
<td>0.60</td>
<td>1.5</td>
</tr>
<tr>
<td>Hela cells</td>
<td>60</td>
<td>40</td>
<td>2.4</td>
<td>0.60</td>
<td>1.5</td>
</tr>
<tr>
<td>Retinal rods (Bovine)</td>
<td>51</td>
<td>49</td>
<td>4</td>
<td>0.51</td>
<td>1.0</td>
</tr>
<tr>
<td>Mitochondrial outer membrane</td>
<td>52</td>
<td>48</td>
<td>2-4</td>
<td>0.52</td>
<td>1.1</td>
</tr>
<tr>
<td>Sarcoplasmic reticulum</td>
<td>33</td>
<td></td>
<td></td>
<td>0.67</td>
<td>2.0</td>
</tr>
<tr>
<td>Chloroplast lamellae spinach</td>
<td>70</td>
<td>30</td>
<td>6</td>
<td>0.70</td>
<td>2.3</td>
</tr>
</tbody>
</table>
Table 1.5 Chemical Composition of Cell Membrane (continued)

<table>
<thead>
<tr>
<th>Membrane</th>
<th>Protein (%)</th>
<th>Lipid (%)</th>
<th>Carbohydrate</th>
<th>Weight fraction of protein</th>
<th>Weight fraction of protein to lipid (%)</th>
<th>Ratio of protein to lipid (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitochondrial inner membrane</td>
<td>76</td>
<td>24</td>
<td>1-2</td>
<td>0.76</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>Gram-positive bacteria</td>
<td>75</td>
<td>10</td>
<td>0.75</td>
<td>3.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Halobacterium purple membrane</td>
<td>75</td>
<td>25</td>
<td>-</td>
<td>0.75</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma</td>
<td>58</td>
<td>37</td>
<td>1.5</td>
<td>0.58</td>
<td>1.6</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1.16. The fluid-mosaic model membrane
Eukaryotic (nucleated) cells also contain numerous intracellular organelles of widely differing structure and function, each of which is specialized in its function: digestion (lysosomes), respiration (mitochondria), photosynthesis (chloroplast), secretion (endoplasmic reticulum and Golgi apparatus), or nucleic acid biosynthesis (nucleus). The flow of information between cells and their environment to external stimuli are processed by specific receptors in the membranes. In turn, each of these organelles is bounded by its own specialized membrane system that has evolved to participate in these specialized functions.

In some specialized membranes signals are generated in electrical or chemical form. Thus, membranes play a crucial role in biological communications. Finally there exist two very important energy conversion processes in biological systems which are carried out by membrane-associated enzymes and other proteins. In photosynthesis, light energy is converted into chemical-bond energy in the inner membranes of chloroplast whereas in oxidative phosphorylation where ATP is generated in the inner membranes of mitochondria. In contrast, prokaryote cells (bacteria) typically have all these functions integrated into the plasma membrane and lack specialized intracellular organelles. It is clear, however that these differences in cell structure do not indicate divergent fundamental biochemical mechanisms, but merely the presence or absence of compartments specifically designed to fulfil essential functions.
Nevertheless a number of fundamental principles have emerged that appear to apply to most systems that have been studied:

1) Membranes are generally very thin sheet like structures, being between 6 and 10nm across. This implies that they can only be a few molecules thick. (compositions)

2) Although the chemical compositions vary, membranes consist mainly of lipid and protein with carbohydrate. The carbohydrate is not free but covalently linked to the molecules of lipid and protein.

3) The lipid components of the membrane are relatively small molecules that provide the basic matrix of the membrane, and serves as a permeability barrier.

4) The protein component of the membrane is responsible for the specific function of the membrane.

5) The lipid and protein components of the membrane are held together by many non-covalent interactions, which are cooperative in character. The protein can be pictured as being in a two-dimensional solution of lipid.

6) Membranes can be viewed as asymmetric, that is, the inside and the outside faces of the membrane are different.

1.15. LIPIDS

Biological membranes are complex structures that are a composite of proteins and lipids. Membrane lipids can be divided into two broad classes, each having several subdivisions. The amphipathic lipids are the phospholipids which include glycolipids, phosphoglycerides, and sphingolipids. Despite their apparent chemical diversity membrane lipids all have one common feature upon
which rests the structure of the membrane. This feature is that they are amphipathic, which is to say, that one end of the molecule is charged or polar and mixes well with water whereas the other end of the molecule is non-polar and mixes poorly with water (Fig 1.17). The second broad class of lipids is the sterol family, whose principal representative in mammalian cell membranes is cholesterol. It, too, is amphipathic, since its hydroxyl group is polar and the steroid ring system and hydrocarbon "tail" are non-polar. However, the affinity of the hydroxyl group for water is much less than the affinity of an ionic head for water, and the amphipathic properties of cholesterol do not dominate its physical properties (Fig 1.17).
Figure 1.17. The amphipathic nature of phospholipids and cholesterol
1.15.1 Fatty Acids

The most abundant lipids are esters of the alcohol, glycerol. If the glycerol is esterified at all three positions, the product is a triacylglycerol (triacylglyceride, Fig. 1.18). Triacylglycerols are also known as neutral lipids, are the major components of adipose tissue and plant oils, but they are not found in membranes. The most common "tails" of amphipathic lipids are esters of long chain fatty acids. These carboxylic acids usually contain an even number of carbon atoms. The fatty acids found in the lipids of mammalian cells are unbranched, with an even number of carbon atoms, generally between 14 and 24. Many contain double bonds with a predominating cis configuration. Fatty acids containing double bonds are referred to as "unsaturated" whereas those with no double bonds are "saturated" (Fig 1.19). The saturated fatty acids can exist in extended conformations that can be tightly packed in well-ordered crystals. Therefore, they have relatively high melting points, the homologues above 12 carbons are solids at room temperature. A cis double bond puts a "kink" in the hydrocarbon chain. Most phospholipids contain one saturated and one unsaturated fatty acid chain. The kinked tails cannot form closely packed, well-ordered crystals as well as saturated fatty acids, and the unsaturated fatty acids have lower melting points than their saturated counterparts.
Figure 1.18. Structure of a triglyceride. $R_1$, $R_2$ and $R_3$ are hydrocarbon chain of the esterified fatty acids.
Figure 1.19. Structural arrangement of the carbon chains in saturated and unsaturated fatty acids. (a) saturated; (b) unsaturated trans form; (c) unsaturated cis form
The degree of unsaturation of the hydrocarbon tails of membrane lipids affects the properties of membranes because the fluidity of membranes is determined by the proportions of saturated and unsaturated fatty acids. About 50% of the fatty acids isolated from membrane lipids from all sources are unsaturated. Bacteria grown at different temperatures have different ratios of saturated and unsaturated fatty acids in their membranes, enabling them to maintain the same fluidity at different temperatures. Straight chain fatty acids predominate in plants and mammals, but microorganisms often contain branched-chain fatty acids.

1.15.2 Phospholipids

The most abundant of all the membrane lipids are the phospholipids. The phospholipids are literally lipids which contain a phosphate group, and there are two main types. The most plentiful are the phosphoglycerides, in which one of the three alcohol groups of glycerol is esterified to a phosphate containing group in place of a fatty acid. These phosphoglycerides can be subdivided on the basis of the particular kind of group in which the phosphate is present. The other common phospholipids are the phosphosphingosides, which are derived from the alcohol sphingosine.

1.15.2.1 Phosphoglycerides

The phosphoglycerides are the most common type of phospholipids and are related to phosphatidic acid, which is a compound of glycerol, two fatty acid residues, and one molecule of phosphoric acid (Fig 1.20). The free
phosphate hydroxyl groups can ionize, and therefore the molecule has both a polar and a non-polar section. The middle carbon of the glycerol portion of the molecule is asymmetric and the L isomer is the naturally occurring form. Most phosphoglycerides found in biological systems have one of the other phosphate hydroxyl groups linked to another compound, which gives the molecule its distinctive properties.

1.15.2.1 Choline Phosphoglycerides

One of the most important of the phospholipid substituents is choline, nitrogenous alcohol, which is found in the group of compounds called choline phosphoglycerides or lecithins. These substances are important constituents of biological membranes. A common name often used for α-phosphatidylcholine is α-lecithin (Fig. 1.20). Since this compound contains both phosphoric acid and the basic nitrogen of the choline, it normally exists as a zwitterion. The particular choline phosphoglycerides found in biological systems contain relatively few of the many possible fatty acids. However, these molecules can have two saturated fatty acids, two unsaturated fatty acids or one of each.
Figure 1.20. Structure of some phosphoglycerides. \( R_1 \) and \( R_2 \) are hydrocarbon chain of the esterified fatty acids
1.15.2.3 Ethanolamine Phosphoglycerides

Ethanolamine phosphoglycerides are particularly prevalent in brain tissues, although they are widely distributed in other tissues. Structurally, they are very similar to the lecithins and differ only in the presence of the amino alcohol ethanolamine, HO-CH₂-CH₂-NH₂ instead of choline. The characteristics of any specific ethanolamine phosphoglyceride depends on the fatty acids that it contains.

1.15.2.4 Other Phosphoglycerides

Choline and ethanolamine phosphoglycerides account for more than half the phosphoglycerides present in most mammalian tissues. However, there are a number of other important groups of phospholipids based on the simple glyceride structure. Many of these compounds have the general structure of phosphatidic acid and differ only in the substance which adjoins phosphoric acid. Phosphatidylserine is one of the most common examples. In place of serine residue either threonine or hydroxyproline may be found. Only amino acids which have a hydroxyl group are found in this type of molecule. A second variant on the more usual types of phosphoglyceride has a derivative of ethanolamine joined to the phosphate group. N-Methyl or N,N-dimethylethanolamine can replace ethanolamine in this way. A carbohydrate, such as the sugar alcohol myo-inositol, can also replace a nitrogenous alcohol of a phosphoglyceride to form the glycolipid phosphatidylinositol. Another group of phospholipids related to the glycerol phosphatides are the plasmalogens, so called because they were first described in blood cells. These
substances have the structure of a phosphoglyceride, except that the fatty acid attached to the carbon at the other end of the glycerol from the phosphate group is replaced by an aldehyde derivative of one of the common saturated fatty acids. When this aldehyde is combined with the glycerol hydroxyl, the result is a double bond between the \( \alpha \)- and \( \beta \)-carbons of the fatty acid derivative. Phosphatidyl choline is an example of one type of which is common in certain kinds of cell membranes.

1.15.2.5 Phosphosphingosides

The other main group of phospholipids, apart from the phosphoglycerides is that based on the 18-carbon amino alcohol sphingosine. The sphingomyelins have phosphoric acid and choline joined to the terminal hydroxyl group and a fatty acid joined to the amino group. It should be noted that the phosphosphingosides are quite similar in general size, shape, and charge to the phosphoglycerides, since the hydrocarbon chain of sphingosine is equivalent to the second fatty acid of the phosphoglycerides. The sphingomyelins are particularly common in mammalian tissue.

1.15.2.6 Properties of Phospholipids

Biologically, the significant feature of the phosphoglycerides and phosphosphingosides is the fact that these molecules possess both polar and nonpolar groups. Since biological systems consist of a suspension of insoluble particles in an aqueous environment, any available phosphoglyceride molecules tend to orient themselves to form micelles with their fatty acid portions,
making an internal hydrophobic region (Figure 1.21). Alternatively, they may form a monolayer coating the surface of a relatively hydrophobic particle so that the polar portions of the phosphoglycerides project out into the aqueous medium. The fact that the polar portion of these molecules is hydrophilic, while the nonpolar end is hydrophobic, also allows them to assemble as bimolecular layers (Fig 1.21). The natural tendency of phosphoglyceride molecules to assemble into biomolecular layers and to arrange themselves on interfaces may well be involved in the assembly of membranes. A bimolecular phosphoglyceride layer has dimensions similar to those of cellular membranes.

1.16 GLYCOSPHINGOSIDES AND STEROIDS

There are vast number of other substances which are included among the lipids. Some of these substances are structurally related to the neutral fats or phospholipids, but others, such as the steroids, are only related to the simpler lipids by the similarity of their physical properties.

1.16.1 Glycosphingosides

Glycosphingosides include the cerebrosides and gangliosides and, like the phosphosphingosides, are found in brain and nervous tissue. They are structurally similar to sphingomyelin, except that the phosphorylcholine is replaced by a hexose, or by a complex carbohydrate (Fig. 1.22). The cerebrosides appear to be an important constituent of the insulating myelin sheaths of nerve fibres.
Figure 1.21  Structures adopted by phospholipids in aqueous solution
Figure 1.22. Structure of galactocerebroside
1.16.2 Steroids

Steroids possess no obvious structural relationship to the triglycerides or phospholipids. They exist in a great variety of structural forms and many are characteristic of a particular group of animals or plants. Another important lipid component found in membranes is cholesterol. It is present in large amounts in eukaryotic cell membranes where it serves to maintain the structure by providing rigidity. Cholesterol is by far the most widely distributed and abundant steroid in animal tissue, and in humans it forms about 0.2% of the body weight. In addition to the steroid nucleus, cholesterol has an eight-carbon side chain, two extra methyl groups, a double bond, and a hydroxyl group (Fig. 1.17). It is a white crystalline substance which is insoluble in water and has a melting point of 150°C. There are a variety of steroids in animal tissues which are related to cholesterol. Some of the well known and important are the steroid hormones formed in the adrenal cortex and the gonads. There are also many steroids which are not hormones and their function is unclear, although it seems likely that some of them play a crucial role in the structure of biological membranes, and others certainly participate in important cellular reactions. Prokaryotic cells on the other hand (which have cell walls) do not have cholesterol in their cell membrane, an observation that supports the view that cholesterol is needed to provide rigidity.

1.17 MEMBRANE FLUIDITY

We now know that all biological membranes behave simultaneously like a fluid and solid, so liquid crystallinity must play a vital role in the function of
the cells. Over the past 23 years there has been a surge of interest in biological membrane fluidity. This has created novel and exciting ideas regarding phase behaviour which may influence the way a cell membrane operates. This led to the development of a simple coherent picture called the fluid mosaic model (Singer and Nicholson, 1972) of the cell membrane (Fig. 1.16). Information supporting the fluid behaviour of the lipid matrix has been gathered applying various techniques, including NMR (Chapman and Salsbury, 1966; Oldfield et al., 1981), IR spectroscopy and DSC calorimetry (Chapman et al., 1969). The data obtained in DSC experiments indicate a major phase transition of a cooperative nature in hydrated lipid bilayers and biomembranes (Chapman et al., 1969; Steim et al., 1969). At lower temperatures limited movement of the acyl chains of the lipid moiety is identified implying an ordered gel-like property. Above the transition temperature, substantial molecular motion is detected spectroscopically from IR and NMR studies. For simplicity, most of the research into the phase behaviour of the membrane, however ignores the proteins, and study only an idealized two-dimensional fluid fabric. Research groups are now trying to go one step further and apply this knowledge of these simplified systems to the behaviour of a living cell.

It can be illustrated that at the main phase transition, denoted by $T_m$, the lipid bilayers undergo transformations. In this mode the acyl chains show an increase in the gauche conformers. With the tendency of the bilayer with its acyl chains adopting a tighter packing order, any disorder imposed into the bilayer will disrupt the chain packing which in turn will be transmitted along the chain creating more disorder. This phenomena is a highly cooperative
interaction operating within a narrow range of temperatures (typically 1-2°C). The terminal portions of the acyl chains are prone to this rotational disorder (Chapman and Salsbury, 1966). The transition temperatures and enthalpies for various one-component bilayers have been measured (Chapman et al., 1974; Hinz and Sturtevant, 1972; Mabrey and Sturtevant, 1976). Two transitions are observed when PCs contain saturated hydrocarbon chains. The phase transition temperature of lipids with ionizable groups can be affected by changing the pH. Also the headgroup and the length of the hydrocarbon chain can influence the phase transition (Chapman et al., 1967, see table 1.6). Thus by varying the composition of the lipid components, the fluidity of membranes can be controlled, characteristic of numerous prokaryotes (Melchior et al., 1970; Baldassare et al., 1976) and protozoa (Dickens and Thomson, 1982) adapting to its environments. Also introduction of complex mixtures of lipids with a loose packing order rather than simple mixtures or single type of lipid can lower the $T_m$ value of the bilayer. The effect of many molecules incorporated into biomembranes can affect the membrane fluidity. As discussed earlier, cholestrol plays a crucial part in maintaining membrane fluidity in mammals. It has been shown that by adding cholesterol to simple bilayers, the $T_m$ and the enthalpy change of the transition are lowered. This is attributed to the cholesterol intercalating within the bilayer thus disrupting the packing order of the lipid molecules (Ladbrooke et al., 1968).
Table 1.6. The Influence Of Headgroup, Chain Length And Unsaturation Upon Transition Temperature ($T_m$) and Enthalpy Changes ($H$)

<table>
<thead>
<tr>
<th>LIPID</th>
<th>COMPOSITION</th>
<th>$H$ (Kcal. Mol$^{-1}$)</th>
<th>$T_m$ ($^\circ$C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMPC</td>
<td>14:0</td>
<td>6.26</td>
<td>23.7</td>
</tr>
<tr>
<td>DPPC</td>
<td>16:0</td>
<td>9.69</td>
<td>41.7</td>
</tr>
<tr>
<td>DPPE</td>
<td>16:0</td>
<td>8.0</td>
<td>63.0</td>
</tr>
<tr>
<td>DSPC</td>
<td>18:0</td>
<td>10.84</td>
<td>58.2</td>
</tr>
<tr>
<td>DOPC</td>
<td>18:1</td>
<td>7.6</td>
<td>-22.0</td>
</tr>
</tbody>
</table>
At higher transition temperatures of the bilayer cholesterol tends to sterically hinder rotational motion and also reduction in the ratio of gauche conformers. This unique property of cholesterol functions as a modulator of membrane fluidity in mammals within certain limits.

1.18 RECONSTITUTION OF MEMBRANE PROTEINS

Because of the complex nature of membrane proteins, studies involving the mechanism of how protein and lipid molecules interact in order to specify function can be a difficult task. A practical approach to studying membrane proteins is to isolate individual components and study them in simple isolated systems. With the technique of developing liposomes and planar lipid bilayers it has become possible to study membrane proteins isolated and purified from natural membranes on reconstitution into these membrane mimetic systems. Partial functions are sometimes restored and this has led to a greater understanding of the mechanisms involved. However there are many pitfalls in this approach as many of the membrane proteins function, in a cooperative manner with other membrane proteins. However gaining an insight into single components can sometimes lead to reasonable assumptions on the functional mechanism of events occurring in vivo which can then be investigated in the intact cell membranes. A great amount of information is known about the lipid bilayer portion of the membrane (section 1.15.2.6) because it is possible to prepare bilayers composed of a single, chemically defined class of phospholipids. To acquire a similar understanding of the interactions between lipid and protein molecules, a necessary criteria is to prepare membranes.
possessing a single class of phospholipid and a single type of protein. Therefore it is essential to reconstitute membranes from defined purified proteins and phospholipids. This is a monumental task because to obtain a single purified protein from a myriad of proteins present in most membranes is difficult. One way of overcoming these difficulties is to select a membrane that has a highly specialized function, the rationale being that the protein (or proteins) responsible for that function is (are) likely to predominate amongst the protein present in the membrane. The sarcoplasmic reticulum found in the skeletal muscle that modulates calcium concentration is rich in Ca\(^{2+}\) pump protein (ATPase), and bacteriorhodopsin found in the purple membrane of the Halobacterium halobium are some of the representative examples. However, with the advent of chemical synthesis of polypeptides, selected domains can now be studied of membrane bound proteins when normally such proteins are impossible to isolate intact. This approach was applied to the study of the pore forming region of K\(^+\) channel protein (in this laboratory) deduced from site directed mutagenesis studies (Haris et al., 1994). The goal of any reconstitution studies is to gain information as to the conformation of the transmembrane proteins and possibly the orientation of domains. Also in functional studies, ligands to receptor binding interactions such as toxins and hormones can give information as to the local structural elements involved. In addition to gaining meaningful functional data derived from solubilization and reconstitution techniques, it can also be used to extract detailed structural information. The use of detergents have played a vital role in the crystallization of membrane proteins suitable for spectral analysis (Michel, 1983). Although
detergents are useful in extracting natural membrane proteins, complications can set in when spectral analyses are carried out, especially when proteins begin to form insoluble aggregates due to their hydrophobic properties. Detergents have not been useful for disaggregation purposes, when applied to synthetic membrane proteins (Chapter 3). To date no particular reconstitution technique or detergent has been effective as a standard for membrane proteins. However when it comes to the solubilization and reconstitution of synthetic membrane polypeptides, an effective technique has been introduced successfully in this laboratory in elucidating conformation of membrane spanning proteins.

1.19 FUNCTIONS OF LIPIDS

Our understanding of their roles in biology is still fragmentary, largely due to the great difficulties encountered in studying lipids in aqueous solutions. In general, it can be stated that the importance of lipids has been underestimated in the past. This is particularly true of lipid metabolism. It has long been recognised that fats are essential for energy storage, but it is becoming increasingly clear that, quantitatively, the metabolism of fatty acids may provide most of the energy supply for some cells. It is even more difficult to obtain information about the structural significance of lipids. Not much is known about the mechanism of how membranes are formed, although it is clear that cholesterol and phospholipids are the major constituents. Future research may well illustrate that phospholipids are more than just building blocks in cellular membranes.
CHAPTER TWO

BIOPHYSICAL METHODS OF STRUCTURAL DETERMINATION
2.1 INTRODUCTION

In this chapter useful physical techniques to probe protein structure are discussed. Each of the techniques by itself provides only indirect and incomplete information, but the combined use of several of them can lead to very detailed and self consistent pictures of protein structure. The emphasis is on the physical nature of what is measured in each application and the way in which the data are used to obtain structural information. These techniques exploit the interaction of electromagnetic radiation (EMR) with the material under investigation. The characteristic interactions are related to the energy and hence the wavelength of the radiation. A schematic representation of the electromagnetic spectrum and its interaction with matter is depicted in figure 2.1.
<table>
<thead>
<tr>
<th>Nuclear electrons</th>
<th>Innershell electrons</th>
<th>Ionisation</th>
<th>Valence electrons</th>
<th>Molecular vibrations</th>
<th>Electron spin orientation</th>
<th>Nuclear spin orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cosmic and Gamma rays</td>
<td>X-rays</td>
<td>(UV) Vacuum Ultraviolet</td>
<td>(UV) Near Ultraviolet</td>
<td>Visible</td>
<td>(IR) Near Infrared</td>
<td>(IR) Infrared</td>
</tr>
</tbody>
</table>

| Wave length (nm) | 0.1 | 10 | 200 | 400 | 800 | 2 μm | 50 μm | 400 μm | 25 cm |

Increasing \( \lambda \)

Figure 2.1. The electromagnetic spectrum and interactions with matter.
2.2 X-RAY DIFFRACTION

The interaction of a beam of X-rays with a regularly arranged array of atoms with a repeating unit, such as is present in a crystal lattice can produce constructive or destructive interference between the scattered waves. Only the constructive interference will give rise to 3-dimensional diffraction pattern which is related to the 3-dimensional electron density distribution with the molecule. Each diffracted beam is contributed by all the electrons in the molecule, but in different beams their contribution will vary due to phase considerations. With suitable phase information, electron density map may be constructed from the diffraction pattern and a protein molecule can then be constructed into a map and the location of the atoms are positioned so that they "accomodate" (i.e. they are adjusted so that the calculated diffraction pattern of the model matches that determined experimentally). The coordinates of the atoms in the model are then obtained from the model.

In conventional X-ray diffraction the crystal is irradiated with a monochromatic beam of X-rays (Fig 2.2). Planes of atoms in the crystal with spacing d produce diffraction spots at an angle $\theta$ for an X-ray wavelength $\lambda$, where the position of the spots is governed by Bragg's Law:

$$2d \sin \theta = n\lambda$$

Equ. 2.1

where n is the order of diffraction ($n = 1, 2, 3, \ldots$)

The crystal must be rotated to put different planes in the right position to diffract. Obtaining the coordinates of atoms from the diffraction only represents the average positions (i.e. not fixed positions).
Figure 2.2. Conventional X-ray diffraction. The Laue method uses many wavelengths at once.
It is feasible to obtain information regarding the motion of each atom within the crystal and mobile segments can be located on this criteria. It is sometimes not unusual for atoms to be highly mobile within a large volume space that their electron density is attenuated by averaging, which can portray the atoms as "invisible". This can occur where a particular segment of the protein can sometimes adopt many conformations and many of these do exist in a crystal and can give rise to this averaging phenomenon which occurs in space as well as over time. This situation can lead to a scenario where polypeptides with differing sequences can co-crytallize if the differences do not interfere with the main structure and are not involved in intermolecular contacts. Successful crystallization is not necessarily an absolute guarantee of total homogeneity.

The structural information as gained by X-ray crystallography can therefore be seen to represent an "average structure" averaged over time and conformational flexibility of the protein in an aqueous environment. The crystal structure of ribonuclease A and its thermal stability at nine temperatures between 98° and 320°K at high resolution has revealed the dynamics and repacking of the molecule during the transition (Tiltin et al., 1992). Most of the protein structures determined by crystallography have been carried out by this routine procedure until two key developments. One was the introduction of area detectors which allows the data gathering from many reflections simultaneously (as opposed to sequentially), reducing the time required for data collection for a complete data set for several intensity measurements. The second and most important, was the availability of high intensity X-ray sources in the form of Brehmstrahlung from synchrotons (Helliwell, 1991) can
reduce the data collection time to hours. The potential application of polychromatic nature of X-rays from synchrotrons and using Laue diffraction for structural determination of proteins have already been demonstrated (Moffat et al., 1984; Moffat, 1989). The disadvantage of conventional X-ray diffraction using monochromatic X-rays, is the limited number of reflections visible in a particular crystal/film orientation to those that meet the Bragg criteria. In Laue diffraction, the crystal is illuminated with a polychromatic beam of X-rays where all (or nearly all) the planes in the crystal can find a particular wavelength to satisfy Bragg's Law. This means that a single exposure to X-rays produces many reflections simultaneously in a relatively short time, with minimal overlap.

This technique was used to generate a complete data set in a few seconds to study the Ha-Ras-21 protein (molecular weight 21,000) in the process of hydrolysing GTP in an experiment previously impossible to conduct (Schlichting et al., 1990). This study confirms the fact that enzyme activity of the molecule was identical in the crystal and in solution (Schlichting et al., 1989). There are still several problems associated with X-ray crystallography namely some proteins especially membrane proteins and enzymes are not accessible in large quantities and are difficult to crystallize. Once crystallized, the crystal are fragile and susceptible to damage, not only by X-rays but also by fluctuations in temperature and in the type of solvent. More serious still is the size, proteins are large, complex molecules. Data interpretation and structure refinement still requires substantial investment in time and effort. Recombinant DNA technology has recently contributed to the elucidation of
structure for polypeptide-hormone receptor complex, that between hGH and the extracellular domain of its receptor (de Vos et al., 1992). This study showed that the receptor formed a dimer on interacting with the hormone monomer which has two binding sites, each capable of binding to the same site on the receptor. Since the pioneering work of crystal structure of myoglobin and lysozyme, over 600 protein structures have been solved by X-ray methods. Several of these are the structures of the same protein from different species or in an alternative crystal form and the population of unique structures is assumed to number about 140. The progress and recent developments in synchrotron radiation and Laue diffraction, may yet see the most exciting and vibrant innovation in crystal studies.

2.3 ELECTRON DIFFRACTION

Most diffraction studies of membrane proteins have utilized the transmission electron microscope. Electron diffraction is a technique which has been used to the study of two-dimensional arrays of membrane proteins. In order to obtain high resolution, analysis have been carried out on membrane proteins oriented naturally (e.g. bacteriorhodopsin) and proteins incorporated into model membranes (e.g. porin).

Electron radiation possess a typical wave character. The wavelength ($\lambda$) of an electron beam accelerated by an electrical potential ($U$) is given by the De Broglie equation:

$$\lambda = \frac{h}{\sqrt{2meU}} = \frac{12.25}{\sqrt{U}}, \quad \text{Equ. 2.2}$$

where $\lambda$ is the wavelength of the electron beam (in Ångstroms),
h is the Planck constant,
m is the mass of an electron,
e is the charge on an electron,
U is an electrical potential (in volts).

With a voltage $U = 40\text{kV}$, the wavelength $\lambda = 0.06\text{Å}$. This is much shorter than the $1.54 \text{ Å}$ usually obtained in X-ray diffraction studies. The geometry of diffraction of electron radiation is analogous to that for X-ray diffraction described in section 2.2 and for computational purposes can be applied to the Bragg equation (Equ.2.1), and the reciprocal lattice construction. There are several important differences between electron and X-ray diffraction analysis of proteins. Because of the short wavelength of electrons, the electron interference maxima display very small diffraction angles ($\theta$), resulting in numerous reflections obtained by electron diffraction from a single crystal as compared to X-ray pattern. Samples smaller than 50-100Å give rise to line broadening from the diffraction compared to 1000Å for X-ray diffraction. The intense scattering of electrons passing through the sample results in diffraction intensities greater by a factor of $10^6$-$10^8$ than those propagated by X-rays under the same conditions. In order to minimize radiation damage, samples are sometimes exposed to a temperature of around -180°C and negative staining is often applied to preserve the sample integrity during the analysis.

The versatility of modern electron microscopes has enabled the study of some membrane proteins and has led to the successful elucidation of the structure of bacteriorhodopsin (Henderson et al., 1990) and porin (Jap et al., 1991). However it has not been widely applied due to electron radiation
inducing the formation of radicals in samples causing severe crosslinking. Also, radiation is less effective at a higher voltage of acceleration. As a result of extra electron radiation effects, alteration of lattice constants and an increase of the number of lattice imperfections are observed. The resolution obtained by electron microscopy is generally inferior to that obtained by X-ray diffraction and lags in atomic detail.

2.4 NUCLEAR MAGNETIC RESONANCE

NMR has been used to investigate the solution structure of biological macromolecules namely, proteins, polysaccharides and nucleic acids, but such determinations are extremely complex (Wüthrich et al., 1986). Advancement in computers has permitted the generation of dedicated processors that can analyse a vast amount of data generated by the more recently developed techniques of multi dimensional NMR experiments which are essential in the elucidation of structure of polypeptides and proteins. This technique is technically and theoretically complex and has the disadvantage of requiring high concentrations of protein and lengthy data acquisition times. Therefore the protein must be sufficiently pure and available in significant quantity and highly stable at high concentration. The amount required is normally around 0.5 ml of a mM solution of protein and this same amount can be used repeatedly if stable. The nuclei of some atoms (such as $^1H$, $^{13}C$, $^{15}N$ etc.) have a quantum mechanical property known as "spin". A charged particle (i.e. the nucleus) in precession generates a magnetic moment which may be aligned either with or against an external magnetic field. This alignment may be reversed if a
quantum of the appropriate energy is absorbed and conversely if the nucleus relaxes (adopts the lower energy state) a quantum will be emitted. With the advent of powerful computers (the use of superconducting magnets) and sensitive detectors, it is possible to measure this phenomenon. In a one-dimensional experiment, the sample to be studied is excited by a broadband pulse of energy and the quanta emitted (as the nuclei relax) are measured by a receiver coil. Applying Fourier transformation on the signal, a spectrum is generated which plots the intensity of the emitted energy as a function of the frequency of the quanta. Information about the environment of the atoms can be obtained, when energy levels of the nucleus result in the shifting of the resonances to the characteristic regions of the spectrum as a result of local structural effects. The application of multidimensional techniques (Clore and Gronenborn 1991), namely the 2-D permits the definition of certain constraints for the ultimate structure that can be obtained. This can be generated from the linking of resonances between nearby atoms, such that when one is excited, the other resonates. In 2-D NMR, the first dimension excites all the resonance at a particular frequency and the second dimension examines the resonance of all the atoms in the structure. A 2-D NMR plot will have the 1-D spectrum as its diagonal and peaks resulting in other regions of the plot are indicative of distances of certain atoms to others. Therefore in determining the structure it is necessary to identify atoms. By examining the characteristic ranges in the chemical shifts it is possible to identify the type of amino acids and chemical groups. The introduction of sequential resonance assignment has made it possible to assign resonances to specific amino acids when the sequence is
known (Wüthrich et al., 1982; Billeter et al., 1982; Wagner and Wüthrich 1982). The examination of the $\alpha$-CH proton resonance assignments has led recently to the resolution of identity, and location of secondary structural domains of proteins rapidly (Wishart et al., 1992). Also incorporation of $^{13}$C amino acids in proteins can enhance resonance of those amino acids that are not easily resolved ($^{12}$C has no spin and is "invisible"). The presence of 10-20% of a particular amino acid labelled with $^{13}$C is sufficient to dominate the resonance at the expense of the other amino acid resonances (Markley, 1991).

Having obtained geometrical information from 1-D, 2-D and 3-D experiments and resonance assignments such as, distances and dihedral angular constraints relating to known atoms of the sequence, a structure can be calculated which accommodates all the constraints. In practice there will be many different structures that will meet the same set of constraints, since it is impossible to compute all the appropriate distances or angles experimentally. Normally a set of several structures will be computed and subjected to a process of energy minimization or simulated annealing applied to produce an "adjusted" set of structures, from which a mean structure can be obtained (Wüthrich, 1986). On the whole NMR can provide information about conformational flexibility from the range of radio frequencies used. As this will overlap with the frequencies of the "mobile" proteins and can discriminate between different conformations co-existing in solution as long as the different "conformers" have lifetimes that are long compared to these frequencies. This is an important characteristic of this technique compared to other techniques (such as X-ray crystallography) which yields information on "average"
conformational properties. In a recent study of glutaredoxin (Mr = 10,000), starting data of 804 distance and 74 dihedral angle constraints were employed to the determination and refinement of structure and information from 34 slowly exchanging amide protons in the $^{15}$N-H 2-D spectra to map the hydrogen bonding network (Xia et al., 1992). This investigation revealed no differences in the overall structure and hydrogen bonding patterns for the reduced and oxidised glutaredoxin, however there were differences in local dynamics involving the active site and C-terminal helix. From this observation it can be seen that this approach has only been successfully applied to small proteins and peptides with MW of 10-15000 currently being the upper limit.

NMR has also provided much detailed information for testing models of the head group conformation of phospholipids in membranes. Many pioneering studies of the bilayer arrangement above and below phase transitions using both $^{31}$P and $^2$H NMR have been studied extensively (Seelig and Gally 1977). In the deuterium investigation, molecules that were selectively deuterated were studied. Valuable information on the configuration of the acyl chains in different bilayer phases were also obtained. The application of this technique to membrane-bound proteins and peptides is very restricting due to lipids restraining the rotational freedom of protein thus causing line broadening. However some researchers have used organic solvents to mimic the membrane environment to study hydrophobic polypeptides in the absence of lipids. Studies using FTIR have cast some doubt on the application of these systems to study membrane-bound proteins (Kennedy et al., 1991). With the advent of powerful computers and superconducting magnets the upper limit of the
2.5 CIRCULAR DICHROISM

Circular dichroism is a technique often used to assess secondary structure in proteins and peptides (Strickland, 1974). The formation of regular and constrained structures such as $\alpha$-helix and $\beta$-sheets etc. as a consequence of the arrangement of the peptide bonds formed between L-amino acids can result in the alteration in their side-chain absorption spectra due to asymmetry producing what is known as dichroism. Equal absorption of right- and left-handed circularly polarized light does not occur, as they would in a symmetrical non-chiral (or racemic environment). The shape and intensity of this circular dichroism signal can be measured as a function of wavelength (in the 190 to 250nm range) and can be related to the secondary structure (Chen et al., 1974; Johnson, 1988). In the classic approach to the application of estimating secondary structure content from CD spectra, spectra from model polypeptides were used. However doubts in the validity of using homopeptide models has led to an alternative approach whereby known secondary structures of proteins (measured from X-ray crystallography), have been applied for determining the spectral contributions of secondary elements in representative globular proteins (Provencher and Gloeckner 1981). Recently using NMR data for $\beta$-turns, the
component curves for this conformation was improved for CD (Perczel and Fasman 1992). CD does not specifically identify which regions are in which conformation but does provide an estimate of overall time-averaged content of \( \alpha \)-helix. This type of information gathering is useful for time course experiments where certain parameters such as pH etc. can influence the overall secondary structure stability in an empirical manner (Otto and Seckler 1991). It is often used for studying the formations of secondary structural elements in folding experiments (Chen et al., 1992). In some denaturation studies, it is usual to use guanidium hydrochloride (Tanford, 1970), thiocynate and SDS, but there are some proteins that can resist such denaturant. One such protein is lipophilin, a proteolipid, which was incubated in 6M guanidium hydrochloride with 0.1M mercaptoethanol at 70°C for 6 hours, after which it was shown to have retained 85% of its initial helical content. Also the use of 6M guanidium chloride or thiocynate was shown to denature incompletely as determined by fluorescence spectroscopy.

In the 240-320nm range the aromatic amino acids (tryptophan, tyrosine and phenylalanine) can also produce spectra whereby these can function as "markers" in an empirical fashion, essentially sensing changes in the tertiary arrangement within the environment. A change in the structure of the molecule may alter the environment of these "marker" groups and shift the CD spectrum which can be used as "sensors" to monitor such changes. As the outcome of CD signals can be positive or negative, it is normally difficult to deduce meaningful data from these spectra, however in some instances individual bands have been resolved and assigned to transitions associated with
respective side-chains (Holladay et al., 1974). Strong signals are associated with a high degree of constraint on the mobility of the side-chains, since free amino acids (with freely rotating side-chains on proteins) produce weak signals. In their excited states tyrosine and tryptophan display asymmetry, this has been resolved by examining the circular polarization of the fluorescence, from these residues. In one study it was demonstrated that the asymmetry of the environments of the luminescent residues in subtilisin was affected by variation in pH in the range from 5-8, consistent with conformational changes promoting some of the changes in enzyme activity in this range (Schlessinger et al., 1975)). Also disulphide bonds function as chromophores and can produce CD bands in this region in contrast to free thiol groups (Strickland, 1974;).

In the interpretation of CD data, great care must be taken and should be complemented by other techniques. In denaturing experiments, the helical content of the molecule can be estimated with increasing the concentration of the denaturant. In this type of study, the amount of native and denatured forms of the molecule at each denaturant concentration can be determined with the free energy of folding estimated (Lim et al., 1992). In this instance the change in CD signal will indicate the extent of denaturation. However, a significant conformational alteration, such as the shifting of an intact α-helix or even the whole domain within the molecule, may not affect the measured overall helix or β-content and could go undetected in a CD experiment. It has been shown, that the binding of an epsilon-aminocaproic acid to plasminogen has no effect on the overall secondary structure of the molecule’s domains, but causes significant conformational change in which the domains untangle with
respect to each other with a 50% increase in its radius of gyration (Mangel et al., 1990) with corresponding major change in its sedimentation properties on ultracentrifugation (Brockway and Castellino 1972). This is clear indication that a change in secondary structure can be detected by a change in the far UV CD, with important major tertiary or quaternary structural changes going undetected. Also for near UV CD, minor differences in tertiary structure can significantly alter the environment of a chromophore and indicate a change in intensity of its CD spectrum significantly with no major change in its secondary structure (Epand et al., 1977). The other major drawback is that CD spectroscopy is inefficient in distinguishing β-sheet from turn structures, and often the spectra contain overlapping contributions from unassigned amino acid side-chains. Also the undesirable effect of light scattering has greatly hampered its application to the study of membrane suspensions. CD spectroscopy is a "helix" active technique and is a useful tool for confirming helical content in proteins.

2.6 RAMAN SPECTROSCOPY

Raman spectroscopy deals with the phenomenon of a change in frequency when light is scattered by molecules. Basically a monochromatic beam of light is incident upon a sample, and the scattered light is captured by a spectrometer and the intensity as a function of its frequency is measured, keeping the frequency of the incident light fixed. The Raman effect arises from an exchange of energy between the scattered molecule and a photon of the incident radiation. This effect depends on the degree to which a molecule can
be polarized. In this process a transition of the molecule from one of its energy states to another state occurs, and a compensating change in the energy follows. The fundamental equation is

$$h v_0 + E_1 = h v_r + E_2$$  \hspace{1cm} \text{Equ. 2.3}

where

- $h$ is the Planck’s constant
- $v_0$ is the frequency of the incident light (photon)
- $v_r$ is the frequency of the scattered light (photon)
- $E_1$ and $E_2$ are the initial and the final energies of the molecule.

The frequency shift $v_r - v_0 = \Delta v$ may be either positive or negative in sign. Its magnitude is referred to as the Raman frequency. The set of Raman frequencies of the scattering species constitutes its Raman spectrum. The frequency difference between the incident and scattered light is of primary importance, since this corresponds to vibrational or rotational frequencies of the molecules on the sample. The quantities measured in an experiment are the frequencies, widths and shapes of the lines or bands, and also the intensities at particular frequencies or the ratio of the intensities at two frequencies. In Raman spectroscopy the photon is, on the whole, never observed, but it perturbs the molecules and induces them to undergo a vibrational or rotational transition. The internal energy level ($E_n$) of a molecule (Fig. 2.3) must not be confused with the ground state level of a molecule ($E_0$).

When $E_n = E_m$ and hence $V_{nm} = 0$, the state of the scattering molecule remains unchanged and Rayleigh scattering is observed.

When $E_n \neq E_m$ we have the Raman effect. The necessary condition for
Raman scattering is that the energy $h\nu_0$ of the incident photon must be greater than the energy difference $E_m - E_n$ between the final and the initial states of the actual transition.

Raman line of frequency $V_r < V_0$ is called a Stokes-Raman line and frequency $V_r > V_0$ is called an anti-Stokes-Raman line (Fig 2.3). The Stokes line correspond to transitions in which the molecule is raised from a lower to a higher energy state and the photon loses its energy. The anti-Stokes lines correspond to transitions in which the molecule drops from an excited state to a lower energy level and the photon increases its energy. The Raman frequency shift is expressed in terms of wavenumbers instead of frequencies.

$$V = \frac{V_0 - V_r}{c} = \frac{E_m - E_n}{hc},$$

Equ. 2.4

where $c$ is the velocity of light.

Following the introduction of lasers as an excitation source (Lord and Yu 1970; Lord and Yu 1971) the application of Raman spectroscopy to biological macromolecules has become widespread, and the quality of obtainable Raman spectra has improved significantly. An experimental limitation of Raman spectroscopy is the fact that most of the input energy of incident light is expressed as Rayleigh scattering, while only a fraction of it appears as Raman scattering.
Figure 2.3 Schematic diagram for the most important processes involved in light scattering by molecules:

(a) Rayleigh scattering

(b) Production of the Stokes-Raman line \( (h\nu_r < h\nu_o) \)

(c) Production of the anti-Stokes-Raman line \( (h\nu_r > h\nu_o) \)
In addition, biological molecules are often fluorescent, thus masking the weak Raman lines. Fluorescence can arise from intrinsic fluorescent moieties such as tryptophan residues or extrinsic fluorescent impurities. A number of techniques, including prolonged illumination, proper selection of excitation wavelength, signal averaging, use of pulsed lasers and coherent anti-stokes Raman scattering (CARS) have been developed to reduce the problem arising from when fluorescence overshadows the weak Raman lines. A vast amount of work have been carried out using Raman spectroscopy on lipids, soluble proteins and membrane proteins (Lippert et al., 1981). Raman spectroscopy is not affected by the presence of water and where chemical interactions occur, the presence of new molecular species can be detected by the appearance of the new Raman lines. Each scattering biological macromolecule gives its own characteristic vibrational Raman spectrum which can be used for quantitative identification. The intensity of a characteristic Raman line is proportional to the volume concentration of the species, which allows quantitative analysis. It is difficult to validate conclusively about quantitative analysis regarding protein secondary structure by this technique. The introduction of Fourier transform near-infrared Raman spectroscopy seems to have solved the problem of background luminescence of chemical samples, however for biological samples good quality spectra have not improved compared to conventional Raman spectroscopy. Also the generation of intense radiation by the laser accompanied by high temperatures can contribute to protein denaturation during the process.
2.7 INFRARED SPECTROSCOPY

Infrared and Raman spectroscopy are both techniques for investigating the vibrational spectra of molecules. However, these two spectroscopic methods are fundamentally different processes. There are some investigations for which each technique is uniquely applied, in other cases infrared and Raman spectroscopy are nicely complementary. The stretching and bending of chemical bonds in a molecule causes rapid motions of atoms and these properties can be exploited to gain information about the effects of the environments of the atoms. Vibrations that modulate bond polarizability are ideally Raman active and those that lead to changes in the dipole moment of a molecule is infrared active. Infrared spectroscopy is frequently used to monitor the transitions involved in the infrared absorption, associated with vibrational changes within the molecule. It is often used to confirm the identity of small molecules because the presence of certain chemical bonds or groups give rise to IR bands of characteristic frequencies. Different bands present in a molecule (C-C, C=O, C-O, C=O, O-H, N-H etc.) have different vibrational frequencies. The presence of these bonds in a molecule can be detected by identifying characteristic frequencies as absorption bands in the infrared spectrum.

Chemical bonds can undergo various types of stretching and bending vibrations (Fig 2.4). In their vibrations, chemical bonds behave as if connected by a stiff spring.
Figure 2.4. Vibrational modes of CH$_2$ groups of lipid acyl chains. (a) Stretching vibrations; (b) ⊙ and ⊗ indicate movement perpendicular to the plane of the page.
The frequency of a given stretching vibration and thus its location in an infrared spectrum can be related to two factors. These are ion mass and bond strength, that is, light atoms vibrate at higher frequencies than heavier atoms and the relative stiffness of the bond. It can be shown that a non-linear molecule of n atoms has 3n vibrational degrees of freedom: three of the degrees of freedom describe rotation and three describe translation, and the remaining 3n-6 degrees of freedom are vibrational degrees of freedom of fundamental vibrations. Not all molecular vibrations promote absorption of energy however. For a vibration to be infrared active, the dipole moment of the molecule must change as the vibration occur.

Infrared spectroscopy has successfully been applied to the study of biological molecules in aqueous environments. The secondary structure of peptides, proteins, nucleic acids and also membrane proteins within their native lipid bilayer matrix has been studied due to its ease in sample handling. A major drawback of IR spectroscopy used to be the intense absorption of the H-O-H vibration of water in the region of the amide I and II bands (1700-1500 cm\(^{-1}\)) of the protein, thus masking the protein signals (Fig 2.5). One approach to this problem was examining the samples in deuterated buffers. A significant change in the spectra of a protein in \(^2\)H\(_2\)O compared to in H\(_2\)O is that the \(^2\)H-O-\(^3\)H bending vibration absorption band shifts approximately 100 cm\(^{-1}\) away from the protein amide I band. Also reduction in intensity that occurs with the amide II band due to isotopic substitution of the peptide bond N-H to N-\(^2\)H shifts the amide II band to lower frequency by 100 cm\(^{-1}\).
Figure 2.5. FTIR transmittance spectrum of H$_2$O (continuous line) and of $^2$H$_2$O (dashed line) recorded in calcium fluoride cell fitted with a 6µm tin spacer.
It is now possible to study samples in H$_2$O buffers in IR spectroscopy after the introduction of computers and the development of fast Fourier transform algorithms. This permits enhanced sensitivity with digital removal of interfering water absorption accompanied by signal-to-noise ratios.

2.7.1 Instrumentation For FTIR Spectroscopy

The construction and operation of FTIR spectrometers are distinctly different compared to the conventional dispersive type instruments. In the dispersive IR instruments, a collimated beam of infrared light is dispersed by a grating or prism onto a monochromator slit which then transmits a narrow range of frequencies to the detector. By varying the angle of the grating with the incoming beam, a whole spectrum can be scanned, allowing only to resolve one spectral element at any one time.

FTIR spectroscopy on the other hand is a technique which utilizes an interferometer (instead of a monochromator), originally designed by Michelson in 1891 (Fig. 2.6). The Michelson interferometer consist of two plane mirrors at right angles to each other, and a semi-reflecting beamsplitter at an angle of 45° to the mirrors. One mirror is fixed in a stationary position and the other can be moved in a direction perpendicular to its front surface at a constant velocity. Between the fixed mirror and the movable mirror is a beamsplitter, which divides the incoming light from the source, that is 50% is transmitted and 50% is reflected.
Figure 2.6. Schematic representation of the Michelson interferometer optical system used in an FTIR spectrometer
A typical beam splitter might be a thin film of coating deposited on an optically flat support material, the choice depending on the frequencies of light entering the instrument. If the incoming radiation is monochromatic, the signal (or interferogram) at the detector goes through a series of maxima (the two light beams will be in phase or constructively interfere when they recombine at the beam splitter) and minima (the two light beams will be out of phase or destructively interfere when they recombine at the beam splitter). If the mirror is continuously moved the signal will oscillate from maximum to minimum for each quarter-wavelength movement of the mirror (Fig. 2.7a).

In the case of polychromatic radiation, the signal or interferogram at the detector is composed of the resultant signal for each frequency present in the incoming radiation. Each input frequency can be treated independently and hence the output will be the sum of all cosine oscillations caused by all the optical frequencies in the incoming polychromatic radiation (Fig. 2.7b). The interferogram also contains information on the intensity of each frequency in the spectrum. The output information from the detector is digitalised in a computer and transformed to the frequency domain where each individual frequency is filtered out from the complex interferogram (Fourier transformations). The signals are then converted into a conventional infrared spectrum. To comprehend the mathematics of the Michelson interferometer, consider an idealized situation where a source of monochromatic radiation produces a collimated beam.
Figure 2.7. Output of the Michelson interferometer as a function of mirror displacement ($x$): (a) a monochromatic source; (b) a polychromatic source. Zero represents the mirror position for equal optical length in both arms of the interferometer.
The intensity of the IR beam at the detector measured as a function of retardation is represented as:

\[ l'(\delta) = 0.5l(v) \left\{ 1 + \cos \frac{2\pi \delta}{\lambda} \right\} \quad \text{Equ. 2.5} \]

\[ = 0.5l(v) \left\{ 1 + \cos 2\pi v \delta \right\} \]

where \( \delta \) = optical retardation

\( v = \) frequency in wavenumbers \( \text{cm}^{-1} \)

\( l(v) = \) intensity of radiation source

It can be seen that \( l'(\delta) \) incorporates a constant component equal to \( 0.5l(v) \) and a modulated component equal to \( 0.5l(v) \cos 2\pi v \delta \). In spectrometric measurements, only the modulated component is important and this is labelled as interferogram \( I(\delta) \). In an ideal interferometer, the interferogram from a monochromatic source is given by the equation:

\[ I(\delta) = 0.5l(v) \cos 2\pi v \delta \quad \text{Equ. 2.6} \]

In general, the amplitude of the interferogram generated after detection and amplification is proportional to the intensity of the radiation source, beamsplitter efficiency, detector response, and amplifier characteristics. While all the factors remain constant, only \( l(v) \) is variable in any measurement for a given system. Therefore Equation 2.5 may be modified by incorporating a single wavenumber-dependent correction factor, \( H(v) \) to yield the equation:

\[ I(\delta) = 0.5H(v)l(v) \cos 2\pi v \delta \quad \text{Equ. 2.7} \]

where \( 0.5H(v)l(v) \) may be set equal to \( B(v) \), the single-beam spectral intensity. Therefore a basic equation to represent the interferogram can be denoted as:

\[ I(\delta) = B(v) \cos 2\pi v \delta \quad \text{Equ. 2.8} \]

where \( I(\delta) \) is mathematically defined as the cosine Fourier transform of \( B(v) \).
Cosine Fourier transform of \( I(\delta) \) can generate an interferogram and the spectrum can be calculated. Alternatively by observing the interferogram of a monochromatic source, direct measurement can be obtained for the frequency of the incident radiation, which is equal to the frequency of the sinusoidal output of the detector. However infrared sources emit light over a broad range of frequencies, where each frequency generates a unique cosine signal. The sum of each cosine wave generated by each individual frequency component of the input light, generates an interferogram. Applying Fourier transformation, the frequency and intensity of each cosine wave in the interferogram can be resolved. Digital computers use this algorithm to transform the measured-intensity-versus-mirror-displacement signal (the interferogram) into a plot of intensity versus frequency (a spectrum).

### 2.7.2 Resolution Enhancement Techniques

Spectral absorption bands are complex and often incorporate two or more bands overlapping each other such that the instrument is unable to separate or resolve them. The resultant spectrum of a broad band envelope will encompass the real component bands. By applying mathematical techniques it is possible to artificially narrow the widths of the component bands allowing the precise measurement of the component band frequencies (Kauppinen et al., 1981; Took, 1984). The techniques most frequently applied are the derivative and deconvolution in FTIR spectroscopic studies involving biological molecules.
2.7.2.1 Second Derivative

The application of this technique relies upon the fact that a point of inflection on the band envelope occurs at the position of a component band. This can be represented as the rate of change of the gradient, \( \frac{d^2I}{dv^2} = 0 \). The intrinsic shape of a single infrared absorption may be approximated by a Lorentzian function (Knappinen et al., 1981; Susi and Byler 1983)

\[
I = \frac{2s}{n} \left( \frac{s^2 + v^2}{s^2} \right)
\]

Equ. 2.9

where \( I \) = intensity of absorption

\( 2s \) = width at half height

\( v \) = frequency of the band

The second derivative of this function produces:

\[
\frac{d^2I}{dv^2} = - \left( \frac{1}{\pi s} \right) \left\{ \frac{2a \left( 1-3av^2 \right)}{(1 + av^2)^3} \right\}
\]

Equ. 2.10

where \( a = \frac{1}{s^2} \)

The half width of the second derivative, \( s'' \), is related to the half width of the original spectrum, \( s \) by \( s'' = s/2.7 \). Thus the component bandwidth is narrowed without altering the frequency (Fig. 2.8). Second derivative is a useful technique in isolating components which would otherwise be embodied in complex band envelopes. A high signal-to-noise ratio (s/n) is an important criteria in absorption spectra when the second derivative technique is applied as the level of noise may also be enhanced in the resultant spectra. Where poor signal to noise ratio is evident, application of the technique often gives rise to band artefacts in the resolution enhanced spectra.
Figure 2.8 Illustration of the process of second derivative of a Lorentzian band. (a) Synthetic Lorentzian with $\gamma = 5\text{cm}^{-1}$; (b) Fourier transform of (a); (c) Parabolic weighting function; (d) Product of (b) and (c); (e) Fourier transform of (d), yielding the second derivative of (a).
2.7.2.2 Fourier Deconvolution

The deconvolution method for resolution enhancement involves transferring the data in the Fourier domain. In the Fourier domain, the rate of decay of the cosine train is determined by the width of the original band, where wide bands impose faster decay than narrow bands. In a finite decaying exponential cosine train, by multiplying with a suitable finite rising exponential function (Fig. 2.9a), a finite constant amplitude cosine train is the product. A sine function can be obtained on inverse transformation of the finite constant amplitude cosine train, where a narrow peak is produced at the position of the original band maximum with small side lobes on either side of the central peak. The associated side lobes have a tendency to distort in the identification of a component band in the deconvolved complex band envelopes. An added undesirable quantity such as noise will also be increased exponentially. In order to suppress the interference of side lobes and increased noise in the deconvolved spectra, the multiplication by an apodization function such as a sine or triangular function with a constant amplitude cosine form can be performed. By producing a decaying exponential train whose damping is minimal compared to the original cosine train, the transformation of this function will produce a narrower band than the original with low intensity of side lobe artifacts (Fig. 2.9b). Thus it can be seen that apodization functions as a smoothing function and also as the apodization function approaches zero, the noise is reduced in the cosine train as well as in the deconvolved spectrum.
Figure 2.9 Illustration of the mechanism of deconvolution: (a) multiplication of the decaying cosine train by an increasing exponential function. (b) multiplication of the finite cosine train by an apodization function.
2.7.3 Interferometer Advantages

An infrared spectrometer incorporating an interferometer offers several theoretical advantages over a dispersive infrared measurements as presented below:

1. Fellgett's or Multiplex Advantage

The fact that an interferometer measures all wavelengths at the same time means that a complete infrared spectrum can be measured in a comparatively short time. Therefore each point in the interferogram contains information from each wavelength present in the input light signal. The ability to co-add spectra means that the signal-to-noise improves extremely rapidly:

\[
\text{Signal-to-noise} \propto \sqrt{\text{no. of co-additives}}
\]

and good quality spectra can be produced in relatively short time.

2. Jacquinot or Throughput Advantage

This relates to the fact that the infrared energy reaching the detector is much greater for a FTIR spectrometer (no slits and fewer optical elements) than for a dispersive infrared spectrometer. Hence, much smaller samples can be analysed using FTIR spectroscopy than with dispersive type and better detection limits are provided.

3. Conne's Advantage or Frequency Precision

Modern high speed FTIR spectrometers incorporate its own internal frequency standard, generally a helium-neon laser to facilitate digitization of
data. By contrast with dispersive instruments, frequency precision and accuracy depends on calibration with external standards and on the ability of electromechanical mechanisms to uniformly move gratings and slits during and between scans. Furthermore, all the frequencies in the output spectrum are calculated from the known frequency of the laser light.

4. **Constant Spectral Resolution**

In an FTIR spectrometer, the resolution of the measured spectrum is the same for all frequencies, unlike in dispersive instruments where it is varied throughout the spectrum.

2.7.4 **Sampling Techniques**

Acquiring meaningful infrared data from aqueous solutions has always been difficult. This is due to the contribution of strong absorption by water throughout much of the mid-infrared spectral region. Studies involving aqueous protein solutions can add further problems because of the bending vibration of water molecules which absorbs strongly in the conformation-sensitive amide I region (Table 2.1). Application of resolution enhancement techniques can highlight narrow bands more than broad bands thus giving rise to band artefacts associated with water vapour in the amide I band.

A remedy to this interference is to purge the instrument continuously with dry compressed air, thus eliminating moisture content and by introducing a sample shuttle, the background can be signal averaged concurrently with the sample. Subtraction of recorded moisture absorption spectrum from the protein
absorption spectrum can further suppress the vapour band. Ideally biological material are investigated in aqueous media. The contribution of strong H-O-H bending vibration of water (Fig. 2.5) limits the sample thickness to 6\textmu m which requires the use of 20-50mg/ml of protein to give a reasonable signal-to-noise ratio. Using \textsuperscript{2}H\textsubscript{2}O as a solvent, 50\textmu m pathlength can be used with subsequent lower concentration of sample as the \textsuperscript{2}H\textsubscript{2}O absorption is located some 400cm\textsuperscript{-1} lower (Fig. 2.5).

In the case of thin films, when applied in conjunction with biological membranes, internal reflection spectroscopy (also known as attenuated total reflectance spectroscopy) can be applied (Fig. 2.10b). This technique which is based on internal reflection where the incident radiation is directed first into a material of high refractive index (n\textsubscript{1}) at an angle \theta greater than the critical angle (\theta\text{c}), and after slight penetration beyond the surface is reflected from the sample, which has lower refractive index (n\textsubscript{2}) (Fig. 2.10a). When a sample which selectively absorbs radiation is deposited, in contact with a reflecting surface, some frequencies of the incident beam will be absorbed while other frequencies are transmitted and reflected.
Condition for internal reflection:
\[ n_1 > n_2, \theta > \theta_C \]

Figure 2.10. Internal reflection spectroscopy:
(a) Conditions for single internal reflection
(b) Multiple internal reflections
This attenuated radiation is measured and plotted as a function of wavelength by a spectrometer which produces an absorption spectrum of the sample called an internal reflection spectrum. By placing the sample on a germanium crystal, the incident infrared beam is repeatedly absorbed and reflected thus increasing the signal-to-noise ratio (Fig. 2.10a). The number of reflection depends on the dimensions of the crystal giving approximately 25 reflections, permitting high quality spectrum from a sample size of around 30\(\mu\)g.

This technique can be adopted for molecular orientation studies (polarized internal reflection spectroscopy) in biomembrane films. This allows spectra to be recorded on ordered bilayers and information can be gathered regarding the orientation of different structures of protein or peptides (Goormaghtigh et al., 1989). The mean orientations of the secondary structures can be determined from the orientations of the peptide bond corresponding to the C = O group. Spectra are recorded with parallel (0°) and perpendicular (90°) polarized incident light with respect to the ATR plate. Polarization is expressed as the dichroic ratio \(R_{\text{str}} = \frac{A_{90}^{\circ}}{A_{0}^{\circ}}\). The mean angle between the C = O band and a normal to the ATR plate surface is calculated from \(R_{\text{str}}\) (Cabiaux et al., 1989).

2.7.5 Significant Vibrations In Biomolecules

2.7.5.1 Proteins

Of the conformationally sensitive regions of the IR spectra studied in detail so far, the most prominent are the amide vibrations (Table 2.1). The presence of peptide bonds in the backbone of proteins and peptides can give
rise to these characteristics infrared active vibrations. The vibrational modes of the peptide moieties present in various secondary structures vary according to their spatial orientation in the polypeptide. Peptide moieties containing similar spatial constraints therefore imparts similar vibrations. The most useful of the amide vibrations for protein secondary structural determination are the amide I and amide II modes (Fig. 2.11). The amide I vibrational mode consist principally of 80% C=O stretching with 10% contribution from N-H in-plane bending and 10% C-N stretching (Miyazawa, 1960). The amide II vibrational mode consist of C-N stretching and N-C\(^\#\) stretchings with minor contribution from N-H in-plane bending.

The amide I vibration, which is principally C=O stretching, can vary according to the nature of the hydrogen bonding found in \(\alpha\)-helical and \(\beta\)-sheet structures. As a result, the frequency of the amide I band can be applied to distinguish between the secondary structures which occur in proteins and polypeptides. Application of theoretical and experimental studies with model polypeptides have shown good correlation between the amide I band frequency and their representative secondary structures (Susi and Byler 1986; Krimm and Bandekar 1986). This has been further corroborated with studies on a range of soluble proteins in various laboratories (Susi et al., 1986; Surewicz and Mantsch 1988; Krimm and Bandekar 1986).
Table 2.1. Characteristic IR Absorption Frequencies Of Proteins

<table>
<thead>
<tr>
<th>Amide Vibration</th>
<th>Frequency cm$^{-1}$</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3300</td>
<td>N—H (s)</td>
</tr>
<tr>
<td>B</td>
<td>3100</td>
<td>N—H (s)</td>
</tr>
<tr>
<td>I</td>
<td>1680-1600</td>
<td>C=O (s) 80%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N—H (b) 10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C—N (s) 10%</td>
</tr>
<tr>
<td>II</td>
<td>1580-1480</td>
<td>N—H (b) 60%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C—N (s) 40%</td>
</tr>
<tr>
<td>III</td>
<td>1300-1230</td>
<td>C—N (s) 30%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N—H (b) 30%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C=O (s) 10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O=C—N (b) 10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other 20%</td>
</tr>
<tr>
<td>IV</td>
<td>770-626</td>
<td>O=C—N (b) 40%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other 60%</td>
</tr>
<tr>
<td>V</td>
<td>800-640</td>
<td>N—H (b)</td>
</tr>
<tr>
<td>VI</td>
<td>605-540</td>
<td>C=O (b)</td>
</tr>
<tr>
<td>VII</td>
<td>200</td>
<td>C—N (t)</td>
</tr>
</tbody>
</table>

s = stretch; b = bending; t = torsion
Similarly it has also been successfully applied to determine secondary structure of membrane proteins in aqueous lipid dispersions and binding of various ligands (Jackson et al., 1989). From these studies, bands in the spectral range 1620-1640 cm\(^{-1}\) are attributed to \(\beta\)-sheet accompanied by minor component bands at 1680-1690 cm\(^{-1}\). Membrane associated proteins possessing \(\alpha\)-helical conformation, will in H\(_2\)O or \(^2\)H\(_2\)O normally display IR absorptions in the range of 1656-1658 cm\(^{-1}\), while soluble proteins absorb in the range of 1648-1656 cm\(^{-1}\). This variation in absorbance may be attributed to the protein-water interaction, which is minimal in hydrophobic membrane environment and hence the increase in the amide I maximum. Unordered structures also absorb in the same \(\alpha\)-helical region in H\(_2\)O, but this can be distinguished by a shift in absorption to 1644 cm\(^{-1}\) upon deuteration.

The majority of membrane proteins investigated to date using FTIR have shown a preference for \(\alpha\)-helical structures with very minor \(\beta\)-sheet like components in membrane mimetic environments. However, the matrix porin protein from *Rhodobacter capsulatus* has shown a predominance of \(\beta\)-sheet (Haris and Chapman 1992). The amide II region has always been weak in intensity in the IR spectra of proteins and polypeptides to be of any value for conformational analysis. It can however be applied in the investigation of hydrogen-deuterium exchange of the peptide groups in proteins and polypeptides. This information can indicate the state of solvation of the peptide or protein in that particular structural arrangement.
Figure 2.11. Amide I, II and III vibrational modes of a peptide moiety.
The most noticeable change in the spectra in $^2$H$_2$O compared to H$_2$O is a significant reduction in the intensity in the amide II band. As the amide II band arises from the N — H bending vibration, substitution of the N — H by N — $^2$H shifts the amide II band to the lower frequency. This form of interaction has greatly contributed to the study of deuterium exchange (Baiman and Rothschild (1988); Haris et al., (1986), which can lead to greater understanding to the study of overall conformational dynamics.

2.7.5.2 Lipids

IR spectroscopy have also been applied extensively to the study of lipids (Table 2.2). The acyl chains present in lipids can impart information due to the IR active vibrations of the CH$_2$ and CH$_3$ groups. The interfacial region of the lipids such as the C — O groups, and the headgroups encompassing the PO$_2^-$, CH$_2$ and NH$_3$ can also be investigated. The solvation of lipids in bilayers and liposomes can be investigated by the PO$_2^-$ stretching mode which can be affected by the presence of H$_2$O. Phase transition behaviour of lipids can also be studied by changes in the frequency of the CH$_2$ asymmetric stretching vibration from 2919 cm$^{-1}$ to 2924 cm$^{-1}$ when temperature is elevated above its T$_m$. The C = O groups in the interfacial region of the lipid molecule are also sensitive to these phase transitions. Resolution Enhancement techniques indicate that the band arising from the C = O stretching vibrations comprises two components. It was proposed originally (Bush et al., 1980), that these absorptions could be attributed to the sn-1 (1743 cm$^{-1}$) and sn-2 (1728 cm$^{-1}$) C = O groups.
Table 2.2. Characteristic IR Absorption Frequencies Of Lipids

<table>
<thead>
<tr>
<th>IR Band Frequency (cm⁻¹)</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2956</td>
<td>CH₃ asymmetric stretch</td>
</tr>
<tr>
<td>2920</td>
<td>CH₂ asymmetric stretch</td>
</tr>
<tr>
<td>2870</td>
<td>CH₃ symmetric stretch</td>
</tr>
<tr>
<td>2850</td>
<td>CH₂ symmetric stretch</td>
</tr>
<tr>
<td>1470</td>
<td>CH₂ bending / scissoring</td>
</tr>
<tr>
<td>1380</td>
<td>CH₂ symmetric deformation</td>
</tr>
<tr>
<td>1380-1190</td>
<td>CH₂ wagging band progression</td>
</tr>
<tr>
<td>1150-700</td>
<td>CH₂ rocking band progression</td>
</tr>
<tr>
<td>3050</td>
<td>CH₃ asymmetric stretch of N(CH₃)₃</td>
</tr>
<tr>
<td>1740-1720</td>
<td>C==O ester stretch</td>
</tr>
<tr>
<td>1490</td>
<td>CH₃ asymmetric bend of N(CH₃)₃</td>
</tr>
<tr>
<td>1405</td>
<td>CH₃ symmetric bend of N(CH₃)₃</td>
</tr>
<tr>
<td>1250</td>
<td>O==P==O asymmetric stretch</td>
</tr>
<tr>
<td>1200-1000</td>
<td>C==O stretch</td>
</tr>
<tr>
<td>1170</td>
<td>C==O from esters</td>
</tr>
<tr>
<td>1085</td>
<td>O==P==O symmetric stretch</td>
</tr>
<tr>
<td>1070</td>
<td>C==O from esters</td>
</tr>
<tr>
<td>1040</td>
<td>C==N stretch</td>
</tr>
<tr>
<td>900-800</td>
<td>P==O stretch</td>
</tr>
<tr>
<td>800</td>
<td>C==N stretch</td>
</tr>
</tbody>
</table>
However FTIR studies using $^{18}$O-substituted lipids later revealed that the two bands arise from the $\text{C} = \text{O}$ groups being hydrated at different levels with the lower frequency band attributed to higher level of hydration (Blume et al., 1988). Elevation of the temperature through the $T_m$ also increases the intensity of the 1728 cm$^{-1}$ component thus indicating a higher level of hydration. This produces a decrease in the frequency of the lipid $\text{C} = \text{O}$ absorption. It is apparent that FTIR spectroscopy is a useful tool in reconstitution studies involving membrane associated natural or synthetic proteins because of its excellent resolving power. In the light of the problems inherent to techniques such as NMR, CD and X-ray, FTIR techniques have been receiving a great deal of attention, and FTIR spectroscopy utilizing the technique of band resolution after enhancement developed over the past decade.
CHAPTER THREE

FTIR SPECTROSCOPIC STUDIES OF CHEMICALLY SYNTHESIZED POLYPEPTIDES CORRESPONDING TO BACTERIOPHAGE Pf1 COAT PROTEIN AND ITS EFFICIENT DISAGGREGATION AND RECONSTITUTION INTO MEMBRANE MIMETIC ENVIRONMENTS
3.1 INTRODUCTION

The current knowledge of biological processes involving membrane-bound proteins is not as detailed as for those originating from cytoplasmic proteins. This can be attributed largely to the difficulty in applying techniques capable of yielding conformational information at atomic resolution on membrane-bound proteins (Eisenberg, 1984). Despite the great progress achieved with diffraction (Leifer and Henderson, 1983) and NMR (Keniry et al., 1984) studies evaluating the structure of bacteriorhodopsin in purple membrane, and porin of Escherichia coli outer membrane (Kleffel et al., 1985), only the diffraction data on the photoreaction centre (Deisenhofer et al., 1985) crystallized from detergent (Michel, 1982) have shown resolution normally reserved for soluble proteins (Richardson, 1981). The isolation, purification and characterization of membrane proteins are often detrimental to their structure and biological activity compared to soluble proteins. Proteins and polypeptides originating from lipid membranes are usually highly hydrophobic and their solubilization and purification entails the use of detergents or organic solvents. However the tendency of these integral membrane proteins to form stable aggregates in aqueous solutions can sometimes destroy their native membrane structure. This has made structural analysis by spectroscopic techniques difficult and unreliable. To circumvent these difficulties, chemically synthesized polypeptides can represent these membrane-bound proteins in structural studies. These synthetic polypeptides can also form highly hydrophobic aggregates in aqueous solutions, detergents or organic solvents. However once disaggregated, there are advantages using chemically synthesized polypeptides.
The fact that, the source of the membrane proteins are non-biological, its homogeneity, quantity and manipulative losses kept to the minimum can be assured. The aim of this chapter is to demonstrate the experimental strategy that has proved successful in the chemical synthesis, disaggregation, purification and reconstitution of synthetic polypeptides corresponding to the filamentous bacteriophage Pf1 coat protein into membrane mimetic environments for structural analysis using FTIR spectroscopy.

Filamentous bacteriophage Pf1 are 2μm long tubular viruses containing a single-stranded, circular DNA of 7349 nucleotides wrapped in a flexible sheath of coat protein subunits (Day et al., 1988). The coat protein plays a vital role in the infectious cycle of the phage by associating with its host bacteria pseudomonas aeruginosa strain K. The amino acid sequence of the coat protein is shown in Figure 3.1 and it is amongst the simplest representative of transmembrane proteins. The coat protein molecules from an infecting virus particle and the newly synthesized coat protein molecules, insert into the host cell membrane (Fig. 3.2). During the assembly of new virus particles at the cell membrane, the NH₂ termini is oriented on the outside of the membrane and the COOH-termini on the inside (Nambudripad et al., 1991). At the COOH-terminus, the basic amino acids associate by electrostatic interactions with the DNA (Greenwood and Perham 1989) and the proteins associate with another to form the long (1 to 2μm) flexible, 65Å diameter virions in the absence of lipids (Fig. 3.3).
Figure 3.1. Amino acid sequence of Pf1 coat protein. The underlined 11 residues indicate very slow proton exchange rates.
This mode of processing can lead to greater understanding of membrane-mediated mechanisms, including assembly of membrane proteins and protein transport. X-ray and neutron diffraction studies have indicated that the viruses are cylindrically hollow composed of numerous subunits of 46 amino acid long protein wrapped around the DNA (Marvin and Wachtel 1975). The coat protein subunits are oriented as helical rods, with a minor tilt to the filament axis and slotting between the neighbouring helical rods, forming left-handed spirals that ascend along the axis of the bacteriophage. The coat protein plays an important dual role during its infectious cycle where it becomes an integral protein (Fig. 3.3) of the host (Smilowitz et al., 1972; Webster and Cashman 1973). During the assembly of the phage, it leaves the host membrane to become one of the numerous subunits that make up the new viruses (Marvin and Wachtel 1975).

The intact filamentous Pf1 bacteriophage in a previous study using CD spectroscopy was shown to be composed of almost 100% α-helical structure (Day et al., 1988). Analysis using 2D NMR indicated that the coat protein in DPC micelles comprises two stable α-helical segments (Schiksnis et al., 1987; Shon et al., 1991). The complete structure and orientation of Pf1 coat protein in DPC micelles has been characterized using solid-state NMR spectroscopy where uniformly $^{15}$N-label and site specific $^{15}$N-labelled Pf1 coat protein were analysed. The data obtained led to the proposal of a model for Pf1 coat protein in the phospholipid bilayer where a stretch of a hydrophobic region of about 24 amino acid residues is a membrane-spanning α-helix.
Figure 3.2. Model for the structural changes that occur in the Pf1 coat protein viral assembly. The coat protein is drawn as a ribbon diagram shown relative to the host cell membrane.
Also the amino terminal region of 18 amino acid residues is projected as an amphipathic $\alpha$-helix oriented parallel to the lipid bilayer plane (Fig. 3.4). The amphipathic N-terminal synthetic polypeptide sequence of the Pf1 coat protein is shown in Figure 3.5.

In this study, FTIR spectroscopy has been applied to assess the suitability of using synthetic polypeptides corresponding to the sequence of Pf1 coat protein and the N-terminal region for conformational analysis in lipid vesicles and detergent micelles. The FTIR spectra of the synthetic polypeptide in lipid vesicles and detergent micelles using the conventional and a new reconstitution technique were compared with the natural protein in similar environments. FTIR spectroscopy has been shown to be a useful tool to probe secondary structures of proteins (Susi and Byler, 1983; Haris and Chapman, 1992) and is especially useful in studying hydrophobic transmembrane polypeptides in membrane mimetic environments.

This chapter describes studies of synthetic polypeptides corresponding to the natural Pf1 coat protein using FTIR spectroscopy combined with resolution enhancement techniques. The characterization of the synthetic coat protein is important in order to identify the aggregated and disaggregated structural forms that can coexist when incorporated into membrane mimetic environments. Once this has been established, the use of chemically synthesized proteins to study membrane-bound proteins can be a useful practical approach.

This work was carried out in collaboration with Dr. Parvez Haris (this department).
Figure 3.3. The expanding cylinder of the virus is shown relative to the assembling coat protein to depict a hypothetical view of the membrane-associated viral assembly and protein conformational changes.
Figure 3.4. Model of Pf1 coat protein in lipid bilayers showing the amino terminal amphipathic helix parallel to the plane of the bilayer.
Figure 3.5. The amphipathic synthetic peptide corresponding to N-terminal sequence of Pf1 coat protein

NH₂-GVIDTSAVESAITDGOGD-NH₂
3.2 MATERIALS AND METHODS

Natural Pf1 coat protein in DMPC and SDS micelles were generous gifts from Dr. P.I. Haris (this department). D_{2}O, DMPC, LPC and SDS were purchased from Sigma Chemicals Ltd., (U.K.). Fmoc-protected amino acids, HMPA, TBTU, and MBHA resins were purchased from Calbiochem-Novabiochem (UK)Ltd. ODHBT, DIPCDI and HBT were purchased from Fluka (UK) Ltd.,

Peptide synthesis grade solvents were used throughout the synthesis and were purchased from Ratburn (Scotland). All other reagents were of analytical grade. Solvents for HPLC were all of protein sequencing grade and all other reagents were of analytical grade.

3.2.1 Preparation of MBHA resin incorporating the acid labile HMPA linker to produce the acid resin

0.5g of the MBHA polystyrene beads with a normal loading of 0.2mmol/g were preswollen in the synthesizer reaction vessel and washed (5ml; 5 x 1 min) with DMF. The washed and drained amino resin was acylated with freshly prepared HMPA active ester. A 5-fold excess (based on normal loading) of acylating species with the following equivalent, linker:TBTU: OHBT: MMP (1:1:1:2) in DMF (4ml) were agitated with oxygen free nitrogen for 30 minutes. The resin was thoroughly washed in DMF and residual amino groups on the resin were acetylated using acetic anhydride in DMF. Acetylation was carried out by adding Ac_{2}O to the amino-resin (0.47 ml: 5mmol) in 15 ml of DMF. The acetylation mixture were mixed in the reaction vessel for 30 minutes at room temperature. On completion of acetylation (a negative test is indicated
by disappearance of yellow colour using ODHBt) the polystyrene resins were washed with DMF (5ml; 5x1 minute) and finally washed with DCM (5ml; 5x1 minute) and dried under vacuum for 2 hours.

3.2.2 Esterification of the first Fmoc-amino acid residue using pre-formed symmetrical anhydrides

The HMPA-MBHA resin was washed with DMF (10ml/g; 5x1 minute). A solution of the Fmoc—alanine—OH (0.31g; 1mmol) in DCM (5ml) containing sufficient DMF (drops from a measured 5ml) to ensure complete dissolution was stirred at room temperature and activated by adding DIPCDI (154μl; 1mmol) directly into the stirred solution. Formation of a solution of diisopropylurea (slight turbidity) ensued after 12 minutes which dissolved on addition of the remaining DMF.

The HMPA-MBHA resin was then esterified with Fmoc-alanine anhydride in the presence of catalytic amount of DMAP (30mg; 0.25mmol) and MMP (277μl; 0.5mmol) for 30 minutes. After the mixture had been agitated for 30 minutes the resin was thoroughly washed and drained with DMF ready for automatic synthesis of the complete peptide.


The HMPA-MBHA resin was placed in a sealed reaction vessel and acylated with active esters formed in-situ from pre-weighted protected Fmoc-
amino acids in the presence of the activating agent TBTU. A 2.5 fold excess (based on the loading) of acylating species with the following equivalents amino acid: TBTU : MMP (1:1:2) in DMF were used in all subsequent coupling. The arginines were all double coupled throughout the synthesis. The synthesis was carried out using the batch synthesis mode on an automated peptide synthesizer Rainin PS3 (Protein Technologies, USA) with standard protocols and scale (0.05 mmol theoretical yield of peptide). The cycle for the addition of activated amino acid consisted of a 7 minute wash of the solid support with 20% piperidine in DMF to cleave the N-Fmoc group, a 12 minute DMF wash, a 30 minute acylation reaction with 5 equivalents of an Fmoc-amino acid TBTU ester, and an 8 minute DMF wash for a total cycle time of about 60 minutes.


0.5g of the Rink-amide resin derivatized with the Fmoc-amide was placed into a glass reaction vessel and a normal deprotection cycle was started with 20% piperidine in DMF for 12 minutes. Acylation was initiated using pre-weighted Fmoc-amino acids in the presence of TBTU in sample 5ml bottles. A 2.5-fold excess (based on the resin loading) of the acylating species in 0.2M MMP in DMF (5 ml) were added automatically via an injection port to the reaction vessel. The synthesis was carried out in an automated peptide synthesizer with standard protocols and scale (0.1 mmol theoretical yield of crude peptide). The cycle for the addition of the activated protected amino acid consisted of the same procedure as the synthesis of the complete Pf1 coat.
protein.

3.2.5 Detachment and deprotection of the synthetic Pf1 coat protein and its 18-residue N-terminal polypeptide

After completion of Pf1 polypeptide assembly and removal of the N-terminal Fmoc group, the HMPA-support was washed with DCM and dried in vacuo. The polypeptide-resin was treated with 95% aqueous TFA (30-50ml) containing 5% scavengers (EDT, phenol and thioanisole) for 3 hours in a sintered glass column agitated with oxygen free nitrogen. The resin was filtered into a round-bottomed flask and washed 3 times with 95% aqueous TFA. The TFA and scavengers were removed by rotary evaporation and the crude product was precipitated with cold anhydrous diethyl ether to remove residual scavenger(s). This was repeated 5 times and the peptide was dried in vacuo.

For the deprotection and detachment of the N-terminal peptide, the Rink-MBHA resin was washed with DCM and dried in vacuo. To the dried resin 10% TFA in DCM was added and agitated with oxygen free nitrogen in a sintered glass column for 60 minutes, collecting the filtrate every 20 minutes into a round-bottomed flask. The TFA and DCM were removed by rotary evaporation until an oily mixture remained. To this oily mixture 95% TFA in H₂O (50ml) containing 5% EDT, phenol, and thioanisole were added and incubated at room temperature for 5 hours. The TFA and scavengers were removed by rotary evaporation and cold diethyl ether was added to precipitate the polypeptide and remove any residual scavengers. The precipitated polypeptide was further
rinsed with diethyl ether and centrifuged to obtain a pellet which was dried in vacuo.

3.2.6 Purification of synthetic Pf1 coat protein and its 18-residue N-terminal polypeptide

The complete Pf1 synthetic coat protein was found to be insoluble in aqueous buffers and so entailed the use of TFA/TFE (1:1, v/v) for the purification scheme. The crude Pf1 coat protein was subjected to multiple reverse-phase HPLC (Varian, USA) runs. The synthetic Pf1 protein was injected onto a Vydac C8 (208TP 1010 RP preparative column) equilibrated in 50% buffer A (TFA/ACN/H2O 1:1:98, v/v) and buffer B (TFA/ACN/H2O 1:90:9, v/v). The polypeptide was purified on a linear gradient of buffer B (from 50 to 70% in 30 minutes) at a flow rate of 4 ml/min. One major peak was observed when the eluate was monitored at 220nm and was isolated and subjected to amino acid analysis to confirm its identity. The purified synthetic polypeptide was further characterized for homogeneity by a second HPLC analysis on a narrow-bore Vydac C8 (208 TP 54 RP, analytical column) equilibrated as described above. The crude N-terminal 18-residue polypeptide was also purified by multiple reverse-phase HPLC runs. The polypeptide was dissolved in 0.1% TFA in water and injected onto a Vydac C18 (218TP1010 RP, preparative column) equilibrated with 100% buffer A (deionized/millipore water containing 0.1% TFA) and 0% buffer B [90% (vol/vol acetonitrile in water containing 0.1% TFA). The peptide was purified on a linear gradient of buffer B (0 to 70% in 30 minutes) at a flow rate of 4ml/minutes. The homogeneity was verified by a second HPLC
analysis on a narrow-bore Vydac C18 (201 TP54 RP analytical column) equilibrated as described. The polypeptide was lyophilized and stored at +4°C until further use.

3.2.7 Amino acid compositional and molecular mass analysis

HPLC purified Pf1 synthetic protein and the N-terminal peptide (∼5nmol) were subjected to hydrolysis with 6M constant boiling HCl containing 0.1% phenol at 110°C for 20-70 hours in oxygen free atmosphere in sealed glass ampoules. On completion of hydrolysis residual HCl was removed using a freeze-drier, the residue resuspended in amino acid sample buffer and applied to the analyser (LKB Biochrom 4151 Alpha Plus). Samples in duplicates were applied onto the amino acid analyser and calibrated using 10 nmol standards. The identity and purity of the synthetic Pf1 coat protein and the N-terminal peptide were further characterized using the laser desorption mass spectrometry (LASERMAT Mass Spectrometer).

3.2.8 Reconstitution of synthetic Pf1 coat protein and its N-terminal polypeptide in lipid vesicles and detergent micelles

The usual conventional reconstitution technique widely used for FTIR spectroscopy was applied, where polypeptides and DMPC or SDS were co-dissolved in chloroform and evaporated to dryness under a stream of nitrogen and further dried in vacuo for 24 hours. This procedure was termed the "before disaggregation" for spectral analysis. In the new disaggregation procedure, the polypeptides were dissolved in neat TFA first and its complete removal by
rotary evaporation under high vacuum and further dried for 24 hours using a freeze-drier to obtain a thin film of the polypeptide. To eliminate residual TFA associated with the polypeptides, it was lyophilised in the presence 0.1M HCl for 24 hours. To this dry film, DMPC or SDS in DCM was added and rotary evaporated at high vacuum to dryness and further dried under vacuum for 24 hours. Finally $^2\text{H}_2\text{O}$ PBS (pD 7.4) was added to this dry mixture and vortexed vigorously for 5 minutes at room temperature. The synthetic Pf1 coat protein to DMPC and SDS molar ratio used was 1:20 and 1:175 respectively. For the N-terminal Pf1 peptide, the peptide to DMPC ratio was 1:20 and in aqueous solution it was 10mg/ml. The polypeptides were allowed to equilibrate for at least 1 hour prior to spectral measurement.

### 3.2.9 FTIR spectroscopy

FTIR spectra were obtained using a Perkin-Elmer 1750 FTIR spectrometer equipped with a fast recovery TGS detector and Perkin-Elmer 7300 computer for data acquisition and analysis. Samples were placed in a Beckman FH-01 CFT micro-cell fitted with CaF$_2$ windows and a 50µm Teflon spacer for measurements in $^2\text{H}_2\text{O}$. For measurements in H$_2$O a 6µm tin spacer and a protein concentration of 50mg/ml were used. Temperature control was achieved by means of a cell jacket of circulating water. The spectrometer was continuously purged with dry air to eliminate water vapour absorptions from the spectral region of interest. A sample shuttle was employed to permit the sample to be signal-averaged with the background. Measurements in $^2\text{H}_2\text{O}$ at 30°C were recorded at a polypeptide concentration of 10mg/ml and 200 scans.
were signal averaged.

All spectra were recorded at a resolution of 4 cm⁻¹. Aqueous buffer spectra were recorded under identical conditions as the sample spectra. Absorption spectra were obtained by digitally subtracting the solvent spectrum from the corresponding sample spectrum. Where appropriate, water vapour contributions were subtracted from the absorption spectrum using a previously recorded water vapour spectrum. Second derivative (Moffat et al., 1986) were calculated over a 13 data-point range (13 cm⁻¹) using the Perkin-Elmer DERIV function to assign features of the composite amide I band to structural features present in the polypeptides.

3.2.10 ATR-FTIR spectroscopy

ATR-FTIR spectra were acquired with a Mattson FTIR spectrometer series 4060, equipped with a vertical ATR accessory located between the sample and the detector. Oriented films of the lipid/polypeptide mixtures were spread on the surface of a germanium ATR crystal (angle of incidence = 45°) as follows: Lipid vesicles dispersed in ²H₂O buffer were spread on the surface of the germanium crystal and oriented by gently stroking with a teflon rod as the ²H₂O slowly evaporated under a gentle stream of nitrogen. The dichroic character of the infrared bands of the terminal methyl chains of the lipid acyl chains at 2874 and 1200 cm⁻¹ served as a reference to indicate the degree of perpendicular orientation of the lipid molecules in the film (Fringeli and Gunthard, 1981). Similarly films of the lipid/polypeptide mixtures in ²H₂O buffer after disaggregation were spread on the surface of the germanium crystal and
oriented with stroking by a teflon rod as the solvent evaporated under a gentle stream of nitrogen. The whole set up was continuously purged with dry air. Spectra of the germanium crystal with and without lipid film were recorded separately as single-beam spectra, ratioed to give transmission spectra, and finally plotted as absorbance = 2\log(R_f/R_i), where $R_f$ and $R_i$ represent the final and initial reflectance, respectively.

### 3.3 RESULTS

#### 3.3.1 Chemical and physical analysis of the synthetic Pf1 coat protein and its N-terminal polypeptide

The synthetic Pf1 coat protein and its N-terminal polypeptide were subjected to multiple purification procedures and analysis. The preparative scale reverse-phase HPLC of 20mg of crude synthetic Pf1 coat protein and its N-terminal peptide are shown in Figures 3.6 and 3.7. The confirmation of the homogeneity of these polypeptides using analytical reverse-phase HPLC are shown in Figures 3.8 and 3.9. The polypeptides eluted as well resolved major peaks in the HPLC runs. After separation by reverse phase HPLC, the polypeptides were identified by their amino acid compositions (Tables 3.1 and 3.2). All the respective peaks could be accounted for as the desired synthesized products.
Figure 3.6. Preparative scale reverse-phase HPLC chromatogram of crude synthetic Pf1 coat protein. The column was eluted isocratically for 3 minutes with 50% B and then with a linear gradient of 50-70% B developed over 30 minutes.
Figure 3.7. Preparative scale reverse-phase HPLC chromatogram of crude synthetic peptide corresponding to the N-terminus sequence of Pf1 coat protein
Figure 3.8. Analytical reverse-phase HPLC of synthetic Pf1 coat protein after purification.
Figure 3.9 Analytical reverse-phase HPLC of synthetic peptide corresponding to the N-terminus sequence of Pf1 coat protein after purification.
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Actual</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>7.10</td>
<td>7</td>
</tr>
<tr>
<td>Asp</td>
<td>3.01</td>
<td>3</td>
</tr>
<tr>
<td>Glx</td>
<td>1.85</td>
<td>2</td>
</tr>
<tr>
<td>Gly</td>
<td>7.25</td>
<td>7</td>
</tr>
<tr>
<td>Ile</td>
<td>6.00</td>
<td>6</td>
</tr>
<tr>
<td>Lys</td>
<td>2.24</td>
<td>2</td>
</tr>
<tr>
<td>Leu</td>
<td>4.15</td>
<td>4</td>
</tr>
<tr>
<td>Met*</td>
<td>1.85</td>
<td>2</td>
</tr>
<tr>
<td>Ser*</td>
<td>2.08</td>
<td>3</td>
</tr>
<tr>
<td>Thr</td>
<td>2.24</td>
<td>2</td>
</tr>
<tr>
<td>Val</td>
<td>5.10</td>
<td>5</td>
</tr>
<tr>
<td>Tyr</td>
<td>2.21</td>
<td>2</td>
</tr>
<tr>
<td>Arg</td>
<td>1.04</td>
<td>1</td>
</tr>
</tbody>
</table>

* Ser and Met partially destroyed by hydrolysis.

Table 3.1. Amino Acid Analysis Of Synthetic Pf1 Coat Protein
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Actual</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>2.01</td>
<td>2</td>
</tr>
<tr>
<td>Asp</td>
<td>1.93</td>
<td>3</td>
</tr>
<tr>
<td>Glx</td>
<td>1.85</td>
<td>2</td>
</tr>
<tr>
<td>Gly</td>
<td>3.00</td>
<td>3</td>
</tr>
<tr>
<td>Ile</td>
<td>2.04</td>
<td>2</td>
</tr>
<tr>
<td>Ser*</td>
<td>1.51</td>
<td>2</td>
</tr>
<tr>
<td>Thr</td>
<td>1.87</td>
<td>2</td>
</tr>
<tr>
<td>Val</td>
<td>2.10</td>
<td>2</td>
</tr>
</tbody>
</table>

* Ser partially destroyed by hydrolysis.

Table 3.2. Amino Acid Analysis Of the 18-residue Synthetic Peptide Corresponding To The N-terminal Sequence Of Pf1Coat Protein
The purified synthetic Pf1 polypeptide shows a calculated mass of 4606.5 which is in good agreement with that determined by mass spectrometry (Figure 3.10). Similarly the purified N-terminal peptide has calculated value of 1732.8 and this is in good agreement with the determined value by mass spectrometry (Figure 3.11).

3.3.2 Studies of the synthetic Pf1 coat protein in DMPC vesicles before disaggregation in $^2$H$_2$O buffer

Information regarding the secondary structure of polypeptides can be obtained from the frequency of the amide I band in the range of 1615-1695 cm$^{-1}$ (Haris and Chapman, 1992). The correlation between the frequency of the amide I band and secondary structure is now well defined. All band assignments made are taken from (Chirgadze et al. 1975; Byler and Susi, 1986; Krim and Bandekar, 1986). Figure 3.12 shows a comparison of the FTIR absorption spectrum of the synthetic Pf1 coat protein with the natural form reconstituted in DMPC vesicles in $^2$H$_2$O PBS (pD 7.4). Both the conventional and the new reconstitution techniques were applied to the synthetic Pf1 coat proteins. The amide I band maximum of the natural coat protein is observed at 1656 cm$^{-1}$ in the FTIR absorbance spectrum. The 1656 cm$^{-1}$ band is assigned to $\alpha$-helical structure (Haris et al., 1986; Jackson et al., 1989) and the 1734 cm$^{-1}$ is attributed to the vibrations of the carbonyl ester groups of DMPC. The amide II band can be observed at 1550 cm$^{-1}$. The component at 1583 cm$^{-1}$ indicates the presence of carboxylate side-groups in the molecule.

The second derivative analysis (Fig. 3.13) shows the main amide I band
Figure 3.10. Laser-desorption mass spectra (LDMS) of synthetic Pf1 coat protein.

Synthetic Pf1 coat protein

GVIDTSAVESAITDGQQDMKAIGGIVGALVILAVAGLISHMLRA

4606.4

int std

5734

Counts/Mass (M/z)
Figure 3.11. Laser-desorption mass spectra of synthetic 18-residue N-terminal region of Pf1 coat protein.
maximum at 1657 cm\(^{-1}\) and very weak components at 1622 cm\(^{-1}\) and 1679 cm\(^{-1}\). The weak bands may be assigned to \(\beta\)-sheet and turn arrangement respectively (Haris et al., 1986; Jackson et al., 1989) however contributions from amino acid side chain groups cannot be ruled out. The 1549 cm\(^{-1}\) band is assigned to the amide II vibrations and the 1586 cm\(^{-1}\) can be attributed to carboxylate groups in the polypeptide chain. The 1517 cm\(^{-1}\) band is assigned to tyrosine side chain vibrations. Hence the FTIR spectrum of the natural Pf1 coat protein when incorporated into lipid vesicles adopts a predominantly \(\alpha\)-helical arrangement. The FTIR absorbance spectrum for the synthetic Pf1 coat protein in DMPC using the conventional reconstitution technique shows two bands at 1624 cm\(^{-1}\) and 1674 cm\(^{-1}\) at the amide I region. Detailed analysis (Fig. 3.13) using the second derivative procedure reveal strong bands at 1622 cm\(^{-1}\) which is usually assigned to \(\beta\)-sheet arrangement and 1674 cm\(^{-1}\) to turn structures (Susi and Byler, 1986; Byler and Susi, 1986; Haris et al., 1986). The spectrum of the synthetic Pf1 coat protein is therefore dominated by a \(\beta\)-sheet type arrangement corresponding to the aggregated form when using the conventional reconstitution technique.
Figure 3.12. The FTIR spectra of Pf1 coat protein from 1800 cm$^{-1}$ to 1500 cm$^{-1}$ of the synthetic (solid trace) before disaggregation and natural forms (dashed trace) in DMPC vesicles in $^2$H$_2$O buffer at 30$^\circ$C.
Figure 3.13. Second-derivative spectra of Pf1 coat protein of synthetic (solid trace) before disaggregation and natural forms (dashed trace) in DMPC vesicles in $^2$H$_2$O buffer at 30° C.
3.3.3 Studies of the synthetic Pf1 protein in DMPC vesicles after disaggregation in $^2$H$_2$O buffer

Figure 3.14 displays a comparison of the FTIR absorbance spectrum of the synthetic Pf1 coat protein after disaggregation with that of the natural Pf1 coat protein in DMPC vesicles. The synthetic coat protein in DMPC vesicles this time displays a predominant band at 1654 cm$^{-1}$ in the amide I band region. The amide II band can be observed at 1550 cm$^{-1}$ together with the carboxylate groups in the molecule located at 1583 cm$^{-1}$ position. The second derivative analysis shows a weak component at 1636 cm$^{-1}$ with the major component at 1658 cm$^{-1}$. This is similar to the spectrum obtained for the natural Pf1 coat protein in DMPC vesicles indicative of $\alpha$-helical arrangement. The strong 1674 cm$^{-1}$ band observed in the conventional reconstitution method can be assigned to TFA and the 1622 cm$^{-1}$ to aggregated population of the synthetic protein.
Figure 3.14. The FTIR absorption spectra of Pf1 coat protein of the synthetic (solid trace) after disaggregation and natural forms (dashed trace) in DMPC vesicles in $^2$H$_2$O buffer at $30^\circ$ C.
Figure 3.15. Second-derivative spectra of Pf1 coat protein of synthetic (solid trace) after disaggregation and natural forms (dashed trace) in DMPC vesicles in $^2$H2O buffer at 30° C.
3.3.4 Studies of amide hydrogen exchange in synthetic Pf1 coat protein in DMPC vesicles after disaggregation $^2\text{H}_2\text{O}$ buffer

The absorption spectrum in the range $1800-1500 \text{ cm}^{-1}$ of the synthetic Pf1 coat protein in DMPC vesicles in H$_2$O buffer pH 7.4 and $^2\text{H}_2\text{O}$ buffer pD 7.4 is depicted in Figure 3.16 (dashed trace). The two main absorption bands present in H$_2$O buffer are assigned to the amide I band centred at 1658 cm$^{-1}$ and amide II band at 1550 cm$^{-1}$. The band at 1734 cm$^{-1}$ corresponds to the stretching vibrations of carbonyl groups of the acyl esters of the phospholipid. The 1583 cm$^{-1}$ band in the $^2\text{H}_2\text{O}$ buffer can be attributed to the carboxylate groups present in the protein. The second derivative spectrum of the Pf1 coat protein is shown in Fig. 3.17. The dominant bands at 1658 cm$^{-1}$ in the H$_2$O and $^2\text{H}_2\text{O}$ spectrum can be assigned to $\alpha$-helical structure. Further components observed at the amide I band are at 1676 cm$^{-1}$ and 1680 cm$^{-1}$ which can be assigned to $\beta$-turns, with 1630 cm$^{-1}$ and 1636 cm$^{-1}$ assigned to $\beta$-sheet and at 1517 cm$^{-1}$ assigned to tyrosine side chains. As the 1630 cm$^{-1}$ and 1636 cm$^{-1}$ bands are of weak intensity they may arise from side chain groups within the polypeptide. The absorption spectrum is similar to the natural coat protein under an identical environment.
Figure 3.16. The FTIR absorption spectra of synthetic Pf1 coat protein in DMPC vesicles in H$_2$O buffer pH 7.4 (dashed trace) and $^2$H$_2$O buffer pD 7.4 (solid trace) at 30° C.
Figure 3.17. Second-derivative spectra of synthetic Pf1 coat protein in DMPC vesicles in H$_2$O buffer pH 7.4 (dashed trace) and $^2$H$_2$O buffer pD 7.4 (solid trace) at 30°C.
Further information can be obtained regarding the exchangeability of the amide protons of synthetic Pf1 coat protein and examining the secondary structure in $^2\text{H}_2\text{O}$ buffers. This allows the study of conformational stability of the synthetic protein and its interaction with solvents when dispersed in the various media. The absorption spectrum of the Pf1 coat protein in $^2\text{H}_2\text{O}$ buffer pD 7.4 is shown in Fig.3.16 (solid trace), showing the amide I band at 1654 cm$^{-1}$. Further analysis using the second derivative shows additional bands (fig.3.17) which can be assigned as follows: 1658 cm$^{-1}$ $\alpha$-helix, 1676-1680 cm$^{-1}$ from $\beta$-turns, 1630-1636 cm$^{-1}$ from random arrangements, 1583-1586 cm$^{-1}$ emanating from ionized carboxylate side chains of aspartates and glutamates residues, 1549-1550 cm$^{-1}$ from the amide II band indicating that the protein has not been subjected to complete H→$^2\text{H}$ exchange, 1517 cm$^{-1}$ from tyrosine side chains.

Using $^2\text{H}_2\text{O}$ buffers a decrease in the intensity of the amide II band near 1547 cm$^{-1}$, can be observed, primarily because N-deuteration of amide hydrogens shifts the amide N-$^1\text{H}/^2\text{H}$ bending absorption to lower frequencies. From the amide II regions of the spectra shown in Figure 3.17 it is evident that a fraction of amide hydrogens are available for exchange rapidly with the bulk solvent. Comparison of the ratio of the amide I/amide II intensities in $^2\text{H}_2\text{O}$ to that in the H$_2$O spectrum indicates that approximately 60% of the amide protons participate in rapid $^1\text{H}/^2\text{H}$ exchange and the remaining amide protons indicating slow exchange. This accessibility to exchange is in full agreement with previous studies conducted using the natural coat protein using $^1\text{H}$-NMR (Schiksnis et al., 1987) and FTIR spectroscopy (Azpiazu et al., 1993).
3.3.5 Studies of the intramembrane orientation of synthetic Pf1 coat protein in DMPC films

The infrared spectra of the synthetic and natural Pf1 coat protein reconstituted into DMPC films utilizing polarized beams are shown in Figure 3.18 and 3.19. It can be observed that the amide I and II bands are located at similar band frequencies to those obtained for in solution. The bands have similar intensities obtained with natural coat protein in lipid films using polarized beam. The amide I band at 1657 cm\(^{-1}\) can be observed to project a strong intensity when recorded with the beam polarized perpendicular to the plane of the lipid film. This is similar to the spectrum recorded with the natural protein when also recorded with the beam polarized perpendicular to the plane of the lipid film. The amide II bands also indicate similar absorbances in the perpendicular and parallel polarized spectra. From the spectra (Fig. 3.20) recorded with the incident light polarized at 90° and 0°, the dichroic ratio \(R_{\text{str}} = A_{90}/A_{0} = 2.3\) corresponding to the helical structure which indicates that the \(\alpha\)-helical is not parallel but adopts a perpendicular orientation to the lipid acyl chain in the lipid bilayer. The dichroic ratio is similar to those obtained for the natural Pf1 coat protein (Azpiazu et al., 1993) and for bacteriorhodopsin in ATR studies (Yang et al., 1987).
Figure 3.18. Internal reflectance infrared spectra of Pf1 coat protein of synthetic (solid trace) after disaggregation and natural forms (dashed trace) in DMPC films with beams polarized perpendicular to the plane of the film.
Figure 3.19. Internal reflectance infrared spectra of Pf1 coat protein of synthetic (solid trace) after disaggregation and natural forms (dashed trace) in DMPC films with beams polarized parallel to the plane of the film.
Figure 3.20. Internal reflectance infrared spectra of DMPC films with reconstituted synthetic Pf1 coat protein after disaggregation with beams perpendicular (solid trace) and parallel (dashed trace) to the plane of the film.
3.3.6 Studies of the synthetic Pf1 protein in SDS micelles before disaggregation in $^2$H$_2$O buffer

A comparison of FTIR absorbance spectrum of the synthetic Pf1 coat protein with that of the corresponding natural coat protein in SDS micelles is shown in Figure 3.21. Again both the conventional and the new reconstitution procedure were applied to the synthetic coat protein for the analysis. The amide I band maximum of the natural coat protein in the FTIR absorbance spectrum occurs at 1652 cm$^{-1}$ and the band is quite symmetrical with no other prominent band features present. The second derivative analysis of the spectrum shows a major band centred at 1656 cm$^{-1}$ with weak a component at 1626 cm$^{-1}$. The 1657 cm$^{-1}$ band can be ascribed to $\alpha$-helical structure. Applying the conventional technique of reconstitution, the FTIR absorbance spectrum of the synthetic PF1 coat protein in SDS micelles display bands at 1630 cm$^{-1}$, 1652 cm$^{-1}$ and 1674 cm$^{-1}$ in the amide I region. Further analysis using the second derivative procedure reveal strong bands at 1626 cm$^{-1}$, 1674 cm$^{-1}$ and 1652 cm$^{-1}$ in the amide I region. The 1626 cm$^{-1}$ band is usually assigned to $\beta$-sheet type structures and the 1674 cm$^{-1}$ to turn structures. The 1657 cm$^{-1}$ band is assigned to $\alpha$-helical arrangement. The spectrum of the synthetic protein is dominated by the $\beta$-sheet type structure accompanied by $\alpha$-helical with the presence of significant turn structures. The $\beta$-sheet type of structure can only be attributed to the aggregated form of the protein and the turn type structure can be assigned to the presence of residual TFA. These bands are only present whenever the conventional mode of reconstitution is undertaken.
Figure 3.21. The FTIR absorption spectra of Pf1 coat protein of the synthetic (solid trace) before disaggregation and natural forms (dashed trace) in SDS micelles in $^2$H$_2$O buffer at 30° C.
Figure 3.22. Second-derivative spectra of Pf1 coat protein of synthetic (solid trace) before disaggregation and natural forms (dashed trace) in SDS micelles in $^2$H$_2$O buffer at 30°C.
3.3.7 Studies of the synthetic Pf1 coat protein in SDS micelles after disaggregation in $^2$H$_2$O buffer

Figure 3.23 shows the FTIR absorbance spectrum of the Pf1 coat protein in SDS micelles after applying the disaggregation technique. A prominent band at 1656 cm$^{-1}$ in the amide I region is featured and this is attributed to the $\alpha$-helical structure. Detailed analysis applying the second derivative procedure, shows a major band at 1657 cm$^{-1}$ with no other major features confirms the $\alpha$-helical arrangement in SDS micelles. This is identical to the FTIR spectrum obtained for the natural protein in SDS micelles. Applying the new reconstitution technique it can be observed that aggregates and TFA can be eliminated. It can be seen that the predominant structure identical to the natural protein in SDS is also adopted.

3.3.8 Studies of the synthetic 18-residue polypeptide corresponding to N-terminal region of Pf1 coat protein before and after disaggregation in $^2$H$_2$O buffer

The FTIR absorbance spectra of the 18 amino acid long peptide in $^2$H$_2$O PBS buffer (pD 7.4) before and after disaggregation are shown in Figure 3.25. Prominent bands at 1648 cm$^{-1}$ and 1630 cm$^{-1}$ accompanied by a shoulder at 1677 cm$^{-1}$ can be observed in the amide I region. Further analysis (Fig. 3.26, solid trace) applying the second derivative procedure reveals additional components at 1674 cm$^{-1}$, 1644 cm$^{-1}$ and 1644 cm$^{-1}$. The 1674 cm$^{-1}$ band is usually assigned to turn type structure and the 1626 cm$^{-1}$ assigned to $\beta$-type arrangement.
Figure 3.23. The FTIR absorption spectra of Pf1 coat protein of the synthetic (solid trace) after disaggregation and natural forms (dashed trace) in SDS micelles in $^2$H$_2$O buffer at 30° C.
Figure 3.24. Second-derivative spectra of P1 coat protein of synthetic (solid trace) after disaggregation and natural forms (dashed trace) in SDS micelles in $^2$H$_2$O buffer at 30° C.
The 1644 cm\(^{-1}\) band can be attributed to random coil conformation. The spectrum of the synthetic N-terminal region of the Pf1 coat protein is dominated by a random coil structure and a \(\beta\)-sheet type arrangement which can be assigned to the aggregated form with TFA in evidence.

The FTIR absorbance spectrum of the N-terminal peptide after disaggregation is shown in Figure 3.25 (dashed trace) and is dominated by the band at 1644 cm\(^{-1}\) in the amide I region. Further analysis applying the second derivative technique reveals only one additional band at 1673 cm\(^{-1}\). This band is very weak in intensity that it may originate from turns or bends within the peptide structure. The main amide I band at 1644 cm\(^{-1}\) can be assigned to random coil structures (Jackson et al., 1989).
Figure 3.25. The FTIR absorption spectra of the synthetic 18-residue N-terminal peptide of Pf1 coat protein before disaggregation (solid trace) and after disaggregation (dashed trace) in $^2$H$_2$O buffer at 30° C.
Figure 3.26. Second-derivative spectra of the synthetic 18-residue N-terminal peptide of Pf1 coat protein before disaggregation (solid trace) and after disaggregation (dashed trace) in $^2$H$_2$O buffer at 30° C.
3.3.9 Studies of the synthetic 18-residue polypeptide corresponding to the N-terminal region of Pf1 coat protein in DMPC before and after disaggregation in $^2 \text{H}_2 \text{O}$ buffer

The FTIR absorbance spectrum (Fig. 3.27) of the N-terminal peptide in DMPC vesicles before disaggregation reveal bands at 1624 cm$^{-1}$ and 1648 cm$^{-1}$ in the amide I region. Detailed analysis using the second derivative procedure show strong bands at 1674 cm$^{-1}$, and 1622 cm$^{-1}$ in the amide I region (Fig. 3.28). The 1674 cm$^{-1}$ band is assigned to turn structures and the 1622 cm$^{-1}$ is usually assigned to $\beta$-sheet arrangement. Again the FTIR spectrum is dominated by the $\beta$-sheet type arrangement which can be assigned to the aggregated form. The 1674 cm$^{-1}$ band in this instance can be assigned to the presence of TFA in the peptide.

Using the disaggregation procedure, the FTIR absorbance spectrum (Fig. 3.27) shows a prominent band at 1648 cm$^{-1}$ in the amide I region. Further analysis using the second derivative procedure indicates a prominent band at 1648 cm$^{-1}$ in the amide I region (Fig. 3.28). The 1648 cm$^{-1}$ band maximum is assigned to random coil structure. Random coil arrangement is also the dominant structure of the N-terminal peptide when in lipid environments.
Figure 3.27. The FTIR absorption spectra of the synthetic 18-residue N-terminal peptide of Pf1 coat protein before disaggregation (solid trace) and after disaggregation (dashed trace) in DMPC vesicles in $^2$H$_2$O buffer at 30° C.
Figure 3.28 Second-derivative spectra of the synthetic 18-residue N-terminal peptide of Pf1 coat protein before disaggregation (solid trace) and after disaggregation (dashed trace) in DMPC vesicles in $^2$H$_2$O buffer at 30° C.
3.4 DISCUSSION

In this study, a unique opportunity using FTIR to analyse chemically synthesized Pf1 coat protein, applying two reconstitution procedures in membrane mimetic systems and then making comparisons with the definitive structural data of the natural Pf1 coat protein derived from separate NMR (Shon et al., 1991) CD (Schiksnis et al., 1971) and FTIR (Azpiazu et al., 1993) was undertaken. Previous NMR, CD and FTIR studies of the natural coat protein have shown a predominantly $\alpha$-helical structure when incorporated in lipid vesicles and detergent micelles. The X-ray studies on the viral fibres have shown that the bacteriophage contains a high amount of $\alpha$-helical structure, with the $\alpha$-helical segments being orientated in the direction of the filament (Marvin and Wachtel, 1975).

The application of the conventional reconstitution technique without prior disaggregation to incorporate the synthetic Pf1 coat protein in lipid vesicles and SDS micelles gave rise to spurious peaks around 1630-1626 cm$^{-1}$ and 1674 cm$^{-1}$. These bands are consistent whenever the conventional technique of reconstitution was applied to incorporate the protein in membrane mimetic systems. Using the new reconstitution procedure where the synthetic protein was disaggregated before incorporation into these environments, these characteristic bands were completely eliminated, thus revealing the true conformation. This is illustrated by comparing the structure of the synthetic protein in DMPC vesicles and SDS micelles with that of the natural protein in these enivroments. The 1620-1626 cm$^{-1}$ band position in the amide I region is often assigned to $\beta$-sheet structure and the 1674 cm$^{-1}$ to turn/bend
arrangements. With the current experimental data it can be shown that this band position is due to aggregates promoted by inter/intramolecular hydrophobic interactions of the amino acid residues.

The 1620-1630 cm⁻¹ band position in the amide I region is often encountered when membrane proteins and hydrophobic synthetic polypeptides are subjected to numerous purification procedures and when exposed to aqueous environments. It has also been observed in proteins and polypeptides undergoing heat or pH denaturation. The intense 1674 cm⁻¹ band position can be unequivocably assigned to TFA whenever proteins or synthetic polypeptides have undergone reverse-phase HPLC purification. TFA has the propensity to associate with residues that are positively charged when it forms an ion pair in HPLC purifications. This band position in the amide I region can be eliminated using the new disaggregation procedure and replacing the bound TFA counterions by chloride ions before reconstitution into the appropriate medium. The results also indicate that elimination of the insoluble aggregates, using the solubilising properties of TFA before reconstitution, enabled the synthetic Pf1 coat protein to adopt its natural membrane conformation when incorporated into membrane mimetic environments. Previous studies applying FTIR, NMR and CD spectroscopy have all shown that the natural Pf1 coat protein does indeed adopt a predominantly α-helical arrangement in lipid vesicles and detergent micelles similar to the synthetic Pf1 coat proteins when disaggregated in these environments.

Recent NMR studies of the natural Pf1 coat protein in lipid membranes and micelles lead to the proposal of a model for its orientation in the lipid
membrane (Shon et al., 1991). Based on the tertiary structure of the coat protein in DPC micelles as determined by 2D combined $^1\text{H}^\text{15N}$-NMR spectroscopy, the model purports that the structure of the protein in detergent micelles is similar to that in the lipid membrane. This model was further supported by FTIR spectroscopic studies of the natural Pf1 coat protein in detergent micelles and lipid vesicles (Azpiazu et al., 1993). Using synthetic Pf1 coat protein it has also been shown to adopt primarily an $\alpha$-helical conformation in lipid vesicles and detergent micelles.

It was also proposed that an $\alpha$-helical domain was arranged with its axis perpendicular to the plane of the membrane and that another domain corresponding to the N-terminal region was orientated parallel to the surface of the lipid bilayer. It is assumed that the helix corresponding to the N-terminal region is amphipathic with its hydrophobic face buried within the lipid acyl chains and its hydrophilic face exposed to the aqueous environment. The FTIR spectroscopic studies using synthetic peptide corresponding to the N-terminal region of the coat protein in aqueous buffer and lipid vesicles indicate a predominantly random coil arrangement. This is an indication that this region of the protein has the propensity to be on the outside of the membrane exposed to the aqueous environment, however it may adopt an $\alpha$-helical structure when being a part of the whole protein. If as suggested that it is amphipathic, a random coil to helix transition would have been observed for the N-terminal peptide when transferred from an aqueous to a lipid environment. This type of conformational transition has been observed in the transmembrane S4 polypeptide corresponding to the voltage sensitive Shaker
potassium channel (Haris et al., 1994). Previous studies using oriented lipid films containing the natural Pf1 coat protein and applying internal reflectance polarized FTIR spectroscopy, it was shown that the average \( \alpha \)-helical structure is arranged perpendicular to the plane of the lipid membrane (Azpiazu et al., 1993). The ATR-FTIR studies of the oriented lipid/polypeptide mixtures indicate that the dichroic ratio of the amide I absorption band is approximately 2.3, indicating that the long axes of the polypeptide helices are predominantly oriented perpendicular to the bilayer plane. This data is a similar to the ATR-FTIR data obtained for the natural Pf1 protein. Fluorescence polarization and calorimetric studies have also shown that the amino-terminal region of the natural protein is composed of a single helix spanning the lipid bilayer (Azpiazu et al., 1993).

From the results obtained for chemically synthesized Pf1 coat protein, some important conclusions can be drawn regarding the conformational properties when compared with natural Pf1 coat protein structure. It can be directly demonstrated that the amino acid sequence corresponding to the natural protein or predicted from cDNA can confer the correct folding in membrane mimetic environments only when disaggregated. The fact that chemically synthesized Pf1 coat proteins are non-biological, can be advantageous ruling out the possibility of contaminating molecules or differently encoded products of the same gene, to be responsible for structures attributed to Pf1 coat protein.

The successful chemical synthesis of Pf1 coat protein suggests that chemically synthesized polypeptides can be applied to the study of other large
hydrophobic polypeptide domains corresponding to transmembrane regions. The approach described in this study have been extended to other membrane proteins and their structure determined in membrane mimetic systems (Haris et al., 1994). Chemical synthesis of membrane proteins by the technique described in this chapter in combination with spectroscopic techniques, may permit further characterization of the structural basis of biological function, for example, analysis of the 3D structure of crystallized synthesized products. The successful application of chemically synthesized Pf1 coat protein to determine secondary structure has lead to new insights into other membrane protein structures and biological functions (Haris et al., 1994).

3.5 SUMMARY

Polypeptide sequences corresponding to transmembrane regions are usually highly hydrophobic, which can promote rigid inter/intramolecular hydrophobic interaction favouring the formation of aggregates and macroaggregates in aqueous environments. The introduction of a new disaggregation process which can be applied to hydrophobic polypeptides so as to enable them to be incorporated into membrane mimetic environments without exhibiting this aggregated form was applied with success. This was clearly illustrated by a comparison of the FTIR spectrum of the natural coat protein in lipid vesicles and detergent micelles prior to and after disaggregation. The natural Pf1 coat protein using FTIR spectroscopy displayed an α-helical structure and the disaggregated synthetic protein also displayed an α-helical structure, whilst the aggregated synthetic protein exhibited intermolecular β-sheet type
arrangement. These aggregated structural forms can often lead to spurious assignments when dealing with denatured or hydrophobic proteins. Also attenuated total reflectance FTIR studies of oriented phospholipid bilayers demonstrate that synthetic Pf1 coat protein is arranged with the long axis perpendicular to the bilayer plane. The FTIR studies of the exchangeability of the amide protons of synthetic Pf1 coat protein indicate that in a lipid environment there are two populations of amide protons exchanging with the bulk solvent at different rates. This is consistent with a proportion of the coat protein being exposed to solvent and the other being buried in the lipid bilayer. Moreover by synthesizing the N-terminal region of the coat protein and comparing its structure in aqueous and lipid environment it was possible to propose its location relative to the membrane.
CHAPTER FOUR

STRUCTURE OF SYNTHETIC POLYPEPTIDES CORRESPONDING TO THE PORE-FORMING AND TRANSMEMBRANE DOMAINS OF A VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN IN MODEL LIPID MEMBRANES AND MICELLES
4.1 INTRODUCTION

The transport of small molecules and ions across biological membranes is a vital process in all living cells. To translocate against or along existing solute concentration gradients specific transport systems such as pumps and channels have evolved. The maximum rates of transport are achieved by channel proteins that when in their activated ("open") state permit flux rates approaching that of free diffusion (Hille, 1984).

In excitable cells, such as nerve and muscle, voltage-gated ion channel proteins provide the molecular basis for electrical signalling and selective information transmission between excitable cells and thus represent a diverse population. Many of the neuronal channel proteins are concentrated at synapses where cell to cell contacts are essential for interneuronal communication. Voltage-sensitive ion channels modulate the propagation of action potentials along the neuronal plasma membrane and regulate the release of neurotransmitters from presynaptic nerve terminals. Channel proteins activated by neurotransmitters function as receptors for rapid transmembrane signalling at the presynaptic membrane of classical "chemical" synapses. In cell to cell communication, gap junction channels connect the cytoplasms and provide direct electric coupling via "electrical" synapses.

In recent years the study of voltage-gated ion channels has been dominated by molecular biology where sequence information and most recently on the relationship between structure and function were elucidated (Miller, 1989) The first voltage-activated K+ channel genes, those encoding the Drosophila Shaker A-type K+ channels, were cloned and expressed (MacKinnon...
et al., 1989. Since then a variety of $K^+$ channels have been discovered and identified. Functionally, the channels are very similar in that they possess two strikingly, uniform properties: they are highly selective for $K^+$ ions as opposed to $Na^+$ or $Ca^{2+}$, and they are activated (opened) by cell membrane depolarization, i.e. they are voltage-gated. Amino acid sequence analysis reveal them as members of the same closely knit gene family of $K^+$ channels. Furthermore, hydropathy analysis and comparison of the primary structure (Fig. 4.1) indicate that they are related to the voltage-activated $Na^+$ and $Ca^{2+}$ channels (Catterall, 1988). A clearer picture of what $K^+$ channels looks like and a better comprehension of how they function have emerged through the combine use of biophysical and molecular genetic techniques.

The voltage-activated $K^+$ channels are multimers which are built through the coassembly of the identical or similar subunits. This multimeric nature of the $K^+$ channels have been demonstrated by coexpressing two kinetically or pharmacologically distinct $K^+$ channels which resulted in the ionic currents displaying a component with hybrid behaviour (Christie et al., 1990; Isacoff et al., 1990; Ruppersberg et al., 1990). This confirmed that independent expression would have resulted in ionic current of the parent channels and not of the hybrid behaviour. Thus, the coassembly of non-identical subunits have resulted in the formation of a unique, hybrid channel. These studies provided the first direct evidence that $K^+$ channels were indeed composed of multimers. It also suggested a possible mechanism for the generation of $K^+$ channel functional diversity (Jan and Jan, 1990). A wide range of intermediate channels becomes possible by starting with just a few different subunit types.
Figure 4.1. Proposed transmembrane topology and subunit number of Na\(^+\), Ca\(^{2+}\) and K\(^+\) channel proteins.
Amino acid sequences for a number of channel proteins such as the voltage-gated sodium, calcium and potassium channels are now accessible (Perney and Kaczmarek, 1991). Site-directed mutagenesis studies have identified functionally important residues (Hartman et al., 1991; Yellen et al., 1991; Yool and Schwarz, 1991). The voltage-gated channels comprise a molecular family of strongly homologous proteins. Na$^+$ and Ca$^{2+}$ channels are assembled from a single polypeptide of about 2000 residues, comprising four homologous domains, each of which is thought to contain six transmembrane segments. K$^+$ channels on the other hand have a single such domain of 600 residues and appear to function as tetramers as shown in Figure 4.1. Recent mutagenesis studies have shown that a localized region of the polypeptide chain linking the fifth and sixth (Fig. 4.2) membrane-spanning $\alpha$-helical stretches ($S5$ and $S6$) forms the K$^+$ channel ion-selective pore. This mapping was only possible with the mechanistic analysis of probe molecules that bind specifically to the pores of K$^+$ channels. The three types of ligand: scorpion venom peptides, tetraethylammonium (TEA) and the permeating ions themselves, known to interact specifically with the pore of K$^+$ channels were applied to locate this region.

In the search for the pore, the S5-S6 linker region was first brought to prominence by the discovery of point mutations in this region specifically and locally which altered the binding of scorpion-venom peptides, such as charybdotoxin, molecules known to occlude the outer vestibule of the channel (MacKinnon and Miller, 1989; MacKinnon et al., 1990; Goldstein and Miller, 1992). It was soon followed by the identification of point mutants that clearly
Figure 4.2. Transmembrane folding model for a voltage-activated $K^+$ channel subunit. Hydropathy profiles indicate the presence of six hydrophobic stretches (S1 through S6). Recent evidence indicates that the polypeptide chain between S5 and S6 forms the channel pore. Amino acid positions known to affect channel blocking by TEA are shown as grey circles and those shown as black circles are known to determine the channel's ion selectivity. The plus symbols in S4 indicate the presence of basic residues at every third position.
affected TEA binding to two distinct positions, one on the external side and one on the internal side of the pore (MacKinnon and Yellen, 1990; Yellen et al., 1991). These selective binding properties were instrumental to the identification of a 20-residue amino acid sequence (Fig 4.3) that spans the membrane twice, to form the narrow part of the pore, and that this short stretch embodies most or all of the pore-determinants (Hartman et al., 1991). It was also subsequently shown that a high affinity TEA-binding site could be assembled with four aromatic rings donated by tyrosine or phenylalanine from each subunit (Heginbotham et al., 1992; Kavanaugh et al., 1992). This information is a clear indication that the conduction pore is assembled on the four-fold axis of symmetry of the tetrameric complex. Also the introduction of point mutations in the inner pore region between the TEA binding sites, the passage of NH$_4^+$ and Rb$^+$ ions with diameters larger than K$^+$, were enhanced without compromising the ability of the channel to exclude the smaller cation Na$^+$. These studies revealed the location of the selectivity-determinants but not the mechanism of the discrimination amongst cations.

Despite the great progress made by these techniques, the details of the molecular mechanism of channel function are only likely to emerge when their 3D structure are known. However, lack of availability of purified K$^+$ ion channel protein and the complexity involved has hindered structural analysis using techniques normally used for protein structural analysis. One practical approach to this problem is to simplify these large complex proteins by conceptually dissecting them into domains/motifs, deduced from the primary structure.
Figure 4.3. The amino acid sequence of the pore-forming region of the ShA $K^+$ channel shown above a postulated membrane folding structure. O, mutations affecting external blockade by TEA; +, mutations affecting scorpion toxin blockade. A mutation at 11 affected only internal TEA blockade. The arbitrary numbering of residues is used as reference.
These isolated domains can then be chemically synthesized and their secondary structure determined by various spectroscopic techniques. The use of synthetic polypeptides as models to study specific ion channels and pore-forming proteins is becoming increasingly popular (Sansom, 1991). This approach is used in this chapter to investigate a functionally important segment of a potassium ion channel protein.

The sequence corresponding to the postulated pore region (designated H5 or SS1-SS2, amino acid residues 431-449) of the Shaker A K⁺ channel from Drosophila (Tempel et al., 1987; Yool and Schwarz, 1991) together with the extended version (Fig. 4.4a and 4.4b) have been chemically synthesized and spectroscopic analysis undertaken to investigate its structure. FTIR spectroscopy and CD have been applied to study the structure of the disaggregated synthetic polypeptides in model lipid membranes and micelles. Furthermore a tetrameric version (Fig 4.4c) of the H5 sequence was also synthesized to simulate the native arrangement that is proposed to exist in the membrane. A polypeptide of 119 amino acid residues (Fig. 4.4d) incorporating the transmembrane domains S4-S5-H5-S6 was also synthesized using the new fragment condensation technique on solid phase (Fig. 4.5). The structures of the synthetic polypeptides are investigated in different environments such as in LPC micelles, DMPC and DMPG vesicles dispersed in ²H₂O buffer. The ion channel activity of these polypeptides in planar lipid bilayers are also examined.

This work was undertaken in collaboration with Dr. P. I. Haris (this laboratory) and Dr. M. S. P. Sansom (Laboratory of Molecular Biophysics, University of Oxford).
Figure 4.4. Amino acid sequences of the synthetic (a) H5 polypeptide; (b) extended version of H5 (additional amino acid residues are underlined); (c) tetrameric H5 polypeptide with branching polylysine core; (d) S4-S5-H5-S6 polypeptide.
4.2 MATERIALS AND METHODS

DMPC, DMPG, CHOL, LPC and PC (type II-S) were all purchased from Sigma Chemicals Ltd., (U.K.) with diPhyPC from (Avanti Polar Lipids). Fmoc amino acids with side-chain protection groups, TBTU, ODHBT, Rink-amide, Rink-MBHA and HMPB-MBHA resins were all purchased from NovaBiochem Ltd., (U.K.). Peptide synthesis grade solvents, TFA, TFE, ACN and chloroform were purchased from Rathburn Chemicals Ltd., (Scotland). EDT, thioanisole and indole were purchased from Fluka Ltd., (U.K.).


0.5g of the Rink resin derivatized with Fmoc-amide was put through normal deprotection cycle with 20% piperidine in DMF (in a reaction vessel for 12 minutes). The resin was washed with DMF (10ml/g; 5x1 minute). A 5-fold excess (based on the loading) of acylating species in 0.2M NMP in DMF (5 ml) were added automatically in all subsequent couplings. The synthesis was carried out using the batch synthesis mode on automated peptide synthesizer (Rainin PS3, Protein technologies, USA) with standard protocols and scale (0.1mmol theoretical yield of crude peptide). The cycle for the addition of protected amino acid consisted of 5x2 minutes wash of the solid support with 20% piperidine in DMF to cleave the N\(^{-}\)-Fmoc-group, 12 minute DMF wash, 25 minutes coupling reaction with 5 equivalents of Fmoc-amino acid TBTU active ester, and an 8 minutes DMF wash for a total cycle time of about 50 minutes.

0.5g of the Rink resin derivatized with the Fmoc-amide was subjected to deprotection cycle with 20% piperidine in DMF in a glass reaction vessel for 12 minutes. The resin was thoroughly washed with DMF (10mg/g, 5x1 minute). A 5-fold excess (based on the resin loading) of acylating species in 0.2M MMP in DMF (5 ml) were delivered to all subsequent couplings. The synthesis was carried out using the batch synthesis procedure on an automated peptide synthesizer (Rainin PS3, Protein Technologies USA) with modified protocols and scale (0.1 mmol theoretical yield of crude peptide). The cycle for the addition of Fmoc-amino acid consisted of a 5x2 minutes wash of the solid support with 20% piperidine in DMF to deprotect the N\textsuperscript{\textalpha}-Fmoc-group, 12 minutes DMF wash, 25 minutes, 25 minutes of coupling reaction with 5 equivalents of TBTU activated Fmoc-amino acids for the first 22 residues, and 35 minutes of coupling reaction for the remainder of the residues, and 8 minutes DMF wash for a total cycle time of 60 minutes.

4.2.3 Assembly of the tetrameric H5 peptidyl-resin sequence: (Fmoc-P-D-A-F-W-W-A-V-V-T-M-T-T-T-V-G-Y-G-D-M-T-P)_4-K\textsubscript{2}-K-Rink-amide-resin

The synthesis of a tetrabranched matrix core with H5 peptide attached was accomplished by a stepwise solid-phase procedure on a Rink-amide-resin in which 0.15 mmol of primary amino groups are present. The synthesis of the first and second level of the template was achieved using 5 molar excess
(based on the resin loading) of activated active ester of N°, N°-Fmoc-K(Fmoc) in DMF (10ml/g resin). The Rink-resin derivatized with Fmoc-K^2(Fmoc)-K-Rink-resin was put through normal deprotection cycle with 20% piperidine in DMF in a reaction vessel for 12 minutes to expose the four functional amino groups. The resin was washed with DMF (10ml/g, 5x1 minute).

A 5-fold excess (based on the new loading) of acylating species in 0.2M MMP in DMF (5ml) was added in all subsequent coupling. The synthesis was carried out using the batch synthesis method using an automated peptide synthesizer, with modified protocols and scale (0.1mmol theoretical yield of crude peptides). The cycle for the addition of activated Fmoc-amino acid consisted of 5x2 minutes wash of the solid support with 20% piperidine in DMF to remove the N°-Fmoc-groups, 12 minutes DMF wash, 25 minutes coupling reaction and 8 minutes DMF wash for a total cycle of about 60 minutes.

4.2.4 Assembly of the 119-residue polypeptide incorporating the transmembrane domains S4-S5-H5-S6

The sequence of 40 amino acids beginning from the carboxyl terminal of the S6 transmembrane region was assembled first onto the Rink-amide resin (Fig. 4.5). The Rink-amide-MBHA (0.5g) functionalized with the Fmoc-amide was placed in a glass reaction vessel and deprotection carried out using 20% piperidine in DMF. After removal of the Fmoc group on the resin, acylation with the first amino acid active esters prepared from pre-weighted Fmoc-amino acid in TBTU was carried out. A 5-fold excess (based on the resin loading) of

230
Figure 4.5. Schematic representation of solid phase fragment condensation synthesis of the 119-residue S4-S5-H5-S6 polypeptide.
acylating species in 0.2M MMP in DMF (5ml) was used throughout the synthesis. The synthesis was performed using the batch method on an automated peptide synthesizer with standard protocols and scale (0.1 mmol theoretical yield of crude peptide). On completion of synthesis, this first protected peptide fragment was left attached to the resin for the subsequent fragment condensation synthesis.

To generate fully protected polypeptide fragments containing the respective transmembrane regions, the next 40-residue polypeptide beginning one after the end of each fragments, were assembled. Finally, the 39-residue protected polypeptide beginning one after where the second 40-residue polypeptide terminated was synthesized. The hyper acid-labile linker (HMPB) containing resin was used in this phase of the synthesis to generate fully protected polypeptide fragments. The polypeptides were synthesized according to the normal protocol already described for the first part of the synthesis.

On completion of synthesis of the respective polypeptide fragments, the polypeptide containing HMPB-resins were pre-swelled in DCM in sintered glass columns and the excess DCM removed by vacuum. To detach the protected polypeptides, cleavage solution containing 1% TFA in DCM (10ml) were added to sintered glass columns and agitated with nitrogen for 3 minutes. The cleavage solutions containing the protected polypeptides were then filtered into round bottomed flask containing 10% pyridine in methanol (5ml). This was repeated 3 or 4 times for all the fragments. Filtrates containing similar protected polypeptides were combined and rotary evaporated to quarter of a volume. Cold ether (40ml) was added to the filtrate to aid precipitation.
of the product. Cold ether was again used to wash the product three more times. The samples were finally dried under high vacuum in a freeze-drier. The dried protected polypeptides were then applied step-wise as activated active esters to the deprotected first protected fragment still attached to the Rink amide-resin. The coupling reaction was allowed to proceed for 30 minutes after which the resin was washed with DMF. The resin was then subjected to a final deprotection step to remove the N"-Fmoc of the second fragment. Finally the third activated protected fragment was added to the resin and allowed to react for 30 minutes.

4.2.5 Detachment and deprotection of the monomeric, extended tetrameric H5 and the 119-residue polypeptide incorporating the S4-S5-H5-S6 transmembrane domains

After the completion of the synthesis of the monomeric, tetrameric, extended H5 and the 119 residue polypeptide, their respective N"-Fmoc-protective groups were removed using 20% piperidine in DMF. The exposed N" groups were blocked by acetylation. This procedure was carried out by adding Ac₂O to the aminopeptidyl-resin (0.47ml; 5mmol) in 15 ml of DMF. The acetylation mixture was agitated in the reaction vessel for 30 minutes at room temperature. After acetylation the polystyrene resins were washed with DMF (5ml x 1 minute) and finally washed with DCM (5ml x 1 minute) and dried under vacuum for 2 hours. The dried amide resins were resuspended in 10% TFA in DCM and agitated with nitrogen in a sintered column, and collected into a round bottomed flask at 10 minute intervals. This procedure was carried out
over a period of 60 minutes for all the polypeptides when synthesis was completed. Following this, the TFA and DCM were completely removed by rotary evaporation and the protected polypeptide amides are usually concentrated at this stage. Finally to remove all the protecting groups completely, 82% TFA in the presence of the scavengers thioanisole (5%), phenol (5%) and ethanedithiol were added to the residual oily solution and left at room temperature with occasional swirling for 4 hours. The TFA together with the scavengers were removed by rotary evaporation at 50°C. The crude products were precipitated and thoroughly washed with chilled anhydrous diethyl ether. All the samples were finally dried under high vacuum and stored at -20°C until further use.

4.2.6 Purification of the synthetic monomeric, extended and tetrameric H5 and the 119-residue polypeptide containing the S4-S5-H5-S6 transmembrane domains

The crude monomeric H5 and the extended H5 were purified by multiple reverse-phase HPLC (Varian, USA) runs. The crude polypeptides were dissolved in TFA/TFE (1:1,v/v) and injected into Vydac C_{18} (218 TP 1010 RP preparative column) equilibrated in 50% buffer A (TFA/ACN/H2O, 1:1:98, v/v) and buffer B (TFA/ACN/H2O, 1:90:9, v/v). The polypeptides were purified on a linear gradient of buffer B (from 50 to 70% in 30 minutes) at a flow rate of 4ml/minute. Homogeneity was assessed by a second HPLC analysis on a narrow-bore Vydac C_{18} (201 TP 54 RP analytical column) equilibrated as described.
The crude tetrameric H5 and the 119 amino acid polypeptide containing the S4-S5-H5-S6 transmembrane domains were also purified by multiple reverse-phase HPLC runs. The polypeptides were solubilized in TFA /TFE (1:1,v/v) injected into Vydac C₄ (214 TP 1010 RP preparative column) equilibrated in 50% buffer A (TFA/CH₃CN/H₂O 1:1:98) and 50% buffer B (TFA/CH₃CN/H₂O 1:90:9). The polypeptides were purified on a linear gradient of buffer B (from 50 to 70% in 30 minutes) at a flow rate of 4.0 ml/minute. The homogeneity of the polypeptides were further confirmed by a second HPLC analysis on a narrow-bore Vydac C₄ (214 TP 54 RP analytical column) equilibrated as described above.

4.2.7 Amino acid compositional and molecular mass analysis

All the HPLC purified polypeptides were subjected to hydrolysis with 6M constant boiling HCL containing 0.1% phenol at 110°C for 20-70 hours in nitrogen rich atmosphere. After hydrolysis, the HCL were removed under high vacuum with the aid of a freeze-drier for 4 hours. The dried hydrolysis mixture was reconstituted in amino acid sample buffer and analysed using an amino acid analyzer (LKB Biochrom 4151 Alpha Plus, amino acid analyser)

Samples in duplicates were prepared and calibrated using 10nmol standards. The identity and purity of the monomeric, extended and tetrameric H5, were further confirmed by matrix assisted laser desorption mass spectrometry (LASERMAT, Finnigan MAT Ltd., UK). The molecular weight of the polypeptide incorporating the domains S4-S5-H5-S6 was estimated using SDS-PAGE electrophoresis (Schägger and von Jagow, 1971) using the
disaggregation process due to its extreme insolubility which made analysis by
laser mass analysis difficult.

4.2.8 Reconstitution of synthetic monomeric, extended, tetrameric H5 and the
119-residue polypeptide incorporating the S4-S5-H5-S6 transmembrane
domains in phospholipid vesicles and micelles

Reconstitution of all the above mentioned synthetic polypeptides
involved using the new disaggregated procedure which was successfully
applied for the study of Pf1 coat protein in model lipid membranes described
in chapter three. Typically polypeptides (1 mg) were solubilized in neat protein
sequencing grade TFA (2 ml) which was then removed by rotary evaporation to
afford a thin film of the polypeptide on the surface of round bottomed glass
flask (5 ml capacity). The polypeptide films were further dried for 24 hours
under high vacuum using a freeze drier. To eliminate residual TFA associated
with the polypeptides, it was lyophilized with 0.1 M hydrochloric acid. To these
dried polypeptide films, DMPC or DMPG or LPC in DCM was added and rotary
evaporated to overlay a thin film of the respective lipids. The mixtures were
further dried for another 24 hours before resuspensions in $^2$H$_2$O buffers were
carried out. For FTIR and CD studies, PBS in $^2$H$_2$O (pD 7.4) was added to the
dried films and vortexed vigorously for 5 minutes at room temperature. The
polypeptide to phospholipid molar ratios used in this study were 1:45.

4.2.9 FTIR spectroscopy

FTIR spectra were recorded as described in section 3.2.9.
4.2.10 CD spectroscopy

CD spectra of the synthetic monomeric H5 polypeptide in LPC micelles were recorded on a Jasco J600 spectrometer. The spectra were recorded using either an 0.01 or 0.02 cm path-length cell. Spectra of aqueous dispersions of phospholipid were subtracted from the polypeptide spectra to remove any effects of slight absorbances from phospholipids. The ratio of polypeptide to phospholipid used was the same as that used for the FTIR measurements. However, the samples were diluted such that a polypeptide concentration of 1.25mg/ml was obtained. Spectra were corrected for concentration and the ΔE values calculated.

4.2.11 Planar bilayer studies

Planar lipid bilayers were formed across a small (100μm) hole in Teflon film by the apposition of two monolayers on either side of a Teflon partition (Montal et al. 1972). The lipid forming solution was phosphatidylcholine and CHL, present in a 9:1 weight ratio, dissolved to 5 mg /ml in pentane. The monomeric H5 polypeptide was present in the lipid-forming solution at a 1:100 molar ratio. In studies using the tetrameric H5, the biayers were formed in a similar manner but using a 5mg/ml solution of diPhyPC in pentane. The tetrameric H5 polypeptide dissolved in 80% TFA was added with stirring to the cis face of the bilayer to a final concentration of 2.5 μg/ml. In all of the bilayer experiments the electrolyte solution was 0.5M KCL, buffered with 10mM BES (monomeric H5 experiments) or 200 mM BES (tetrameric H5 experiments) to pH7. The transbilayer currents were measured using an Axopatch ID amplifier.
and were recorded using a sony VCR and PCM.

4.2.12 Molecular modelling

Molecular modelling was performed using QUANTA 3.2 (Molecular Simulations, Waltham, MA) running on a Silicon Graphics Indigo workstation. Simulated annealing/molecular dynamics was carried using Xplor 3.0 (Brunger, 1992) as previously described (Nilges et al., 1988; Nilges and Brunger, 1991; Kerr and Sansom, 1993b). Structure diagrams were generated using MolScript (Kraulis, 1991).

Molecular hydrophobicity potentials (MHPs) were utilized to assess the degree of amphipathicity of the monomeric H5 α-helices (Brasseur, 1991), with the modification that the potential was set to a large positive (i.e. unfavourable) value whenever a grid point fell within the van der Waals radius of an atom. The MHP is based on the free energy of transfer of constituent atoms from a hydrophilic to a hydrophobic phase. A hydrophobic region of a molecule would correspond to a negative MHP. MHP’s were calculated at successive positions on a cylindrical polar grid defined by \((z, \phi)\), where \(z\) is coincident with the helix axis, \(r\) is the distance from the centre of the helix axis and \(\phi\) is the angle subtended in the \((z, \phi)\) plane. Typically \(z\) and \(r\) were incremented in 1Å steps and \(\phi\) in 10° steps, corresponding to ~20,000 grid points for \(z = -25\) to \(+25\)Å. The MHP calculation generates an array containing a potential at each grid point \(E(z, r, \phi)\). For each value of \((z, \phi)\), the minimum of \(E\) with respect to \(r\) is determined, thus defining an \(E_{\text{min}}(z\phi)\) surface, i.e. a hydrophobic surface map (Kerr and Sansom, 1993a). Averaging
\[ E_{\text{MIN}}(z, \phi) \] over all \( z \) values corresponding to the length of the helix generates a plot of \( <E_{\text{MIN}} > \) versus \( \phi \) which may then be used to identify the centres of the hydrophobic and hydrophilic faces of the helix.

4.3 RESULTS

4.3.1 Chemical and physical analysis of the synthetic monomeric, extended, tetrameric H5 and the 119-residue polypeptides incorporating the transmembrane S4-S5-H5-S6 domains

The synthetic monomeric, extended, tetrameric H5 and the 119-residue polypeptide containing the S4-S5-H5-S6 transmembrane sequences (Fig.4.4) were subjected to preparative scale reverse-phase HPLC purification and characterization. The polypeptides eluted as well-resolved peaks which afforded easy isolation. Homogeneity was confirmed using analytical reverse-phase HPLC (Figs. 4.6, 4.7, 4.8, and 4.9). The HPLC purified polypeptides were identified by their amino acid composition as shown in Tables 4.1, 4.2, 4.3, and 4.4). All the respective major peaks could be accounted for as the correct synthesized products. The monomeric, extended, and tetrameric H5 synthetic polypeptides show calculated molecular masses of 2190.97, 4356.03 and 9924.60 respectively which are in good agreement with those determined by mass spectrometry (Fig. 4.10, 4.11 and 4.12).
Figure 4.6. Analytical reverse-phase HPLC of synthetic H5 polypeptide after purification.
Figure 4.7. Analytical reverse-phase HPLC of synthetic extended H5 polypeptide after purification.
Figure 4.8. Analytical reverse-phase HPLC of synthetic tetrameric H5 polypeptide after purification.
Figure 4.9. Analytical reverse-phase HPLC of synthetic S4-S5-H5-S6 polypeptide after purification.
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Actual</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>1.85</td>
<td>2</td>
</tr>
<tr>
<td>Ala</td>
<td>2.01</td>
<td>2</td>
</tr>
<tr>
<td>Phe</td>
<td>0.87</td>
<td>1</td>
</tr>
<tr>
<td>Trp*</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Val</td>
<td>3.00</td>
<td>3</td>
</tr>
<tr>
<td>Thr</td>
<td>3.75</td>
<td>4</td>
</tr>
<tr>
<td>Met*</td>
<td>1.65</td>
<td>2</td>
</tr>
<tr>
<td>Gly</td>
<td>2.01</td>
<td>2</td>
</tr>
<tr>
<td>Tyr</td>
<td>1.03</td>
<td>1</td>
</tr>
</tbody>
</table>

* Met partially destroyed by hydrolysis. Tryptophan completely destroyed by hydrolysis.

Table 4.1. Amino Acid Analysis Of Synthetic H5 Polypeptide.
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Actual</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu</td>
<td>1.95</td>
<td>2</td>
</tr>
<tr>
<td>Ala</td>
<td>3.01</td>
<td>3</td>
</tr>
<tr>
<td>Gly</td>
<td>4.75</td>
<td>5</td>
</tr>
<tr>
<td>Ser*</td>
<td>2.56</td>
<td>3</td>
</tr>
<tr>
<td>Asx</td>
<td>3.01</td>
<td>1</td>
</tr>
<tr>
<td>Phe</td>
<td>4.01</td>
<td>4</td>
</tr>
<tr>
<td>Lys</td>
<td>1.85</td>
<td>2</td>
</tr>
<tr>
<td>Ile</td>
<td>0.82</td>
<td>1</td>
</tr>
<tr>
<td>Pro</td>
<td>1.95</td>
<td>2</td>
</tr>
<tr>
<td>Trp</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>Val</td>
<td>4.03</td>
<td>4</td>
</tr>
<tr>
<td>Thr</td>
<td>3.85</td>
<td>4</td>
</tr>
<tr>
<td>Met*</td>
<td>1.65</td>
<td>2</td>
</tr>
<tr>
<td>Tyr</td>
<td>0.95</td>
<td>1</td>
</tr>
</tbody>
</table>

* Ser and Met partially destroyed by hydrolysis.

Table 4.2. Amino Acid Analysis Of Synthetic Extended H5 Polypeptide.
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Actual</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>7.97</td>
<td>8</td>
</tr>
<tr>
<td>Ala</td>
<td>8.03</td>
<td>8</td>
</tr>
<tr>
<td>Phe</td>
<td>3.85</td>
<td>4</td>
</tr>
<tr>
<td>Trp*</td>
<td>—</td>
<td>8</td>
</tr>
<tr>
<td>Val</td>
<td>12.01</td>
<td>12</td>
</tr>
<tr>
<td>Thr</td>
<td>15.87</td>
<td>16</td>
</tr>
<tr>
<td>Met*</td>
<td>7.01</td>
<td>8</td>
</tr>
<tr>
<td>Gly</td>
<td>8.00</td>
<td>8</td>
</tr>
<tr>
<td>Tyr</td>
<td>4.01</td>
<td>4</td>
</tr>
<tr>
<td>Pro</td>
<td>7.98</td>
<td>8</td>
</tr>
<tr>
<td>Lys</td>
<td>3.00</td>
<td>3</td>
</tr>
</tbody>
</table>

* Met partially destroyed by hydrolysis. Trp completely destroyed by hydrolysis.

Table 4.3. Amino Acid Analysis Of Synthetic Tetrameric H5 Polypeptide.
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Actual</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>2.85</td>
<td>3</td>
</tr>
<tr>
<td>Leu</td>
<td>14.01</td>
<td>14</td>
</tr>
<tr>
<td>Ile</td>
<td>10.97</td>
<td>11</td>
</tr>
<tr>
<td>Arg</td>
<td>7.03</td>
<td>7</td>
</tr>
<tr>
<td>Val</td>
<td>15.00</td>
<td>15</td>
</tr>
<tr>
<td>Phe</td>
<td>10.95</td>
<td>11</td>
</tr>
<tr>
<td>Lys</td>
<td>5.00</td>
<td>5</td>
</tr>
<tr>
<td>Ser*</td>
<td>8.02</td>
<td>9</td>
</tr>
<tr>
<td>His</td>
<td>1.03</td>
<td>1</td>
</tr>
<tr>
<td>Gly</td>
<td>10.98</td>
<td>11</td>
</tr>
<tr>
<td>Glx</td>
<td>3.97</td>
<td>4</td>
</tr>
<tr>
<td>Ala</td>
<td>8.85</td>
<td>9</td>
</tr>
<tr>
<td>Met*</td>
<td>2.56</td>
<td>3</td>
</tr>
<tr>
<td>Pro</td>
<td>4.00</td>
<td>4</td>
</tr>
<tr>
<td>Trp*</td>
<td>—</td>
<td>3</td>
</tr>
<tr>
<td>Tyr</td>
<td>1.97</td>
<td>2</td>
</tr>
<tr>
<td>Cys</td>
<td>0.87</td>
<td>1</td>
</tr>
</tbody>
</table>

*Ser, and Met partially destroyed by hydrolysis. Trp completely destroyed by hydrolysis.

Table 4.4. Amino Acid Analysis Of Synthetic S4-S5-H5-S6 Polypeptide.
Figure 4.10. Laser-desorption mass spectra of monomeric synthetic H5 polypeptide.
Figure 4.11. Laser-desorption mass spectra of synthetic extended H5 polypeptide.
Figure 4.12. Laser-desorption mass spectra of tetrameric synthetic H5 polypeptide.
4.3.2 Structure of the synthetic monomeric H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer

The FTIR spectra of the monomeric H5 polypeptide incorporated into DMPC vesicles are presented (Fig. 4.13 and 4.14). The prominent band around 1734 cm$^{-1}$ in the absorption spectrum arises from vibrations of the carbonyl ester groups of DMPC and the spectrum in the amide I region shows a maximum at 1651 cm$^{-1}$ and the band shape is quite symmetrical with no other distinct features apparent. This band frequency indicates that the monomeric H5 when incorporated into DMPC vesicles adopts a predominantly an $\alpha$-helical structure (Haris et al., 1986). Detailed analysis using the second-derivative procedure (Fig. 4.14) does not reveal any further major bands in the amide I region except for a strong band at 1650 cm$^{-1}$ and a weak component at 1671 cm$^{-1}$. This strong band position indicates the presence of $\alpha$-helix with the weak band assigned to $\beta$-turn structures in the polypeptide structure (Haris et al., 1986). As it is of weak intensity, contributions from amino acid side chain groups cannot be completely ruled out. The absence of any bands in the region 1625-1635 cm$^{-1}$ indicates that the polypeptide does not contain any $\beta$-sheet structures. Hence the FTIR spectrum displays predominantly an $\alpha$-helical arrangement when incorporated into DMPC vesicles.
Figure 4.13. The FTIR absorption spectrum of synthetic monomeric H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30° C.
Figure 4.14. The second-derivative spectrum of synthetic monomeric H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30°C.
4.3.3 Structure of the synthetic monomeric H5 polypeptide in LPC micelles in $^2$H$_2$O buffer

The FTIR and CD spectra of the monomeric H5 polypeptide in LPC micelles are presented (Figures 4.15, 4.16 and 4.17). The absorption spectrum shows predominant bands at 1725 cm$^{-1}$ and in the amide I region a band at 1651 cm$^{-1}$. The strong band at 1731 cm$^{-1}$ arises from the vibration of the phospholipid ester. Second-derivative analysis shows absence of any major bands to indicate the presence of $\beta$-sheet type structures. The 1650 cm$^{-1}$ band in the amide I region indicates a predominantly $\alpha$-helical structure accompanied by weak components at 1673 cm$^{-1}$ and 1614 cm$^{-1}$. These bands are assigned to turns and $\beta$-sheet arrangement. These components are so weak in appearance that they can sometimes be assigned to contributions from amino acid side chains. Hence the polypeptide adopts predominantly an $\alpha$-helical arrangement in micelles.

The CD spectrum of the monomeric H5 polypeptide in LPC micelles is shown Figure 4.17. The spectrum shows a minimum at 207 nm and another at 220 nm. This double minimum is characteristic of proteins and polypeptides with an $\alpha$-helical structure. Thus, both the CD and FTIR data are consistent in indicating that the H5 polypeptide mainly adopts an $\alpha$-helical arrangement in membrane mimetic environments.
Figure 4.15. The FTIR absorption spectrum of synthetic monomeric H5 polypeptide in LPC micelles in $^2$H$_2$O buffer at 30° C.
Figure 4.16. The second-derivative spectrum of synthetic monomeric H5 polypeptide in LPC micelles in $^2$H$_2$O buffer at 30° C.
Figure 4.17. CD spectrum of synthetic H5 polypeptide in LPC micelles in $^2$H$_2$O buffer.
4.3.4 Structure of the synthetic monomeric H5 polypeptide in DMPG vesicles in $^2$H$_2$O buffer

The FTIR spectra of the monomeric H5 in DMPG vesicles is presented (Fig. 4.18 and 4.19). The amide I maximum is centred at 1651 cm$^{-1}$ in the absorption spectrum, which is assigned to an $\alpha$-helical arrangement. The second derivative analysis does not reveal any further major bands in the amide I region, although there is a weak component at 1673 cm$^{-1}$ (Fig. 4.19). This may reflect the presence of $\beta$-turns in the polypeptide conformation. As this band is weak in its appearance, it is possible some contribution from side chain groups of amino acids could give rise to this band position. However, the monomeric H5 adopts a predominantly $\alpha$-helical structure in an anionic lipid environment such as DMPG vesicles. These results are thus consistent with H5 adopting predominantly $\alpha$-helical conformations when dispersed in lipid assemblies.

4.3.5 Structure of the synthetic extended H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer

The FTIR spectra of the extended H5 polypeptide in DMPC vesicles is depicted (Fig. 4.20 and 4.21). The absorption spectrum shows a band maximum at 1654 cm$^{-1}$ in the amide I region. The band appears symmetrical and there are no indications of any other major components. The amide I maximum at 1654 cm$^{-1}$ can be assigned to an $\alpha$-helical conformation. Further analysis applying the second-derivative method reveals only very weak components in the regions 1680-1674 cm$^{-1}$ and 1631-1615 cm$^{-1}$.
Figure 4.18. The FTIR absorption spectrum of synthetic monomeric H5 polypeptide in DMPG vesicles in $^2$H$_2$O buffer at 30° C.
Figure 4.19. The second-derivative spectrum of synthetic monomeric H5 polypeptide in DMPG vesicles in $^2$H$_2$O buffer at 30° C.
These components can be ascribed to turns and $\beta$-sheet structures respectively. However contributions from amino acid side chains cannot be ruled out. The preponderance of an $\alpha$-helical arrangement of the extended H5 polypeptide in DMPC vesicles is again evident.

4.3.6 Structure of the synthetic tetrameric H5 polypeptide in DMPC micelles in $^2$H$_2$O buffer

Figures 4.22 and 4.23 show the spectra of the synthetic tetrameric H5 polypeptide in DMPC micelles. The absorption spectrum shows the amide I band centred at 1649 cm$^{-1}$ which indicates an $\alpha$-helical structure. Further detailed analysis using the second-derivative method reveals the main amide I band centred at 1649 cm$^{-1}$. This is consistent with the presence of an $\alpha$-helical arrangement revealing only minor components at 1671 cm$^{-1}$ and 1615 cm$^{-1}$. These minor band components may emanate from turns and amino acid side chains respectively. Hence the tetrameric H5 polypeptide also has the tendency to adopt a predominantly an $\alpha$-helical arrangement when in phospholipid vesicles.
Figure 4.20. The FTIR absorption spectrum of synthetic extended H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30° C.
Figure 4.21. The second-derivative spectrum of synthetic extended H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30° C.
Figure 4.22. The FTIR absorption spectrum of synthetic tetrameric H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30° C.
Figure 4.23. The second-derivative spectrum of synthetic tetrameric H5 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30°C.
4.3.7 Structure of the synthetic 119-residue polypeptide incorporating the transmembrane S4-S5-H5-S6 domains in DMPC vesicles in \( \text{H}_2\text{O} \) buffer

Figures 4.24 and 4.25 show the FTIR spectra of the synthetic 119 amino acid long polypeptide containing the transmembrane domains. The absorption spectrum (Fig. 4.24) shows a band maximum at 1654 cm\(^{-1}\) and the band shape appears quite symmetrical with no other major distinct features apparent. This band frequency indicates that the polypeptide in DMPC vesicles primarily adopts an \( \alpha \)-helical conformation. Detailed structural analysis using the second-derivative procedure (Fig. 4.25) does not reveal any major bands to indicate the presence of \( \beta \)-sheet arrangement in the amide I region. However, weak components at 1678 cm\(^{-1}\) and 1631 cm\(^{-1}\) may reflect the presence of turns and \( \beta \)-sheet structures. As these band components are of weak intensity, contributions from amino acid side chains cannot be discounted. Hence the 119 amino acid long polypeptide encompassing the transmembrane domains tends to adopt a predominantly \( \alpha \)-helical structure whilst incorporated in DMPC vesicles.
Figure 4.24. The FTIR absorption spectrum of synthetic 119-residue S4-S5-H5-S6 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30° C.
Figure 4.25. The second-derivative spectrum of synthetic 119-residue S4-S5-H5-S6 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30° C.
4.3.8 Planar bilayer measurements

The extreme hydrophobic nature of the monomeric H5 polypeptide in aqueous solution entailed its incorporation by direct addition to the lipid solution with the use of PC/CHL as the bilayer-forming phospholipid. This procedure has proved successful for several very hydrophobic channel-forming polypeptides (Oiki et al., 1988). A relatively low molar ratio (1:100) of polypeptide : lipid rules out any channel activity observed being attributed to detergent-like action of the polypeptide or formation of non-bilayer lipid phases. A 5s section from a recording made from H5 monomeric-containing bilayer is shown (Fig. 4.26). For the first 4s or thereabouts, the channel switches rapidly between the closed state (C) and the first open level (O1). The channel then opens to (O1) and occasional transitions to a second open level (O2) are observed. An expanded segment (length 200ms) from the recording is shown (Fig. 4.26b). It can be seen that the channel switches between C and O1 modes. Channel openings are of duration 5-10 ms and of conductance 30 pS. Overall, a range of conductance levels from 12 to 50 pS were observed for bilayers incorporating monomeric H5.

The hydrophobic nature of tetrameric H5 polypeptide in TFA allowed its incorporation into preformed bilayers via the aqueous phase. In these experiments the diPhyPC was employed for the bilayer formation, as in previous studies it was found to generate very low background noise level. Control experiments have shown that 1.6% TFA (the final concentration in planar bilayer experiments) had no effect on stability of diPhyPC bilayers in the presence of 200 mM BES.
Figure 4.26. (a) and (b) Channel formation by monomeric H5 peptide. Applied transmembrane voltage of 140 mV. Positive (upward) deflections indicate channel openings. (b) A segment of the recording in (a). (c) and (d) Channel formation by H5 tetramer.
A 5s segment from a recording made from a diPhyPC bilayer exposed ~ 0.4μM tetramer is shown (Fig. 4.26d). There was a lag of ~30 minutes between addition, with stirring, of the polypeptide to the aqueous phase and appearance channel activity. There are occasional brief openings to an open state (O). A 200ms expanded segment is depicted (Fig.4.26d). The channel switches from a closed state (C) to an open state. The open state is of brief duration (5-10 ms) and has conductance of 13 pS.

4.3.9 Molecular modelling and analysis of amphipathicity

The FTIR and CD results indicate that H5 polypeptide adopts mainly an α-helical arrangement in a membrane mimetic environment. Information on possible side chain arrangement cannot be provided by these spectroscopic techniques. Simulated annealing/molecular dynamics were used to generate an ensemble of 25 structures for H5 in an α-helical conformation. Five models selected at random from the final 25 are shown Figure 4.27. Some heterogeneity in side chain conformations is detected on comparing the 25 H5 models whilst the helical backbone geometry is conserved. In particular, the aromatic side chains of F3, W4, W5 and Y15 exhibit multiple conformations, corresponding to "flipping" of the rings. Analysis using Procheck (Morris et al., 1992) confirmed that the stereochemistry of the models was consistent with that found in highly refined crystallographic structures of proteins. The overall r.m.s.d. between the 25 models was 1.7Å for all non-H atoms and 0.9Å for backbone atoms. The residue by residue r.m.s.d.s. were larger at the helix termini (~1.5Å) than in the centre (~0.5Å), as has been observed in several
Figure 4.27. Modelling the H5 helix by simulated annealing together with molecular hydrophobicity potential analysis.
molecular dynamics studies of α-helices (Levy et al., 1982). The results of MHP analysis of the H5 models are given (Fig. 4.27c) and (Fig. 4.27d), in the form of plots of $<E^{\min}>$ versus $\phi$ for the five selected models and the corresponding plot averaged over the five models. This analysis allows a hydrophilic face (from $\phi = -50$ to $+100^\circ$) and a hydrophobic face (from $\phi = +100$ to $-50^\circ$) to be defined. The hydrophilic face is defined by residues T9 and T12 and the hydrophobic face is defined by residues F3, W4 and Y15 (Fig. 4.27d). This implies that, averaging over five different sets of side chain conformations, the H5 polypeptide forms a moderately amphipathic helix. On the basis of current models of channel formation by amphipathic helical peptides (Sansom, 1991), it can be inferred that the channels formed in planar bilayers by monomeric H5 helices oriented such that their hydrophilic residues line a central ion-permeable pore.

In the presence of phospholipids, the helical conformation and stabilization of the H5 polypeptide can be explained by its amphipathicity. Several studies have indicated that channel-forming and related polypeptides may be induced to adopt amphipathic helices at the water-phospholipid interfaces. CD and NMR studies of amphipathic helices bound to phospholipid bilayers (Vogel, 1987; Demsey and Butler, 1992) suggest that such helices may lie with the helix axis in the plane of the bilayer i.e. perpendicular to the bilayer normal. In such an orientation the hydrophobic face of the helix interacts with the fatty acyl chains of the lipid molecules, whilst the hydrophilic face forms favourable interactions with the lipid head groups and aqueous interface. In the presence of a transbilayer electrostatic field, the polypeptide may then undergo
reorientation causing the membrane-spanning helices to aggregate, thus forming ion channels.

4.4 DISCUSSION

Based on mutagenesis studies of K⁺ channels and pharmacological manipulations, some of the structural elements involved in channel activation, inactivation, and ion permeation have been identified (Hartman et al., 1991). The voltage-gated K⁺ channels are postulated to possess six transmembrane segments designated S1 to S6. However, recent studies (Hartman et al) have revealed a linker region (designated as SS1-SS2 or H5) linking the S5 and S6 domains as the pore-forming sequence. Employing site-directed mutagenesis experiments on the Shaker H5 region and substituting the H5 regions of two different potassium channels (drk1 with the Ngk2 sequence) have indicated that this region determines the conductance and ion selectivity properties of the potassium channel. Previously it was proposed that this region formed the extracellular loop. The experimental evidence that proposes that the H5 sequence spans the membrane has come from interaction of the potassium channel blocker, TEA ion, when applied to either the extracellular or intracellular side of the potassium channel, when inhibition of ion conductance was observed. Mutation applied to residues 1 and 19 of H5 (numbering scheme refers to the peptide sequence given (Fig. 5.1) affects external but not internal blockade by TEA. In contrast to this, mutation at residue 11 affected only internal and not external TEA occlusion. Based on these findings it was interpreted that residues near either end of H5 were at the extracellular mouth.
of the channel and residues in the middle H5 to be in the vicinity of the intracellular mouth. This interpretation assumes that the effects of single point mutations are exerted directly and allosteric effects on the protein conformation are not considered. However a model has been postulated on the basis of site-directed mutagenesis studies in which the H5 segment (residues 431-449 of the K⁺ channel) forms a β-hairpin structure. The co-assembly of four such hairpins have been proposed for the formation of a channel, and detailed models have been constructed. Despite the advancement in molecular and pharmacological studies, there are as yet no experimental structural data available to support these modelling studies based on the proposal of the β-sheet arrangement.

The biophysical studies obtained using FTIR and CD spectroscopic indicate that the monomeric, extended, tetrameric H5 and the S4-S5-H5-S6 synthetic polypeptides all adopt predominantly an α-helical structure in membrane mimetic environments. The planar bilayer studies indicate that H5 polypeptides can form channels with conductances in the range 12-50 pS, which should be compared with a conductance range 5-30 pS for physiological K⁺ channels (Hille et al., 1992). This suggests that an α-helical arrangement of H5 may be of significant relevance to its channel-forming properties. Simulated annealing/molecular dynamics has also been applied to generate an ensemble of α-helical models of monomeric H5, so as to further understand channel formation by this synthetic polypeptide. Amphipathicity analysis of these models helps to explain the helical conformation of H5, observed by spectroscopic measurements, when in the presence of lipid bilayers and
suggests a possible role for H5 helices as a component of the intact channel protein. The present data together with the mutagenesis results gives rise to three possible interpretations. The first is that the H5 polypeptides are \( \alpha \)-helical as an isolated domain but forms a \( \beta \)-hairpin only when part of the total protein structure. A second interpretation is that the H5 polypeptide is in fact helical in the intact channel. This interpretation seems difficult to resolve with the mutagenesis results and also would entail modifications to current topological models of \( K^+ \) and related channels in order for the N- and C-terminal extramembraneous domains to remain on the intracellular face of the membrane. However, as discussed above, it is possible that the point mutations when expressed in host systems do not exert their effects directly. Furthermore, a definitive topology of the \( K^+ \) channel has yet to be established experimentally. In the light of the compelling spectroscopic data and given the results of the bilayer studies, it does seem logical to propose that in the intact channel the pore is formed by coassembly of four H5 helices. A third interpretation would be that H5 undergoes a conformational transition, from an \( \alpha \)-helix structure to a \( \beta \)-hairpin arrangement, upon activation of the channel. This is similar in spirit to the model for \( Na^+ \) channel activation proposed (Guy and Conti, 1990). In the closed channel, H5 polypeptide would form an amphipathic helix, as in the structural and bilayer studies. Subsequent to activation of the channel, the H5 polypeptide would then form a \( \beta \)-hairpin structure. While this model is speculative, in essence it reconciles both structural data of H5 in vitro and mutagenesis data on H5 in situ whilst also stressing the essentially dynamic nature of ion channel structures. However,
as yet there is no evidence for such an \( \alpha \) to \( \beta \) transition characteristics for the \( K^+ \) channel activation. In this study, by using chemically synthesized polypeptides corresponding to the pore, together with the transmembrane domains, it still shows a predominantly \( \alpha \)-helical structure when incorporated into membrane mimetic environments. The H5 and its derivatives have also demonstrated ion conductivity characteristics when incorporated into lipid membrane systems.

4.5 SUMMARY

Recent mutagenesis has located a stretch of amino acid sequence which constitutes the ion-selective pore of the voltage sensitive potassium channel. It has been postulated that this stretch of sequence adopts a \( \beta \)-barrel arrangement in the lipid bilayer. Polypeptides corresponding to this amino sequence (residues 431-449) of the A-type potassium channel protein from *Drosophila* have been chemically synthesized. A tetrameric version of this sequence and a polypeptide of 119-residues containing the transmembrane domains S4-S5-H5-S6 were also synthesized. FTIR and CD spectroscopy of the H5 indicate a predominantly \( \alpha \)-helical structure when incorporated into membrane mimétic enviroments. There is no evidence of the prominence of \( \beta \)-sheet like arrangement of the various representations of the pore forming domain. Ion channel activity was detected when the monomeric and tetrameric H5 were incorporated into planar lipid bilayers. Molecular modelling studies support the accomodation of an \( \alpha \)-helical arrangement of the H5 within the lipid bilayer.
CHAPTER FIVE

CONFORMATIONAL ANALYSIS OF A SYNTHETIC POLYPEPTIDE
CORRESPONDING TO THE VOLTAGE-SENSING S4 DOMAIN OF A VOLTAGE-
GATED POTASSIUM CHANNEL PROTEIN
5.1 INTRODUCTION

Voltage-gated ion channel proteins have the ability to respond to fluctuations to transmembrane voltage and this constitutes one of the basic mechanisms underlying electrical excitability of nerve and muscle membranes. The voltage-gated potassium channel is an integral membrane protein which plays a fundamental role in the generation and propagation of action potentials in most multicellular organisms. Sequences of potassium channel proteins have recently been accessible for various species (Perney and Kaczmarek 1991). Analysis of the sequence for hydrophobicity assumes six putative membrane-spanning segments (Catterall, 1988). The stretches of hydrophobic residues in the core region of the K\(^+\) channel polypeptide are depicted schematically in Figure 5.1. The fourth of these segments shows remarkable homology with the S4 segments of the voltage-activated Na\(^+\) and Ca\(^{2+}\) channels (Noda, 1984; Tanabe, 1987; Papazian et al., 1991). The overall arrangement of the postulated transmembrane segments of the K\(^+\) channel, is similar to that of each of the four homologous transmembrane domains of the Na\(^+\) (Noda et al., 1986) and Ca\(^{2+}\) channel alpha subunit (Tanabe et al., 1987). The S4 segment is highly conserved in sodium channels of different species and are also evident in transmembrane domains of other voltage-sensitive ion channels.

The S4 sequence is unusual in that it contains basic (positively charged) residues at every third or fourth position, (Fig. 5.2) interspersed with generally hydrophobic residues, and has been located in each member of the four K\(^+\) channel subfamilies and in each of the domains of the Ca\(^{2+}\) and Na\(^+\) channels alpha subunits (Noda et al., 1986; Tanabe et al., 1987). Because of its unique
composition, it was proposed to be membrane-spanning so that positively charged residues are located in the hydrophobic interior of the membrane and function as the channel's voltage sensor (Catteral et al., 1986; Greenblatt et al., 1985). It has also been suggested that the basic residues are ion paired with the acidic residues on neighbouring transmembrane helices. Site-directed mutagenesis studies, where uncharged polar residues were substituted one at a time, revealed alteration in the voltage-dependent activation process (Papazian et al., 1991). A similar wide range mutant effect was also observed in mutagenesis studies of the S4 basic residues of the Na⁺ channel (Stühmer et al., 1989). These studies indicate that these basic residues in the S4 region play a central role in the voltage-dependence of activation of K⁺ channels.

The mechanism by which the S4 segment functions as a voltage sensor is, however, unclear. Although mutagenesis studies have provided valuable information on the possible functional role of the S4 sequence, little is known about its structure. Recently, some information has been directed towards obtaining a structural understanding of the S4 sequence. This was achieved by 2D NMR spectroscopy on the solution structure of the S4 sequence corresponding to the rat brain Na⁺ channel (Mulvey et al., 1989). The ion conductance property of a synthetic Na⁺ channel S4 peptide in planar lipid bilayers has also been investigated (Tosteson et al., 1989). Various models have been proposed to explain the charged movements that occur in the "gating-currents" observed upon sodium channel activation.
Figure 5.1. Putative transmembrane folding scheme of voltage dependent $K^+$ channel showing the S4 voltage sensor (red). The S4 sequence of Shaker $K^+$ channel is shown below the folding scheme.
Figure 5.2. Sequence of the synthetic S4 polypeptide.

\[ \text{CH}_3-\text{ILRVIRLVRVFRIFKLSRHS-NH}_2 \]
The "sliding helix" (Catterall 1986) or "Helical screw" (Guy & Seetharamulu) models propose that, at rest, all the positive charges on the S4 are stabilized by the negative charges from surrounding transmembrane α-helical segments. Membrane depolarization causes the S4 helix to slide outward into the extracellular membrane surface. This sliding motion imposes a 60° rotation which causes a lateral movement of 4.5Å. Other models favour conformational changes of the arginine side chains as the prime mover of the charges. In all likelihood, a combination of conformational changes of the arginine side chain groups and partial movement of S4 helices are probably involved in activation (Guy, 1988). Another recent model proposed as the propagating helix, also has a possible explanation for the channel-activating mechanism of S4. In this model, mainly the middle section of the S4 adopts an α-helical structure when in a closed state, while the N-terminus (external) and the C-terminus (internal) adopts a β-sheet arrangement. Upon activation, the helical section of the S4 propagates in the direction of the N-terminus approaching closer to the extracellular surface, whereby the central region of the channel comes apart. While limited structural data are available for the S4 sequence of the Na⁺ channel, little or no structural data are available for the corresponding sequence of the K⁺ channel. In this chapter, a chemically synthesized polypeptide corresponding to the S4 sequence of the voltage-sensitive potassium channel is examined in various environments. FTIR and CD spectroscopy have been employed to study the structure in aqueous solution, in TFE solvent, and also after reconstitution into lipid vesicles and micelles. This study was undertaken in collaboration with Dr. P.I. Haris (this department).
5.2 MATERIALS AND METHODS

DMPC, LPC, TFE, acetic anhydride and $^3$H$_2$O were purchased from Sigma Chemicals Ltd. (U.K.). Fmoc-protected amino acids, Rink Amide MBHA resin, TBTU, ODHB, and HBT were purchased from Calbiochem-Novabiochem (UK) Ltd. Solvents used in the synthesis were of peptide synthesis grade and were obtained from Rathburn Ltd. (Scotland). All other reagents were of analytical grade. All buffers were prepared using Milli Q purified water.

5.2.1 Assembly of the S4 peptidyl-resin sequence: Fmoc-I-L-R-V-I-R-L-V-R-V-F-R-I-F-K-L-S-R-H-S-Rink-amide-MBHA-resin

The S4 polypeptide was synthesized using the stepwise solid-phase procedure using the Fmoc protecting strategy. The Rink-Amide-MBHA-resin was placed in a sealed glass reaction vessel and acylation of the first amino acid as active ester generated \textit{in-situ} using TBTU was carried out. A 5-fold excess (based on the resin loading) of Fmoc-amino acid was dissolved with the following equivalents: Amino acid:TBTU:NMP (1:1:2) in DMF was used in all subsequent couplings. All the arginines in the assembly were double coupled to ensure efficient coupling. The Fmoc arginines had their side chains protected with 2,2,5,7,8-Pentamethylchroman-6-sulphonyl groups (Pmc). The synthesis was performed using the batch synthesis method on an automated peptide synthesizer Rainin PS3 (Protein Technologies, USA) with modified protocols and scale (0.1mmol theoretical yield of crude peptide). The cycle for addition of Fmoc-amino acid consisted of a 5x2 minutes wash of the solid support with 20% piperidine in DMF to deprotect the N°-Fmoc-group, a 12 minute DMF
wash, a 25 minute acylation reaction with activated protected amino acids, and an 8 minute DMF wash for a total cycle time of about 60 minutes.

5.2.2 Detachment and deprotection of the synthetic S4 polypeptide

On completion of the synthesis of the S4 polypeptide, the N-terminal N\textsuperscript{\textregistered}-Fmoc protective groups were removed using 20% piperidine in DMF. The deprotected N-termini of the S4 polypeptide was then blocked by acetylation. This was effected by addition of Ac\textsubscript{2}O to the aminopeptidyl-resin in the reaction vessel. The acetylating mixture was agitated for 30 minutes at room temperature. After completion of acetylation the polystyrene beads were washed with DMF (5x1 minute) and finally washed with DCM (5mlx1 minute) and dried \textit{in vacuo} for 5 hours. The dried amide resin was then agitated with dry nitrogen in the presence of 10% TFA in DCM (cleavage solution) in a sintered glass column. At intervals of 10 minutes the cleavage solution was collected into a round bottomed glass flask. This procedure was carried out for 1 hour when detachment is normally complete.

At this stage, the TFA and DCM were completely removed by rotary evaporation where the protected peptide amide are concentrated. Finally to remove all the protecting groups still present on the polypeptide amide, 82% TFA in the presence of the scavengers, phenol (5%), thioanisole (5%) and ethanedithiol (2.5%) were added to the residual oily solution at room temperature with occasional swirling for 4 hours. At the end of 4 hours, the TFA together with the scavengers were removed by rotary evaporation at 50°C. The crude products were precipitated and rinsed with cold anhydrous
diethyl ether (5x2 minutes). The precipitated S4 polypeptide was finally dried under high vacuum in a freeze drier.

5.2.3 Purification of the synthetic S4 polypeptide

The dried crude S4 synthetic polypeptide was resuspended in 0.1% TFA in water and purified by multiple reverse-phase HPLC runs. The S4 polypeptide was injected into a C\textsubscript{18} (semi-preparative) TSK ODS 120T column equilibrated in 100% buffer A (deionized/Milli Q water containing 0.1% HPLC-grade TFA) and 0% buffer B [90%(v/v) acetonitrile in water containing 0.1% TFA]. The polypeptide was purified on a linear gradient of buffer B (from 0% to 70% in 30 minutes) at a flow rate of 2.5 ml/minute, when a major peak was observed. Homogeneity was assessed by a second HPLC analysis on a narrow-bore Vydac C\textsubscript{18} (analytical) 201 TP54 RP column equilibrated as described above.

5.2.4 Amino acid compositional and molecular mass analysis

The freeze dried HPLC purified S4 polypeptide was hydrolysed with 6M constant boiling HCl containing 0.1% phenol at 110°C for 20-70 hours. After hydrolysis, residual HCl was removed under vacuum with the aid of a freeze drier. The dried hydrolysed sample was then resuspended in amino acid sample buffer analysed using the amino acid analyser (LKB Biochrom 4151 Alpha Plus amino acid analyser). Samples in duplicates were prepared and calibrated using 10 nmol standards. The identity and purity of the S4 was further confirmed by laser desorption mass spectrometer (LASERMAT Mass Spectrometer).
5.2.5 Reconstitution of the synthetic S4 polypeptide in 98% TFE, DMPC vesicles and LPC micelles in $^2$H$_2$O buffer

Reconstitution of the synthetic S4 polypeptide involved applying the new disaggregation procedure which was shown to be reliable in earlier studies involving the Pf1 coat protein in chapter 3. Typically, the S4 polypeptide was dissolved in neat TFA after which the TFA was completely removed by rotary evaporation and further dried for 24 hours under high vacuum. To further remove trace amounts of TFA from the sample, it was lyophilized with HCL (0.1M), for 24 hours. Appropriate lipids or detergents in DCM were overlayed onto the polypeptide film and the DCM rotary evaporated. The mixture was then subjected to further drying under high vacuum for 24 hours. Finally for FTIR studies, $^2$H$_2$O PBS (pD 7.4) was added to the dried mixture and vortexed for 5 minutes at room temperature. Similarly for studies in aqueous TFE and buffer, a similar procedure was adopted except that the buffer was added to the polypeptide film instead of lipid solution.

5.2.6 FTIR and CD spectroscopy

FTIR spectra were recorded as described in section 3.2.9. CD spectra of S4 polypeptide in LPC micelles and PBS were recorded on a JASCO J-600 spectrometer. The spectra were recorded using 0.01cm pathlength cell. Spectra of aqueous dispersions of the phospholipid micelles were subtracted from the peptide spectra to remove absorbances contributed by phospholipids. The ratio of polypeptide to phospholipid used was the same as that used for the FTIR measurement. The new reconstitution procedure was
applied and the sample was diluted such that a peptide concentration of 1.25 mg/ml was obtained. Spectra were corrected for concentration and ΔE values calculated.

5.3 RESULTS

5.3.1 Chemical and physical analysis of the synthetic S4 polypeptide

The polypeptide was subjected to HPLC purification (Fig. 5.3) and homogeneity confirmed (Fig. 5.4). The S4 polypeptide eluted as a well-resolved peak with no trace of any major contaminating peptides. On the basis of the HPLC analysis the synthesis can be considered as a success. The HPLC purified polypeptide was also identified by their amino acid composition (Table 5.1). The HPLC peak could be accounted for as the correct synthesized product. The S4 polypeptide shows a calculated molecular mass of 2548.60 which is in good agreement with that determined by mass spectrometry (fig 5.5).
Figure 5.3. Preparative scale reverse-phase HPLC chromatogram of crude synthetic peptide corresponding to S4 domain of the Shaker K⁺ channel protein.
Figure 5.4 Analytical reverse-phase HPLC of synthetic peptide corresponding to the S4 domain of Shaker K\(^+\) channel protein after purification.
Figure 5.5. Laser-desorption mass spectra of synthetic 20-residue S4 domain of Shaker K⁺ channel protein.
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Actual</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ile</td>
<td>3.05</td>
<td>3</td>
</tr>
<tr>
<td>Leu</td>
<td>2.95</td>
<td>3</td>
</tr>
<tr>
<td>Arg</td>
<td>4.85</td>
<td>5</td>
</tr>
<tr>
<td>Val</td>
<td>2.85</td>
<td>3</td>
</tr>
<tr>
<td>Phe</td>
<td>1.89</td>
<td>2</td>
</tr>
<tr>
<td>Lys</td>
<td>1.01</td>
<td>1</td>
</tr>
<tr>
<td>Ser*</td>
<td>1.65</td>
<td>2</td>
</tr>
<tr>
<td>His</td>
<td>1.00</td>
<td>1</td>
</tr>
</tbody>
</table>

* Ser partially destroyed by hydrolysis.

Table 5.1. Amino Acid Analysis Of Synthetic S4 Polypeptide.
5.3.2 Conformation of synthetic S4 polypeptide in 98% TFE in $^2$H$_2$O buffer

Analysis of the Amide I band in the FTIR spectra of polypeptides and proteins impart useful information on their secondary structure (Haris and Chapman 1992). The relationship between the frequency of the amide I band and secondary structure is now well understood. The organic solvent TFE is known to promote and enhance intermolecular interactions and therefore has often been employed as an $\alpha$-helical structure-inducing and stabilizing medium. FTIR spectra of the S4 polypeptide were therefore recorded in TFE. The absorption spectrum (Fig. 5.6) shows a maximum centred at 1654 cm$^{-1}$ in the amide I region. This can be attributed to an $\alpha$-helical arrangement. Detailed analysis using second-derivative procedure (fig. 5.7) reveals the main amide I maximum at 1655 cm$^{-1}$. Weak components can be observed at 1634 cm$^{-1}$ and 1683 cm$^{-1}$. These bands are weak and may originate from some turn type structure although contributions from $\beta$-sheet structure may also occur. Contributions from amino acid side chains are also possible. The CD spectrum (not shown) in TFE displays double minima at 208 nm and 222nm which is also typical of polypeptides and proteins in an $\alpha$-helical conformation. The two biophysical measurements illustrate that the S4 polypeptide adopts an $\alpha$-helical arrangement when in TFE solvent. This result is also in very good agreement with the previous 2D NMR study of the S4 polypeptide corresponding to the rat brain sodium channel (Mulvey et al., 1989).
Figure 5.6. FTIR absorption spectrum of synthetic S4 polypeptide in 98% trifluoroethanol in $^2$H$_2$O at 30° C.
Figure 5.7. Second-derivative spectrum of synthetic S4 polypeptide in 98% trifluoroethanol in $^2$H$_2$O at 30° C.
5.3.3 Conformation of the synthetic S4 polypeptide in LPC micelles in $^{2}$H$_2$O buffer

The absorption spectrum of S4 in LPC vesicles (Fig. 5.8) displays a band maximum at 1650 cm$^{-1}$ in the amide I region. This band position can be attributed to an $\alpha$-helical conformation. Applying the second-derivative procedure (Fig. 5.9) for a detailed analysis, weak components at 1630 cm$^{-1}$ and 1670 cm$^{-1}$ are observed. These band positions may be contributed by $\beta$-sheet and turn structures. The main predominant band at 1650 cm$^{-1}$ in the amide I region is consistent with an $\alpha$-helical structure. Further evidence for the presence of the $\alpha$-helical structure in the LPC micelles is also supported by CD spectroscopy (Fig. 5.10). The spectrum is characterised by a bilobed display, with minima at 205 nm and 221 nm typical of a helical arrangement.
Figure 5.8. FTIR absorption spectrum of synthetic S4 polypeptide in LPC micelles in $^2$H$_2$O buffer at 30°C.
Figure 5.9. Second-derivative spectrum of synthetic S4 polypeptide in LPC micelles in $^2$H$_2$O at 30$^\circ$C.
Figure 5.10. CD spectra of synthetic S4 polypeptide in aqueous solution (red trace) and LPC micelles (black trace) in $^2$H$_2$O buffer.
5.3.4 Conformation of the synthetic S4 polypeptide in DMPC vesicles in $^{2}$H$_{2}$O buffer

Figure 5.11 shows the absorption spectrum of the S4 polypeptide in DMPC vesicles. A strong band at 1652 cm$^{-1}$ in the amide I region is revealed. This band position can be assigned to an $\alpha$-helical structure. Further analysis using the second-derivative (Fig. 5.12) reveals the main amide I band at 1653 cm$^{-1}$. This is consistent with the presence of $\alpha$-helical structure in DMPC vesicles. Hence the conformation in lipid environment is predominantly an $\alpha$-helical. Similar results were also obtained for the S4 polypeptide when reconstituted in DMPG vesicles.

5.3.5 Conformation of the synthetic S4 polypeptide in PBS (pD 7.4) in $^{2}$H$_{2}$O buffer

The FTIR absorption spectrum of the synthetic S4 polypeptide in PBS solution is shown (Fig. 5.13). The amide I band is predominantly centred at 1643 cm$^{-1}$. This band position is assigned to random coil structure. Detailed analysis of the amide I band applying the second-derivative procedure reveals only one more additional band at 1670 cm$^{-1}$ (Fig. 5.14). This band is of so weak appearance that it could arise from turns/bends within the polypeptide structure. The main amide I maximum however is 1640 cm$^{-1}$ can be attributed to random coil arrangement. This assignment is further supported by CD spectroscopy which shows a spectrum typical of unordered arrangement (Fig. 5.10).
Figure 5.11. FTIR absorption spectrum of synthetic S4 polypeptide in DMPC vesicles in $^2$H$_2$O buffer at 30°C.
Figure 5.12. Second-derivative spectrum of synthetic S4 polypeptide in DMPC vesicles in $^2$H$_2$O at $30^\circ$C.
Figure 5.13. FTIR absorption spectrum of synthetic S4 polypeptide in $^2$H$_2$O buffer at 30° C.
Figure 5.14. Second-derivative spectrum of synthetic S4 polypeptide in $^2$H$_2$O buffer at 30° C.
5.4 DISCUSSION

From mutagenesis studies it is evident that the S4 sequence plays a vital role in the conformational transition which accompanies gating (Pongs, 1992; Stuhmer et al., Perzo et al., 1994). While these studies have been invaluable in providing information regarding the possible functional role of the S4 sequence, there is as yet no hard structural data to explain the mechanism by which it achieves this important function. The lack of availability of sufficient amount of the voltage-gated ion channels has prevented structural studies using biophysical techniques. One practical approach towards obtaining some structural information on these large proteins is to chemically synthesize polypeptides corresponding to various domains of its primary structure which can then be analysed using biophysical techniques. Using this approach the S4 sequence of the Na\(^+\) channel was studied using NMR spectroscopy (Mulvey et al., 1989). However these studies were carried out using the organic solvent TFE a good promoter of helical type conformations. This approach has been adopted to the study of structure of the polypeptides corresponding to the ion-selective pore of voltage-gated potassium channels in various membrane mimetic environments (Haris et al 1994).

In the absence of structural data various models have been proposed as to how S4 might be involved in voltage-sensing (Guy and Seetharamulu 1986; Leuchtag, 1994). It has been postulated that the voltage-dependent conformational change involves a "helical screw" movement of the charged S4 segment towards the extracellular surface. The S4 is in its "innermost" position in the protein structure when the channel is in the D, deactivated state. The
movement of S4 from the D1 to the open conformation by such a helical screw mechanism have been described (Durell and Guy 1992). In the open conformation all the positively charged amino acids of the S4 are located on or near the extracellular surface, forming salt bridges with negatively charged residues from S1, S2, S3 and S5. The model as proposed purports that a substantial portion of the S4 segment moves outside the lipid bilayer in the open state. For this to occur, the polypeptide must possess conformational flexibility, and the transfer from a membrane environment to a hydrophilic aqueous environment probably involves some rearrangement of the polypeptide structure. The results obtained in this study clearly indicates such conformational transition for the synthetic S4 polypeptide. It has also been shown that in environments such as TFE solvent, LPC micelles and in DMPC vesicles a predominantly α-helical structure is adopted. However, upon transfer to an aqueous environment the polypeptide undergoes a major conformational rearrangement and adopts a random coil structure (Fig. 5.15 and 5.16).

The FTIR spectroscopic results indicate that the S4 polypeptide possess the conformational flexibility required for it to move from the lipid bilayer to the extracellular aqueous medium. There is in fact some evidence to indicate that S4 does get exposed to the extracellular medium. By mutating one positive charge (lysine) near the NH₂-terminal of S4, it has been shown to alter its influence of extracellular Ca²⁺ ions but not that of intracellular Mg²⁺ (Pusch, 1990). This information indicates that this charged amino acid is exposed to the extracellular medium. Further evidence is provided by way of antibodies directed against S4 sequence of the Repeat 1 in Na⁺ channels appear to bind.
Figure 5.15. FTIR absorption spectrum of synthetic S4 polypeptide in LPC micelles (dashed trace); in DMPC vesicles (red trace); in aqueous solution (bold trace) in $^2$H$_2$O buffer at 30°C.
Figure 5.16. Second-derivative spectrum of synthetic S4 polypeptide in LPC micelles (dashed trace) ; in DMPC vesicles (red trace) ; in aqueous solution (bold trace) in $^2$H$_2$O buffer at 30° C.
on the extracellular surface only when the membrane is depolarized (Schwartz, 1990). The observation of the FTIR data of a helix-coil transition for the S4 sequence is also consistent with a recent theoretical study which suggests that such a major conformational change occurs during the closing and opening of the Na⁺ channel (Leuchtag, 1994).

5.5 SUMMARY

The S4 region of the Drosophila Shaker voltage-gated K⁺ channel has been proposed to function as a voltage sensor. In order to elucidate its conformation, the polypeptide representing this sequence was chemically synthesized. Structural studies on the synthetic S4 polypeptide were conducted using FTIR and CD spectroscopy. Spectra were obtained for the synthetic polypeptide dissolved in aqueous solution, in trifluoroethanol solvent and also after reconstitution into lipid vesicles and micelles. The polypeptide in trifluoroethanol adopts an α-helical conformation which is in good agreement with the results obtained recently using 2D NMR study on the structure of a S4 polypeptide corresponding to the rat brain sodium channel. A predominantly α-helical structure is also observed when the S4 polypeptide is present in LPC micelles, and in DMPC vesicles. In contrast, the S4 polypeptide in aqueous medium is in random coil conformation. The coil-to-helix transition observed for the S4 polypeptide upon its transfer from aqueous medium to lipid membrane indicates that it has a high degree of conformational flexibility and can undergo large changes in its structure in response to its environment. This may have important implications for its role in the voltage activation process during
which the S4 polypeptide has been postulated to, at least partially, move from a lipid bilayer to an aqueous extracellular medium. The data obtained using synthetic S4 in this study lends support to such a model.
CHAPTER SIX: FUTURE PERPECTIVES
The studies presented in this thesis illustrate the advances made in the development of methods for the chemical synthesis of polypeptides, as well as the corresponding advances in FTIR spectroscopy in our understanding of biomembrane polypeptide conformation and dynamics provides unique opportunities to apply designed synthetic polypeptides for obstacles encountered in biology and medicine. These advances when coupled to the advances in molecular biology can provide hitherto unavailable and unimaginable application in areas of drug design, chemotherapy, immunology, technology and many other applications.

FTIR spectroscopic techniques provide invaluable insight into membrane-bound proteins providing researchers with precise model of protein structure. The resolving power of the technique is illustrated by its ability to distinguish structural arrangements between intramolecular aggregates and disaggregated transmembrane polypeptides, where incorrect structural assignment has been made in the past. This is illustrated by the successful elucidation of the synthetic Pf1 coat protein structure in chapter three, which should make it possible to accurately determine the IR absorption frequencies and in addition to provide quantification for membrane-bound proteins. Work is now in progress where the seven transmembrane helices of bacteriorhodopsin have been chemically synthesized so as to enable it to be used as a reference molecule in estimating transmembrane helices. For example, cytoplasmic and extracellular strands of membrane proteins could be cleaved by proteolytic enzymes, thus permitting the quantitation of known transmembrane proteins. The original study in the application of IR and FTIR
spetroscopy to protein conformation was based solely on theoretical calculations of proteins in the absence of water, and by empirical methods of assignment of IR absorptions relating to known structures of protein from X-ray diffraction models. With the advent of FTIR microscopy, it is now possible to obtain spectra of single protein crystals, which should enable the direct comparison of the secondary structures obtained from X-ray diffraction with those deduced by FTIR spectroscopy. As membrane protein crystals are difficult to obtain, reliable IR band assignment for soluble proteins can be established and a direct comparison of the protein structure in the crystalline and aqueous environment made.

The precise band allocation of IR absorptions of β-turns of type I, II and III, together with side chains of acetylated amino acid amides is needed to boost the resolving power of FTIR spectroscopy. This is now possible by applying the techniques as described in chapter four by using a comparison of synthetic transmembrane polypeptide, of which the definitive structural information of the natural form is known from NMR spectroscopy, neutron, and X-ray diffraction studies, to afford a more reliable band assignments of proteins. This can be a practical approach to the study of unavailable transmembrane proteins and encoded protein products, especially when molecular biology can provide sequence information. In the past, there have been several approaches to the study of protein structure-function studies: (a) By chemical means protein structure can be modified (Glazier, 1976); (b) By site-directed mutagenesis, gene structure can be modified with the resulting altered protein produced in suitable expression systems; (c) Altered proteins
may be produced by total or partial chemical synthesis of the polypeptide chain. The first technique is not always suitable because of the limited types of alterations that are feasible and the non-specificity of those changes, and the second may not always be convenient because of the time factor involved in the techniques and the fact that expression systems and purification of the modified product is not always clear-cut. The third technique has only been applied in very rare occasions, usually by conjugating synthetic polypeptides with natural protein fragments.

The technique described in this thesis lies at the interface between synthetic chemistry and molecular biology, and largely stems from the ability to chemically synthesize, analyse and visualize protein structures. Synthetic peptide chemistry can provide a far greater repertoire of fine protein manipulations than conventional site-directed mutagenesis and molecular biology. The application of FTIR spectroscopy in the study of protein-protein interactions requires further investigation. The absorption spectra of these interactions are at present complex and complicated to interpret because the amide I and II bands from the proteins in the sample tend to overlap. In order to overcome this problem it may be possible to chemically synthesize one of the participating polypeptides incorporating $^{13}$C at its polypeptide backbone. The displacement induced in its amide I band position may be adequate in reducing the overlap of the amide bands, thus permitting meaningful information to be gained on each of the proteins.

With the development of time-resolved FTIR spectroscopy, the improvement in scanning times have paved the way for obtaining information
on protein dynamics to be gathered at a time resolution of $\approx 10 \mu s$. This technique has been applied to the characterization of the L, M and N intermediate states of the of photocycle of bacteriorhodopsin (Braiman et al., 1987) when changes in the protein can be detected during proton pumping and the proton uptake cycle. This technique opens up new avenues for monitoring protein dynamics in protein folding and ligand binding. Another important field of research is the use of FTIR to investigate DNA structure. The easy access to milligram quantities of highly purified synthetic oligonucleotides permits the probing of sequence-dependent effects on DNA-structural transitions and ligand binding interactions. NMR spectroscopy has been applied in this field of work, however FTIR spectroscopy has yet to be fully exploited in this field. Another interesting area of study ideally suited for FTIR is the peptide-DNA binding interactions. Conformational changes imposed during these interactions requires investigation to better comprehend their structural basis. Specialized proteins such as restriction enzymes, polymerases and transcription factors which are able to differentiate between a binding site on DNA and random sequences, whereas the nucleases and chromosomal proteins have affinities to numerous different sequences awaits FTIR investigation. FTIR spectroscopy is capable of probing the secondary structure of the polypeptide and DNA simultaneously. In this overview of applications of synthetic polypeptides to modern biology and medicine, the emphasis has been the need for both physical/chemical and biological considerations in polypeptide design as well as the application of modern synthetic chemistry. Because macromolecular recognition and selectivity are essential features of most biological systems,
it is paramount that highly purified polypeptides are used to examine these
systems, whether the analysis is in the form of biophysical or biochemical
measurement or whether it is in the form a binding, in vitro, or in vivo
bioassay. It is also clear that chemically synthesized polypeptides and their
analogues, and peptidomimetics are potentially the major tools of the future.
This finding raises many issues in need of further study. As already outlined
herein, the process of discovery for obtaining polypeptide sequences that are
receptors (or acceptors) type or subtype selective for a specific ligand, a
specific enzyme, a specific toxin, a specific antibody or antigen and so on have
been accelerated in recent years by great advances in synthetic peptide
chemistry, molecular biology, biophysical techniques and bioassay systems.
Further studies are needed in these areas with special emphasis on peptide
delivery into biological systems, polypeptide incorporation into various
biomembrane barriers and to accumulate more reliable techniques for verifying
peptide conformations and dynamic properties.
REFERENCES


317


318


321


323

324
Schlessinger, J., Roche, R.S. and Steinberg, I. Z. (1975) Biochemistry 14, 255-262


325


Wang, S.S. (1973) J. Am. Chem. Soc. 95, 1328-1333


Wieland, T. (1968) Science 159, 946-952


326


"Studies of the pore-forming domain of a voltage-gated potassium channel protein."

"The conformational analysis of a synthetic S4 peptide corresponding to a voltage-gated potassium ion channel protein."

"Fourier transform infrared spectroscopy of membrane proteins and synthetic polypeptide domains."

"Disaggregation and reconstitution of hydrophobic synthetic polypeptides into lipid membranes."
Studies of the pore-forming domain of a voltage-gated potassium channel protein

Parvez I.Haris1, Bala Ramesh, Mark S.P.Sansom1, Ian D.Kerr1, Kaila S.Srai and Dennis Chapman

Department of Protein and Molecular Biology, Royal Free Hospital School of Medicine, University of London, Rowland Hill Street, London NW3 2PF and 1Laboratory of Molecular Biophysics, University of Oxford, Rex Richards Building, South Parks Road, Oxford OX1 3QX, UK

To whom correspondence should be addressed

Recent mutagenesis studies have identified a stretch of amino acid residues which form the ion-selective pore of the voltage-gated potassium channel. It has been suggested that this sequence of amino acids forms a β-barrel structure making up the structure of the ion-selective pore [Hartman, H.A., Kirsch, G.E., Drewe, J.A., Taglialatela, M., Joho, R.H. and Brown, A.M. (1991) Science, 251, 942–944; Yellen, G., Jurman, M.E., Abrams, T. and MacKinnon, R. (1991) Science, 251, 939–942; Yool, A.J. and Schwarz, T.I. (1991) Nature, 349, 700–704]. We have synthesized a polypeptide corresponding to this amino acid sequence (residues 431–449 of the Shaker potassium channel from Drosophila). A tetrameric version of this sequence was also synthesized by linking together four of these peptides onto a branching lysine core.

Introduction

Ion channels are integral components of biological membranes which permit controlled flux of selected ions across lipid bilayers. The elucidation of the mechanism by which ion channel proteins function remains an important challenge yet to be solved. Amino acid sequences for a number of channel proteins such as the voltage-gated sodium, potassium and calcium channels are now available [see Perney and Kaczmarek (1991) for a recent review] and in some cases site-directed mutagenesis studies have been employed to identify functionally important residues (Hartman et al., 1991; Yellen et al., 1991; Yool and Schwarz, 1991). The latter studies in combination with electrophysiological measurements have progressed remarkably over the last few years and are providing valuable information about various properties of ion channels. Despite the progress made by these techniques, the details of the molecular mechanism of channel function are only likely to emerge once their 3-D structures are determined at the atomic level. The large size and complexity of transmembrane ion channels often restricts their study using techniques normally used for protein structural analysis. One approach is to simplify these large complex proteins by conceptually dividing them into domains/motifs, identified from the primary structure. These domains can then be chemically synthesized and their structure determined by various spectroscopic techniques. The use of peptide domains as models to study specific ion channels and pore-forming proteins is becoming increasingly popular [for a recent review see Sansom (1991)]. We use this approach to investigate a functionally important segment of a potassium ion channel protein.

The sequence corresponding to the postulated pore region (called H5 or SS1-SS2, amino acid residues 431–449) of the Shaker potassium channel [Tempel et al., 1987; Yool and Schwarz, 1991] has been synthesized and spectroscopic measurements undertaken to investigate its structure. We have also synthesized an extended form of H5 that includes residues on its C- and N-termini. Furthermore, a tetrameric version of H5 was also synthesized as it is suggested that in the native protein, the pore is lined by four H5 chains, one from each subunit. The structures of the synthetic polypeptides are investigated in different environments such as in lyso phosphatidylcholine micelles and in phospholipid bilayers composed of dimyristoyl phosphatidyl-choline and dimyristoyl phosphatidyl-glycerol. The techniques of Fourier transform infrared (FT-IR) and circular dichroism (CD) spectroscopy are used for the structural analysis. Preliminary accounts of these structural studies have been reported (Haris and Chapman, 1992a; Haris et al., 1993). The ion channel activity of these peptides in planar lipid bilayers is also studied. Molecular modelling based upon the propensity of these peptides to form α-helical structures has also been performed.

Materials and methods

Peptide synthesis

The peptides were synthesized on an automated peptide synthesizer Rainsyn FS3 (Protein Technologies, Inc., USA) by a step-wise solid-phase procedure (Merrifield, 1963) using α-9-fluorenylmethylloxycarbonyl (Fmoc) protecting group strategy (Carpino and Han, 1972). The polypeptide was cleaved from the resin using 95% trifluoroacetic acid with scavengers and was purified isocratically on a reverse-phase HPLC column (250 x 10 mm i.d.) using 0.1% trifluoroacetic acid and acetonitrile as eluents. One major peak was observed when the eluate was monitored at 280 nm. The sequence of the 19-residue polypeptide (referred to as the monomeric H5 in this paper) corresponding to the H5 region of the Shaker potassium channel is as follows: CH3CO-DAFWAVVMTTTGQDM-GT-CONH2.

An extended version of the above sequence was also synthesized and the sequence is shown (the additional amino acid residues...
The polypeptide were obtained by digital subtraction of a spectrum of D$_2$O containing the buffer from the sample spectrum (Harris et al., 1986). The spectra of the aqueous buffer and the sample were recorded under identical conditions. Detailed analysis of the amide I band was carried out using second-derivative and deconvolution procedures (Harris et al., 1986).

CD spectra of the polypeptide in LPC micelles were recorded on a Jasco J600 spectrometer. The spectra were recorded using an 0.01 or 0.02 cm path-length cell. Spectra of aqueous dispersions of phospholipid were subtracted from the polypeptide spectra to remove any effects of slight absorbance of the phospholipids. The ratio of polypeptide to phospholipid used was the same as that used for the FT-IR measurements. However, the samples were diluted such that a polypeptide concentration of 1.25 mg/ml was obtained. Spectra were corrected for concentration and the $\Delta E$ values calculated.

**Planar bilayer studies**

Planar lipid bilayers were formed across a small (100 μm) hole in Teflon film by the apposition of two monolayers on either side of a Teflon partition (Montal and Mueller, 1972). The lipid-forming solution was phosphatidylcholine (Sigma, type II-S) and cholesterol, present in a 9:1 weight ratio, dissolved to 5 mg/ml in pentane. The monomeric H5 peptide was present in the lipid-forming solution at a 1:100 molar ratio. In the studies of the H5 tetramer, the bilayers were formed in a similar manner but using a 5 mg/ml solution of 1,2-dipalmitoyl-3-phosphatidylcholine (diPhyPC; Avanti Polar Lipids) in pentane. The tetrameric H5 peptide dissolved in 80% TFA was added with stirring to the cis face of the bilayer to a final concentration of 2.5 μg/ml.

In all of the bilayer experiments the electrolyte solution was 0.5 M KCl, buffered with 10 mM BES (monomeric experiments) or 200 mM BES (tetramer) to pH 7. Transbilateral currents were measured using an Axopatch 1D amplifier and were recorded using a Sony VCR and PCM.

**Molecular modelling**

Molecular modelling was performed using QUANTA 3.2 (Molecular Simulations, Waltham, MA) running on a Silicon Graphics Indigo workstation. Simulated annealing/molecular dynamics was carried out using Xplor 3.0 (Brünger, 1992) as previously described (Nilges et al., 1988; Nilges and Brünger, 1991; Kerr and Sansom, 1993b). Structure diagrams were generated using MolScript (Kraulis, 1991).

The degree of amphipathicity of H5 $\alpha$-helices was analysed via evaluation of molecular hydrophobicity potentials (MHPs), evaluated as described by Bricogne (1991), with the modification that the potential was set to a large positive (i.e. un favourable) value whenever a grid point fell within the van der Waals radius of an atom. The MHP is based on the free energy of transfer of constituent atoms from a hydrophilic to a hydrophobic phase. Thus, a negative MHP corresponds to a hydrophobic region of the molecule. MHPs were calculated at successive positions on a cylindrical polar grid defined by $(\tau, \phi, z)$, where $z$ is coincident with the helix axis, $\tau$ is the distance from the centre of the helix axis and $\phi$ is the angle subtended in the $z, \tau$ plane. Typically $\tau$ was incremented in 1 Å steps and $\phi$ in 10° steps, corresponding to ~20 000 grid points for $z = -25$ to $+25$ Å. The MHP calculation generates an array containing a potential at each grid point, $E(\tau, \phi, z)$. For each value of $(\phi, z)$, the minimum of $E$ with respect to $\tau$ is determined, thus defining an $E_{\text{min}}(\phi, z)$ surface, i.e. a hydrophobic surface map (Kerr and Sansom, 1993a). Averaging $E_{\text{min}}(\phi, z)$ over all $z$ values corresponding to
Fig. 2. FT-IR spectra of synthetic monomeric H5 polypeptide dispersed in dimyristoyl phosphatidylcholine vesicles in $\text{H}_2\text{O}$ buffer. The spectra were recorded at 30°C. (a) Absorbance spectrum. (b) Second-derivative spectrum.

Fig. 3. FT-IR spectra of synthetic monomeric H5 polypeptide dissolved in lysophosphatidylcholine micelles in $\text{H}_2\text{O}$ buffer. The spectra were recorded at 30°C. (a) Absorbance spectrum. (b) Second-derivative spectrum.
the length of the helix generates a plot of $E_{\text{max}}$ versus $\phi$ which may then be used to identify the centres of the hydrophobic and hydrophilic faces of the helix.

**Results**

**Spectra of H5 polypeptides in phospholipid dispersions**

The FT-IR spectra of the monomeric H5 polypeptide dispersed in DMPC vesicles are presented in Figure 2. An intense band at 1724 cm$^{-1}$ arises from vibrations of the carbonyl ester groups of DMPC. Information regarding the secondary structure of the polypeptide can be obtained from the frequency of the amide I band in the range 1615 - 1695/cm [for a recent review see Haris and Chapman (1992b)]. The amide I band occurs at 1651 cm$^{-1}$ and the band shape is quite symmetrical with no other distinct features apparent. This band frequency indicates that the polypeptide in DMPC vesicles has a predominantly $\alpha$-helical conformation (Haris et al., 1986; Jackson et al., 1989). The second-derivative procedure (Figure 2(b)) does not reveal any further major bands in the amide I region although there is a weak component at 1671 cm$^{-1}$. This may reflect the presence of some $\beta$-turn arrangements in the polypeptide structure (Haris et al., 1986; Jackson et al., 1989). This band is, however, of very weak intensity and some absorbance from amino acid side chains cannot be ruled out. The lack of any bands in the region 1625 - 1635 cm$^{-1}$ shows the absence of $\beta$-sheet structure in the H5 polypeptide. Hence, the FT-IR spectrum shows that in DMPC vesicles the H5 polypeptide is predominantly $\alpha$-helical in structure. The FT-IR spectrum of the H5 polypeptide obtained for a sample dispersed in DMPC phospholipid membranes exhibits its amide I maximum at 1652 cm$^{-1}$ consistent with a predominantly $\alpha$-helical structure.

Figure 3 presents the FT-IR absorbance spectrum of the monomeric H5 polypeptide in LPC micelles in $^2$H$^2$O buffer. The strong band at 1731 cm$^{-1}$ arises from the vibration of the phospholipid ester groups. From the analysis of the amide I band it can be concluded that the H5 polypeptide in LPC micelles is also predominantly $\alpha$-helical.

A high degree of hydrogen–deuterium exchange for the H5 polypeptide in both LPC micelles and also in DMPC vesicles was evident from the weak intensity of the amide II band near 1550 cm$^{-1}$.

The CD spectrum of the H5 polypeptide in LPC micelles is shown in Figure 4. The spectrum shows a minimum at 207 nm.
and another at 220 nm. This double minimum is characteristic of proteins and polypeptides with an α-helical conformation.

Thus, both the CD result and the FT-IR results are consistent in indicating that the H5 polypeptide adopts an α-helical structure in a membrane environment.

The FT-IR spectra of the extended and tetrameric version of H5 are presented in Figures 5 and 6 respectively. The amide I maxima for the extended and tetrameric forms of H5 dispersed in DMPC vesicles are centred at 1654 cm⁻¹ and 1649 cm⁻¹ respectively. The bands are highly symmetrical and there are no indications of any other major components. The amide I maxima at 1654 cm⁻¹ and 1649 cm⁻¹ can be assigned to an α-helical structure. Detailed analysis using the second-derivative method [Figures 5(b) and 6(b)] reveals only very weak components in the regions 1680–1671 cm⁻¹ and 1631 cm⁻¹. These spectra indicate there is little contribution from a β-sheet structure in both of these peptides.

The results obtained for the three forms of the H5 polypeptide in aqueous dispersions of LPC, DMPC and DMPG show the secondary structure is predominantly α-helical with only minor contributions from other secondary structural arrangements.

Planar bilayer measurements
Insolubility of the H5 monomer in aqueous solution necessitated its incorporation via direct addition to the lipid solution and, consequently, use of phosphatidylcholine/cholesterol as the bilayer-forming phospholipid. This method has proved successful for several hydrophobic channel-forming peptides (Oiki et al., 1988; Balaram et al., 1992). A relatively low molar ratio (1:100) of peptide:lipid ensures that any channel activity seen is not the result of detergent-like action of the peptide or formation of non-bilayer lipid phases. A 5 s section from a recording made from an H5 monomeric-containing bilayer is shown in Figure 7(a). For the first 4 s or thereabouts, the channel switches rapidly between the closed state (C) and the first open level (O1). The channel then opens to O1 and occasional transitions to a second open level (O2) are seen. An expanded segment (length 200 ms) from the recording is shown in Figure 7(b). The channel switches between C and O1. Channel openings are of duration 5–10 ms and of conductance 30 pS. Overall, a range of conductance levels from 12 to 50 pS were seen for bilayers incorporating monomeric H5.

The solubility of tetrameric H5 in TFA enabled it to be incorporated into preformed bilayers via the aqueous phase. In these experiments the diPhyPC was employed for bilayer formation, as in previous studies we have found this to generate a lower background noise level. Control experiments demonstrated that 1.6% TFA (the final concentration in the planar bilayer experiments) had no effect on stability of diPhyPC bilayers in the presence of 200 mM BES. A 5 s segment from a recording made from a diPhyPC bilayer exposed to ~0.4 μM tetramer is shown in Figure 7(c). There was a lag of ~30 min between addition, with stirring, of the peptide to the aqueous phase and the appearance of channel activity. There are occasional brief openings to an open state (O). A 200 ms expanded segment is shown in Figure 7(d). The channel switches from a closed state (C) to an open state. The open state is of brief duration (5–10 ms) and has a conductance of 13 pS.

Molecular modelling and analysis of amphipathicity
The FT-IR and CD results indicate that the H5 polypeptide is mainly α-helical when in a membrane mimetic environment. However, they do not provide information on possible side chain conformations. Simulated annealing/molecular dynamics was used to generate an ensemble of 25 structures for H5 in an α-helical conformation. Five models selected at random from

---

Fig. 6. FT-IR spectra of the tetrameric version of the H5 polypeptide dispersed in dimyristoyl phosphatidylcholine vesicles in 2H₂O buffer. The spectra were recorded at 30°C. (a) Absorbance spectrum. (b) Second-derivative spectrum.
Amphipathicity of the H5 helix may also explain stabilization of the helical conformation in the presence of phospholipids. As discussed by e.g. Sansom (1991), several studies have demonstrated that channel-forming and related peptides may be induced to form amphipathic helices at water—phospholipid interfaces. CD and NMR studies of amphipathic helices bound to phospholipid bilayers (Vogel, 1987; Dempsey and Butler, 1992) suggest that such helices may lie with the helix axis in the plane of the bilayer, i.e. perpendicular to the bilayer normal. In such an orientation the hydrophobic face of the helix interacts with the fatty acyl chains of the lipid molecules, whilst the hydrophilic face forms favourable interactions with the lipid head groups and the aqueous interface. Such a location of the peptide would be consistent with its accessibility to solvent as shown by the FT-IR hydrogen—deuterium exchange measurements. Impositions of a transbilayer electrostatic field may then cause peptide reorientation resulting in membrane-spanning helices which subsequently aggregate, to form ion channels.

Discussion

Recent studies employing molecular biological techniques have led to the identification of the amino acid sequence which constitutes the ion-selective pore of potassium, sodium and calcium channels (Hartman et al., 1991; Yellen et al., 1991; Yool and Schwarz, 1991). Voltage-gated ion channels are suggested to have six transmembrane segments designated S1 to S6. However, recent studies (Hartman et al., 1991; Yellen et al., 1991; Yool and Schwarz, 1991) indicate that a linker region (designated as SS1—SS2 or H5) connecting the S5 and S6 segment is the pore-forming sequence. Site-directed mutagenesis experiments on the Shaker H5 region and swapping of the H5 regions of two different potassium channels (drk1 with the Ngk2 sequence) indicate that this region determines the conductance and selectivity properties of the potassium channel. Previously this region was thought to be an extracellular loop. The suggestion for the H5 sequence spanning the membrane comes from experiments which demonstrate that the potassium channel blocker, tetracylammonium (TEA) ion, when applied to either the extracellular or intracellular side of the potassium channel, inhibits ion conductance. Mutation of residues 1 and 19 of H5 (where the numbering scheme refers to the peptide sequence given above) affects external but not internal blockade by TEA. In contrast to this, mutation of residue 11 affected only internal and not external TFA blockade. On the basis of these findings residues near either end of H5 were interpreted to be at the extracellular mouth of the channel and residues in the middle of H5 to be near to the intracellular mouth. Of course, this interpretation presumes that the effects of single point mutations are exerted directly and not via more subtle changes in protein conformation. A model has been postulated on the basis of the site-directed mutagenesis studies in which the H5 segment (residues 431—449 of the K+ channel) forms a β-hairpin structure. Four such hairpins are proposed to come together to form a channel, as in the detailed models described by Darrell and Guy (1992) and by Bogutz et al. (1992). However, there are as yet no experimental structural data available to support these modelling studies based on the proposal of the β-sheet structure.

The FT-IR and CD studies indicate that the H5 polypeptide
is largely helical when in a membrane mimetic environment. The planar bilayer studies indicate that H5 peptides may form channels with conductances in the range 12–50 pS, which should be compared with a conductance range of 5–30 pS for physiological K⁺ channels (Hille, 1992). This suggests that a helical conformation of H5 may be of relevance to its channel-forming
properties. We have therefore employed simulated annealing/molecular dynamics to generate an ensemble of \( \alpha \)-helical models of monomeric H5, in an attempt to further understand channel formation by this synthetic peptide. Amphipathicity analysis of these models helps to explain the helical conformation of H5, observed by spectroscopic measurements, when in the presence of lipid bilayers and suggests a possible role for H5 helices as a component of the intact channel protein.

There are three possible interpretations of the present data and the mutagenesis results. The first is that the H5 polypeptide is \( \alpha \)-helical as an isolated peptide but forms a \( \beta \)-hairpin only when part of the total protein structure. A second interpretation is that the H5 polypeptide is in fact helical in the intact channel. This seems difficult to resolve with the mutagenesis results and also would require changes to current topological models of K\(^+\) and related channels in order for the N- and C-terminal extra-membraneous domains to remain on the intracellular face of the membrane. However, as discussed above, it is possible that the point mutations do not all exert their effects directly. Furthermore, a definitive topology of the K\(^+\) channel has yet to be established experimentally. In the light of the compelling spectroscopic data and given the results of the bilayer studies, it does not seem unreasonable to propose that in the intact channel the pore is formed by a bundle of four H5 helices. A third interpretation would be that H5 undergoes a conformational transition, from an \( \alpha \)-helix structure to a \( \beta \)-hairpin arrangement, upon activation of the channel. This is similar in spirit to the model for Na\(^+\) channel activation proposed by Guy and Conti (1990). In the closed channel H5 would form an amphipathic helix, as in the structural and bilayer studies. Subsequent to activation of the channel, the H5 polypeptide would then form a \( \beta \)-hairpin structure. This model, whilst speculative, has the virtue of rationalizing both structural data on H5 in vitro and mutagenesis data on H5 in situ whilst also stressing the essentially dynamic nature of ion channel structures. However, as yet there is no evidence for such an \( \alpha \) to \( \beta \) transition underlying K\(^+\) channel activation. What is, however, clear from our study is that these synthetic polypeptides which have been shown to adopt an \( \alpha \)-helical arrangement in lipid membrane systems do display ion conductivity characteristics.

Acknowledgements

We wish to thank the Wellcome Trust for supporting work in both laboratories. I.D.K. is an SERC CASE student with Shell Research Ltd.

References


The conformational analysis of a synthetic S4 peptide corresponding to a voltage-gated potassium ion channel protein

Parvez I. Haris*, Bala Ramesh, Stephen Brazier, Dennis Chapman

Department of Protein and Molecular Biology, Royal Free Hospital School of Medicine, University of London, Rowland Hill Street, London, NW3 2PF, UK

Received 17 June 1994

Abstract

The S4 region of the Drosophila Shaker voltage-gated K⁺ channel has been proposed to function as a voltage-sensor. We have synthesised a peptide corresponding to this S4 region. Structural studies on the S4 peptide were conducted using Fourier transform infrared (FTIR) spectroscopy. Spectra were obtained for the peptide dissolved in aqueous solution, in trifluoroethanol solvent and also after reconstitution into lipid bilayers and micelles. The peptide in trifluoroethanol adopts an α-helical conformation which is in good agreement with the results of a recent 2D NMR study on the structure of a S4 peptide corresponding to the rat brain sodium channel [1989] FEBS Lett. 257, 113-117]. A predominantly α-helical structure is also observed when the S4 peptide is present in aqueous lysophosphatidylcholine micelles, in dimyristoyl phosphatidylcholine and dimyristoyl phosphatidylglycerol lipid bilayers. In contrast to this, the S4 peptide in aqueous solution is in a random coil conformation. The coil-to-helix transition observed for the S4 peptide upon its transfer from aqueous solution to lipid membrane indicates that it has a high degree of conformational flexibility and can undergo large changes in its structure in response to its environment. This may have important implications for its role in the voltage activation process during which the S4 peptide has been postulated to, at least partially, move from a lipid bilayer to an aqueous extracellular media [1992] Biophys J. 62, 238-250]. The results of our study lend support to such a model.

Key words: Protein conformation; FTIR spectroscopy; S4 domain; Voltage sensor; Potassium channel

1. Introduction

Voltage-gated ion channels are integral membrane proteins which conduct ions across biomembranes and are responsible for electrical excitability. Voltage-gated ion channels are thought to be constructed from four similar domains with each domain consisting of six transmembrane segments, S1 to S6 (for reviews, see [1-3]). The relatively small size of voltage-gated K⁺ channels makes them particularly attractive for structure-function analysis. Recent studies employing site-directed mutagenesis in conjunction with electrophysiological measurements have identified the functional role of several regions of this protein [2,3]. The most recent advance has been the identification of the sequence which constitutes the ion-selective pore [4-6]. Prior to this, a highly conserved S4 transmembrane domain has been suggested to have a role in sensing the transmembrane electric potential [7,8]. Mutations in the S4 segment of voltage-dependent ion channels have been shown to affect both voltage-sensing and gating currents [7-12]. The S4 segment contains a positively charged amino acid residue (lysine or arginine) at every third position. Mutagenesis studies have shown that these basic residues in the S4 region play a central role in the voltage-dependence of activation of K⁺ channels. Further support indicating that S4 is involved in voltage sensing comes from experiments in which S4 domains are swapped. In these situations the voltage-dependence is conferred largely by the substituted S4 domain [13,14]. Although mutagenesis studies have provided valuable information on the possible functional role of the S4 sequence, little is known about its structure. Recently, some studies have been directed towards obtaining a structural understanding of the S4 sequence. For example, the solution structure of a synthetic peptide corresponding to the S4 sequence of the rat brain Na⁺ channel has been determined using 2D NMR spectroscopy [15]. The ion conductance property of a synthetic Na⁺ channel S4 peptide in planar lipid bilayers has also been investigated [16]. While, some structural data are available for the S4 sequence of the Na⁺ channel, little or no data is as yet available for the corresponding sequence of the K⁺ channel. In the present study we employ FTIR spectroscopic measurements for conformational analysis of a synthetic peptide corresponding to the S4 sequence of the Shaker Drosophila K⁺ channel. FTIR spectroscopy is a useful technique for the structural analysis of peptides and proteins in an aqueous lipid bilayer environment (for a recent review see [17]).

*Corresponding author. Fax: (44) (71) 431 8107.

Abbreviations: FTIR, Fourier transform infrared spectroscopy; NMR, nuclear magnetic resonance spectroscopy; CD, circular dichroism spectroscopy; DMPC, dimyristoyl phosphatidylcholine; DMPG, dimyristoyl phosphatidylglycerol; LPC, lyso-phosphatidylcholine; TFE, trifluoroethanol; MBHA, 4-methyl benzhydrylamine.

0014-5793/94/$7.00 © 1994 Federation of European Biochemical Societies. All rights reserved.

SSDI 0014-5793(94)00704-7
2. Materials and methods

2.1. Materials

Dimyristoyl 

L -a-phosphatidyl choline (DMPC), dimyristoyl 

L -a- 

phosphatidyl-UL-glycerol (DM PG) and Lyso phosphatidyl choline (LPC), were purchased from Sigma Chemicals Ltd., UK. Protected amino acids were obtained from Novabiochem, UK.

2.2. Peptide synthesis and purification

The S4 peptide was synthesised on an automated peptide synthesiser (Rainin PS3, Protein Technologies, Inc.) by a stepwise solid-phase procedure (18) using /Va-9-fluorenylmethyloxycarbonyl (Fmoc) protecting group strategy (19). All the arginines had their side chains protected with 2,2,5,7,8-Pentamethylchroman-6-sulphonyl group and were double coupled throughout the synthesis. The Rink amide MBHA resin was used to generate the peptide with C-terminal amides and their N terminal acylated on completion of synthesis. The peptide was initially cleaved from the resin using 10% trifluoroacetic acid (TFA) in dichloromethane and final cleavage achieved with 95% aqueous TFA in the presence of scavengers. The crude peptide was purified on a semi-preparative reverse-phase HPLC column TSK ODS 120T column (300 mm x 7.8 mm ID) using 0.1% trifluoroacetic acid and acetonitrile as eluents. One major peak was observed when the eluate was monitored at 220 nm. The sequence of the 20 amino acid S4 peptide synthesised is as follows:

CHj-ILRVIRLVRVFRIFKLSRHS-NHj

The peptide was further characterised by amino acid analysis and the correct molecular weight was confirmed by laser desorption mass spectrometry. The peptide was found to be readily soluble in H2O.

2.3. Spectroscopic measurements

Spectra of the S4 peptide were recorded for samples in aqueous solution, trifluoroethanol (TFE) with 2% H2O, LPC micelles and in phospholipid bilayers composed of DMPC and DMPG. In order to reconstitute the S4 peptide into phospholipid vesicles, the peptide was dissolved in trifluoroactic acid and the solvent removed by rotary evaporation to produce a thin film. Phospholipid dissolved in dichloromethane was subsequently added to the peptide film, and once again the solvent removed. Further drying of the thin film of lipid-peptide mixture was carried out under vacuum for 24 h. Finally, 1H2O buffered with phosphate (150 mM sodium phosphate, pH 7.4) was added to the lipid-peptide film and mixed thoroughly using a Vortex mixer. The peptide to phospholipid molar ratio used in this study was 1:30.

Infrared spectra were recorded on a 1750 Perkin-Elmer spectrometer equipped with a TGS detector (20). The spectrometer was continuously purged with dry air to reduce water vapour absorption in the spectral region of interest. Samples were placed in a Specac cell fitted with a 50 μm Teflon spacer. For each sample 200 scans were signal averaged at a resolution of 4 cm−1. The peptide concentration employed for the IR measurements was 10 mg/ml. Absorbance spectra of the peptide were obtained by digital subtraction of the solvent spectrum i.e. either buffered 1H2O or 96% TFE. Detailed analysis of the amide I band was carried out using the second-derivative method (20).

3. Results

Analysis of the amide I band in the FTIR spectra of proteins and peptides provide useful information on their secondary structure (for a recent review see (17) ). The relationship between the frequency of the amide I band and secondary structure is now well established. The organic solvent TFE is known to enhance intermolecular interactions and therefore has been often used as an α-helical structure-inducing and stabilizing media. Spectra of the S4 peptide were therefore recorded for samples in TFE (Fig. 1). The amide I maximum is centred at 1,654 cm−1. This can be attributed to an α-helical structure. Second-derivative analysis shows the main amide I maximum at 1,655 cm−1. Minor components are revealed at 1,634 cm−1 and 1,683 cm−1. These bands are
weak and may originate from some turn type structure although overlap from β-sheet structure can also occur. The CD spectrum (not shown) displays double minima at 208 nm and 222 nm which is also typical of peptides and proteins in an α-helical conformation. Thus both types of measurement indicate that the S4 peptide has an α-helical structure when the S4 peptide is present in TFE solvent. This result is also in good agreement with the previous 2D NMR study of the S4 peptide corresponding to the rat brain sodium channel [15].

FTIR spectra of the S4 peptide were recorded in aqueous LPC micelles. The FTIR absorbance spectrum (Fig. 2a) shows a maximum at 1649 cm⁻¹. Detailed analysis using the second-derivative method (Fig. 2b) reveals a component of weak intensity at 1630 cm⁻¹ and 1670 cm⁻¹. These bands may arise from some turn/β-sheet structures. The main amide I band at 1649 cm⁻¹ is consistent with an α-helical structure. Evidence for an α-helical structure for the S4 peptide is also supported by the result of the CD study (not shown).

Fig. 2a also shows the absorbance spectrum of the S4 peptide in aqueous DMPC vesicles. The amide I band is centred at 1652 cm⁻¹. Detailed analysis using the second-derivative method (Fig. 2b) reveals the main amide I band at 1653 cm⁻¹. This is consistent with the presence of predominantly α-helical structure for the S4 peptide in DMPC lipid. Similar results were also obtained for the S4 peptide when reconstituted in DMPG phospholipid. Interestingly, raising the temperature above the lipid phase transition temperature of DMPC results in an increase in α-helical structure.

Fig. 2a shows the FTIR absorbance spectrum of the S4 peptide in aqueous solution (pH 7.4). The amide I band is centred at 1643 cm⁻¹. Detailed analysis of the amide I band using the second-derivative method (Fig. 2b) reveals only one additional component at 1670 cm⁻¹. The latter band is very weak in intensity, it may originate from turns/bends within the peptide structure. The main amide I maximum of 1640–1643 cm⁻¹ can be attributed to random coil structure. This assignment is further supported by our CD study which shows a spectrum typical of unordered conformation.

4. Discussion

Mutagenesis results clearly suggest that the S4 sequence play an important role in the conformational change which accompanies gating [3,7–12]. While these studies have provided valuable information about the possible functional role of the S4 sequence, very little hard structural data is available to explain the mechanism by which it achieves this function. Voltage-gated ion channels are large membrane proteins which are as yet not available in sufficient quantities for biophysical studies. This has been a major barrier for carrying out structural studies. One approach towards obtaining some structural information on these proteins is to syn-
that the S4 sequence of the Na+ channel has been studied in this way using NMR spectroscopy [15]. We have also used this approach to investigate the structure of peptides corresponding to the ion-selective pore of voltage-gated potassium channels [21].

In the absence of structural data various models have been proposed as to how S4 is involved in voltage-sensing [22-25]. In one such model it is postulated that the voltage-dependent conformational change involves a 'helical screw' movement of the charged S4 segment towards the extracellular surface [23]. The S4 is in its 'innermost' position in the protein structure when the channel is in the D0, deactivated state. Durell and Guy [24] have described how the S4 can move from the D0 to the open conformation by such an 'helical screw' mechanism. In the open conformation all the positively charged amino acids of S4 are on or near the extracellular surface, forming salt bridges with negatively charged residues in S1, S2, S3 and S5. The model of Durell and Guy [24] indicates that a substantial fraction of the S4 segment moves outside the lipid bilayer in the open state. This requires the peptide to be conformationally flexible, and the transfer from a membrane environment to a hydrophilic aqueous environment is likely to involve some rearrangement in the peptide structure. The results presented here provide information on such conformational changes for the S4 peptide. We have shown that the peptide is predominantly z-helical in TFE solvent, lysophosphatidylcholine micelles and in phospholipid membranes. However, upon transfer to an aqueous solution environment the peptide undergoes a major conformational rearrangement and adopts a random coil structure.

An increase in z-helicity was detected with the peptide in DMPC bilayers at temperatures above the lipid phase transition. We attribute this increase in z-helicity to greater penetration of the peptide into the lipid bilayer due to the increase in lipid fluidity. We envisage that as the lipid chains crystallize the peptide may be extruded from the lipid bilayer. The results of our FTIR spectroscopic study indicate that the S4 peptide has the conformational flexibility necessary for it to move from the lipid bilayer to the extracellular aqueous media. There is in fact some evidence to indicate that S4 does get exposed to the extracellular medium. Mutating one positive charge (lysine) near the NH-terminus of S4 alters the influence of extracellular Ca2+ ions but not that of intracellular Mg2+ [26]. This is interpreted to signify that this charge is exposed to the extracellular medium.

Furthermore, antibodies to the S4 sequence of Repeat I in Na channels appear to bind to the extracellular surface when the membrane is depolarized [27]. Our observation of a helix-coil transition for the S4 sequence is consistent with a recent theoretical study which suggests that such a major conformational change occurs during the closing and opening of the Na+ channel [25].

References