2-D NMR AND FTIR STUDIES ON AMPHIPATHIC AND HYDROPHOBIC PEPTIDES

by

QI ZHONG

B. Sc., University of Science and Technology of China 1983
M.Sc., Beijing University 1986

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
Doctor of Philosophy.

in the Department
of
Chemistry

© QI ZHONG 1992

SIMON FRASER UNIVERSITY

July 1992

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.
APPROVAL

Name: Qi Zhong

Degree: DOCTOR OF PHILOSOPHY

Title of Thesis: 2-D NMR AND FTIR STUDIES ON AMPHIPATHIC AND HYDROPHOBIC PEPTIDES

Examining Committee:

Chair: Dr. D. Sutton

Dr. R.J. Cushley, (Professor)  
Senior Supervisor

Dr. I.D. Gay, (Professor)  
Committee Member

Dr. A. Tracey, (Supervisor NMR Services)  
Committee Member

Dr. R.H. Hill, (Assistant Professor)  
Internal Examiner

Dr. F. Dahlquist
External Examiner
Department of Chemistry
University of Oregon
Eugene, Oregon U.S.A. 97403

Date Approved: November 26, 1992
PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay:

2-D NMR AND FTIR STUDIES ON AMPHIPATHIC AND HYDROPHobic Peptides

Author: Qi Zhong

(signature) (name) (date)
ABSTRACT

Two peptides, C-PEPTIDE (segment of apo C-I) and KIT (segment of c-KIT gene product), were studied extensively by 2-D NMR spectroscopy. DQF-COSY and TOCSY experiments were performed to achieve complete proton assignments for C-PEPTIDE in solution and in complexes with perdeuterated dodecylphosphocholine (DPC-d38), and a partial assignment for KIT in solution. A series of NOESY experiments with mixing times ranging from 75 ms to 550 ms was recorded and analyzed to obtain information on the 3-D structures of the peptides.

From the NOE measurements, the structural features of C-PEPTIDE in solution, C-PEPTIDE/DPC-d38 complexes and KIT in solution were determined. C-PEPTIDE adopted a conformation containing a certain amount of β-sheet with a mobile N-end. In complexes with DPC, C-PEPTIDE contained a partial helical conformation with a mobile N-terminal end. Proton-deuteron exchange experiments suggested that the DPC binding domain was at the C-end of the peptide. The binding mode remained unclear, although the exchange of back-bone protons in the helical region were relatively slower than those of other back-bone protons. KIT formed multimers in solution.

In the second part of this study, interactions between an amphipathic peptide, and three other peptides, and lipids were studied by Fourier transform infrared (FTIR). Hydrophobic moment profiles of the peptides suggested the possibility of β-like structures for the amphipathic peptide. The binding of the peptides to lipid membranes was verified by electron microscopy and turbidity measurements.

FTIR difference spectra were determined on peptides alone in solution and on peptide/dimyristoylphosphatidylycholine (DMPC) discoidal complexes. Second derivative spectra were calculated to eliminate the contribution from water vapor. The spectra were simulated using deconvoluted line heights and widths by an iterative procedure; a good fit being indicated by a small root mean square value. Band assignments were made according to the literature values. The results showed predominantly β-sheet content for the amphipathic peptide in lipid-peptide complexes and an intermolecular H-bonded β-structure when in solution. Another peptide, KIT, adapted mainly the β-structure in both solution
and the DMPC complexes. This agreed with the recent FTIR finding in this lab that the predominant secondary structure is the β-structure in high density lipoprotein. The β-sheet structures of amphipathic peptides in complexes disagreed with circular dichroism (CD) results and predictive methods of amphipathic helical secondary structure.
Dedications

To my parents and my wife Ming Yi
Acknowledgments

I would like to express my sincere thanks to my supervisor, Dr. R.J. Cushley, for his guidance, encouragement and support in the course of this research project.

I am indebted to Dr. F.W. Dahlquist for his enlightening discussion which helped me to achieve the complete assignments of C-PEPTIDE. His kindness and time spent on our discussion are very much appreciated.

I also wish to thank Dr. Brandhorst for helping me to prepare my Cumulative II exam, Dr. A. Tracey and Dr. I. Gay, as my supervision committee, for their time spent on my Cumulatives and this examination and Dr. R. Hill to be the internal examiner. My thanks extend further to Mrs. M. Tracey for collecting the spectra in the proton-deuteron exchange experiments, to Dr. D. Treleaven and Mrs. Lum (Jenny Yang) for routine collaborations, to Dr. V. Bourne in Biosciences, for help in preparation of electron micrographs, to Mr. T. Fischer and Mr. V. Varma for help in using the NMR spectrometer, and to people in this department in general.
# Table of Contents

Approval........................................................................................................... ii  
Abstract........................................................................................................ iii  
Dedications....................................................................................................... v  
Acknowledgments............................................................................................ vi  
Table of contents ............................................................................................ vii  
List of tables ................................................................................................... xii  
List of figures ................................................................................................... xiii  
List of abbreviations ...................................................................................... xx  
Contents in the appendix ............................................................................... xi  

Chapter 1. Introduction...................................................................................... 1  

1.1 Amphipathic Helix...................................................................................... 2  
  1.1.1 Amphipathic model.............................................................................. 2  
  1.1.2 Occurrence of amphipathic helices ....................................................... 3  
  1.1.3 Amphipathic peptides ......................................................................... 5  
    1.1.3.1 Fragments from native proteins...................................................... 5  
    1.1.3.2 Arbitrary synthetic peptides ......................................................... 5  
      1.1.3.2.1 LAP......................................................................................... 6  
      1.1.3.2.2 18A........................................................................................ 8  
  1.2 Problems and approaches......................................................................... 8  
    1.2.1 Circular dichroism............................................................................ 8  
    1.2.2 Hydrophobic moment....................................................................... 9  
    1.2.3 Fourier transform infrared................................................................. 10  
    1.2.4 Nuclear magnetic resonance experiments........................................ 11  
    1.2.5 Micelles............................................................................................ 13
1.3 Research objectives ................................................................. 14

Chapter 2. Theories ...................................................................... 16

2.1 Nuclear magnetic resonance spectroscopy ......................................... 16

   2.1.1 Principles ............................................................................. 16

   2.1.1.1 Bloch equations ................................................................. 16

   2.1.1.2 Relaxation mechanisms ....................................................... 18

      2.1.1.2.1 Dipole-dipole relaxation .............................................. 18

      2.1.1.2.2 Electric quadrupole relaxation ................................... 19

      2.1.1.2.3 Spin-rotation relaxation .............................................. 19

      2.1.1.2.4 Chemical shift anisotropy relaxation ......................... 20

      2.1.1.2.5 Scalar coupling relaxation ......................................... 20

   2.1.1.3 The nuclear Overhauser effect ............................................ 21

   2.1.2 Classical approach to 2-D NMR ................................................ 23

      2.1.2.1 Correlation spectroscopy ................................................. 23

      2.1.2.2 Nuclear Overhauser effect spectroscopy .......................... 23

   2.1.3 Product operator formalism ..................................................... 24

      2.1.3.1 The product operator formalism ....................................... 24

      2.1.3.2 Rules ............................................................................. 24

      2.1.3.3 COSY ............................................................................ 25

      2.1.3.4 Other COSY-type experiments ....................................... 26

         2.1.3.4.1 Double quantum filtered COSY ................................. 26

         2.1.3.4.2 Total correlation spectroscopy .................................. 27

   2.1.3.5 NOESY ................................................................. 27

   2.1.4 The relaxation processes and distance measurements .................. 30

   2.1.5 Phase cycling ....................................................................... 34

   2.1.6 Protein 3-D structure and 2-D NMR spectra .............................. 36
2.1.6.1 Sequential assignment strategy............................................36
   2.1.6.1.1 Side chain spin-system and sequential assignments..................36
   2.1.6.1.2 Conformation determinations..............................................38
2.1.6.2 Main-chain-directed strategy .............................................40

2.2 Fourier Transform Infrared Spectrum........................................41
   2.2.1 Interferogram ................................................................41
   2.2.2 Apodization ..................................................................42
   2.2.3 Self-deconvolution ..........................................................43
   2.2.4 The amide I band and the secondary structure of proteins............45
      2.2.4.1 Fitting ..................................................................45
      2.2.4.2 Band assignments ......................................................45
      2.2.4.3 Structural studies .......................................................46

Chapter 3. Materials and Methods ...............................................47
   3.1 Materials ..........................................................................47
   3.2 Methods .........................................................................48
      3.2.1 NMR spectra ...............................................................48
         3.2.1.1 C-Peptide ...............................................................48
         3.2.1.2 KIT ..................................................................49
      3.2.2 Assignment strategies .....................................................49
         3.2.2.1 Spin systems ............................................................50
         3.2.2.2 Structural studies ......................................................51
      3.2.3 Proton exchange experiments ............................................51
      3.2.4 Secondary structure predictions ........................................51
         3.2.4.1 Secondary shift of peptides ........................................51
         3.2.4.2 Hydrophobic moment ...............................................53
3.2.5 FTIR studies ................................................................. 54
  3.2.5.1 Sample preparation ................................................. 54
  3.2.5.2 FTIR spectra ......................................................... 54
  3.2.5.3 Deconvolution and curve fitting.............................. 55
  3.2.5.4 Assignments for the deconvoluted bands ................. 55

3.2.6 Electron Microscopy .................................................. 56

Chapter 4. C-PEPTIDE ......................................................... 57

  4.1 NMR of C-PEPTIDE in solution ..................................... 57
    4.1.1 Spin Systems ...................................................... 58
    4.1.2 Sequential Assignments and Conformation .................. 63
  4.2 C-PEPTIDE in complexes ............................................. 68
    4.2.1 Spin System ...................................................... 69
    4.2.2 Sequential assignments ......................................... 82
    4.2.3 Conformation .................................................... 89
  4.3 Proton Exchange ........................................................ 92
  4.4 Discussion .................................................................. 95

Chapter 5. KIT ................................................................. 101

  5.1 Spin System .............................................................. 102
  5.2 Sequence-specific assignment ..................................... 114
  5.3 Discussion .................................................................. 119

Chapter 6. Amphipathic Peptides ..................................... 120

  6.1 Structure prediction ................................................. 120
    6.1.1 Amphipathic helix .............................................. 120
    6.1.2 Hydrophobic Moment ......................................... 121
  6.2 Formation of Complexes ............................................ 124
  6.3 FTIR studies .......................................................... 126
6.3.1 LAP-20 .................................................................................. 126
  6.3.1.1 LAP-20 in solution ......................................................... 126
  6.3.1.2 LAP-20 in complexes with DMPC ............................... 129
  6.3.1.3 Discussion ................................................................. 131

6.3.2 Other peptides ................................................................. 135
  6.3.2.1 KIT ........................................................................ 135
  6.3.2.2 B19 ..................................................................... 139
  6.3.2.3 SH3 ................................................................. 143
  6.3.2.4 Discussion-Hydrophobic Peptides ............................ 146

6.4 The amphipathic helical model .......................................... 147

CONCLUSION .................................................................................. 149

REFERENCES .................................................................................. 157

Appendix ....................................................................................... 165
  1. Fitting program .................................................................. 165
  2. Secondary derivative ......................................................... 172
List of Tables

Table 1.1 Seven classes of amphipathic helices and their different properties.................................................................................................................. 4

Table 1.2 Characteristics of dodecylphosphocholine (DPC) micelles ....................... 13

Table 2.1 Sequential and medium $^1$H-$^1$H NOE in common secondary structures of proteins.......................................................... 39

Table 3.1 Sequences of the peptides ........................................................................ 47

Table 3.2 Canonical values of chemical shifts of 20 common amino acids.......................................................... 52

Table 4.1 Proton assignments of C-PEPTIDE in solution......................................... 65

Table 4.2 Complete assignments of C-PEPTIDE in complexes.............................. 88

Table 5.1 Proton assignments of KIT in solution .................................................. 118

Table 6.1 Hydrophobic Indexes of the peptides ...................................................... 120

Table 6.2 The diameters of the different complexes of peptide and DMPC .......................................................... 124

Table 6.3 Position and fractional areas of the amide I bands of LAP-20 in solution and in complexes with DMPC................................................ 131

Table 6.4 Position and fractional areas of the amide I bands of KIT in solution and in complexes with DMPC................................................ 136

Table 6.5 Position and fractional areas of the amide I bands of B19 in solution and in complexes with DMPC................................................ 140

Table 6.6 Position and fractional areas of the amide I bands of SH3 in solution and in complexes with DMPC................................................ 146

Table 6.7 The percent secondary structure of the peptides determined by FTIR................. 147
List of Figures

Figure 1.1 A helical wheel expression of an amphipathic helix............................................. 3
Figure 1.2 Helical expressions of amphipathic peptides............................................................... 7
Figure 1.3 Hydrophobic moment profile of LAP-20. ................................................................. 10
Figure 1.4 A micelle model ........................................................................................................ 14
Figure 2.1 Vectors used to mimic the behaves of a magnetization under magnetic fields H₀ and H₁ which represent main magnetic field and the second radio frequency, respectively.................... 17
Figure 2.2 The energy diagram for a two-spin system. The Ws are the transition probabilities. The subscripts indicate the number of spin changes. The I subscripts specify the relaxations of the spin I while the S represents those for the spin S. The energy levels are labelled by the numbers. ................................................................. 22
Figure 2.3 The pulse sequence for a COSY experiment. The t₁ and t₂ are two time variables. The y subscripts indicate the pulses along y axis.................................................................................. 23
Figure 2.4 The pulse sequence for a NOESY experiment. The τ is called the mixing time. Referring to Figure 2.3 for meanings of other symbols................................................................. 24
Figure 2.5 The product operator presentation of a COSY experiment for the spin I in a two spin (S,I) system. On the right hand side, the operators represent pulses and interactions. The bottom row shows magnetizations detected. The (od) and the (oc) represent diagonal and cross peaks, respectively. .................................................. 26
Figure 2.6 Evolution of product operators in a NOESY experiment. The (od) and (oc) represent diagonal and cross peaks, respectively................................................................. 29
Figure 2.7 Coherence transfer pathways in an NOESY experiment. ....................................... 35
Figure 2.8 Side chains of 20 amino acids. ............................................................... 37
Figure 2.9 Spin system of 20 amino acids. ................................................................. 37
Figure 2.10 NOESY-COSY connectivities in sequential assignments........................................ 38
Figure 2.11 The light paths of an infrared spectrometer (a) and the signal detected from the detector D (b). The S is a light source. D is a detector. The mirror 2 is a movable mirror. Δd is the distance difference between the two light paths d₁ and d₂.................................................. 42
Figure 4.1 1-D spectrum of 5 mM C-PEPTIDE in H$_2$O/D$_2$O (9/1, v/v). The spectrum was recorded at 400 MHz at 37 °C.

Figure 4.2 An αH relay region of 400 MHz TOCSY spectrum of 5 mM C-PEPTIDE solution in H$_2$O/D$_2$O (9/1, v/v) at 37 °C. 450 experiments were recorded with 32 scans and 2K complex data points for each scan. The spectrum width was 3846 Hz. A 1K x 1K real matrix was created by zero-filling once in the $t_1$ domain. Shifted squared sine bell functions were applied to $t_1$ and $t_2$ dimensions before double Fourier transform. Only positive peaks are shown.

Figure 4.3 Side chain region of the TOCSY spectrum of C-PEPTIDE in solution.

Figure 4.4 Fingerprint region of the TOCSY spectrum of C-PEPTIDE in solution.

Figure 4.5 NH relay region of TOCSY spectrum of C-PEPTIDE in solution.

Figure 4.6 Amide and aromatic region of TOCSY spectrum of C-PEPTIDE in solution.

Figure 4.7 Fingerprint region of 400 MHz NOESY spectrum of 5 mM C-PEPTIDE in solution in H$_2$O/D$_2$O (9/1, v/v) at 37 °C. The spectrum was recorded, using a 350 ms mixing time, and a digital resolution of 1.9 Hz with 450 $t_1$ increments. The spectrum was stored in a 1K x 1K real matrix. Zero-filling was performed once in the $t_1$ dimension, followed by application of shifted squared sine bell functions to both dimensions before double Fourier transform. Positive peaks were plotted.

Figure 4.8 Fingerprint region of 400 MHz NOESY spectrum of 5 mM C-PEPTIDE solution in H$_2$O/D$_2$O (9/1, v/v) at 37 °C. The identical conditions with Figure 4.7 were used except the mixing time of 500 ms and 350 $t_1$ increments.

Figure 4.9 1-D spectrum of C-PEPTIDE in complexes with DPC-d38 in H$_2$O/D$_2$O (9/1, v/v) at 37 °C.

Figure 4.10 Aliphatic region of DQF-COSY of 5 mM C-PEPTIDE in complexes with DPC-d38 at a ratio of 1 to 5 (mol/mol) in H$_2$O/D$_2$O (9/1, v/v) at 37 °C. A total of 208 scans was collected for each experiment and 320 experiments were recorded. The digital resolution of FID was 2.0 Hz. The apodization function was shifted sine bell function for both dimensions.

Figure 4.11 Cross peaks between αH and βH of aromatic residues of C-PEPTIDE in complexes in the DQF-COSY spectrum.
Figure 4.12 Fingerprint of the DQF-COSY of C-PEPTIDE in complexes .............................................. 72

Figure 4.13 Extended fingerprint region of the DQF-COSY of C-PEPTIDE in complexes ......................... 73

Figure 4.14 Aromatic region of the DQF-COSY of C-PEPTIDE in complexes ........................................ 74

Figure 4.15 Side chain region of TOCSY spectrum of 5 mM C-PEPTIDE in DPC-d38 complexes at a ratio of 1 to 5 (mol/mol) at 37 °C. The spin lock time was 70 ms and field strength was 9.9 kHz. 2K complex data points were collected for each FID and 512 experiments were recorded. The spectrum width was 4032 Hz in $t_2$. The spectrum was zero-filled to 1K in $t_1$ dimension and apodized with a phase-shifted squared sine bell function. .................................................. 75

Figure 4.16 Cross peaks between $\alpha$Hs and side chain protons in the TOCSY spectrum of C-PEPTIDE in complexes .......................................................... 77

Figure 4.17 Cross peaks between $\alpha$Hs and $\beta$Hs for amino acids in the medium chain group of C-PEPTIDE in complexes in the TOCSY spectrum ........................................... 78

Figure 4.18 Cross peaks of serines of C-PEPTIDE in complexes in the TOCSY spectrum .......................... 79

Figure 4.19 Relay cross peaks of C-PEPTIDE in complexes in the TOCSY spectrum .................................. 80

Figure 4.20 Fingerprint region of TOCSY spectrum of C-PEPTIDE in complexes ..................................... 81

Figure 4.21 Cross peaks from aromatic ring of tryptophan of C-PEPTIDE in complexes in the TOCSY spectrum ................................................................. 83

Figure 4.22 Cross peaks from two phenol groups of C-PEPTIDE in complexes in the TOCSY spectrum .......... 84

Figure 4.23 Fingerprint region of NOESY spectrum of 5 mM C-PEPTIDE in complexes in H$_2$O/D$_2$O (9/1, v/v) at 37 °C. The mixing time was 300 ms. 48 FID of 2K complex data points were co-added in each of 512 experiments. The spectrum was stored in a 1K x 1K real matrix after zero-filling once in the $t_1$ domain and apodization with shifted squared sine bell functions in both $t_1$ and $t_2$ dimensions were performed. Positive peaks were plotted ................................................................. 85

Figure 4.24 Side-chain NOE cross peaks of C-PEPTIDE in complexes ................................................. 86

Figure 4.25 Relay NOE cross peaks of C-PEPTIDE in complexes ...................................................... 87
Figure 4.26 NOE cross peaks between amides and aromatic ring protons of C-PEPTIDE in complexes. ................................................................. 90

Figure 4.27 Sequence of C-PEPTIDE and the NOE data used to characterize the secondary structure of C-PEPTIDE in a complex. ................................................................. 91

Figure 4.28 Secondary shifts of \(\alpha\)H of C-PEPTIDE in complexes. The canonical values (Wüthrich, 1986) were used as the standards. ................................................................. 91

Figure 4.29 Secondary shifts of amide NH of C-PEPTIDE in complexes. The standards were canonical values of amino acids ................................................................. 92

Figure 4.30 1-D spectra of C-PEPTIDE in solution. 16 scans were co-added for each experiment at room temperature. The digital resolution of FID was 0.12 Hz ................................................................. 93

Figure 4.31 1-D spectra of C-PEPTIDE in complexes. Experiments were performed at room temperature. 32 scans were recorded in each experiment. Same digital resolution with that for C-PEPTIDE in solution was used ................................................................. 94

Figure 4.32 Two binding modes of C-PEPTIDE to a micelle. ................................................................. 97

Figure 4.33 Secondary shifts of \(\alpha\)H of C-PEPTIDE in complexes. The chemical shift values of C-PEPTIDE in solution were taken as the standards ................................................................. 97

Figure 4.34 Secondary shifts of amide NH of C-PEPTIDE. The same calculation procedure as in Figure 4.33 were followed ................................................................. 97

Figure 4.35 Amphipathic helix of C-PEPTIDE in complexes. The numerals are sequential number of the residue. The bold characters are residues contained in the helix as shown by 2-D NMR studies ................................................................. 99

Figure 5.1 A 1-D spectrum of KIT solution in H\(\text{H}_{2}\)O/ D\(\text{H}_{2}\)O (9/1, v/v) at 37 °C and pH 2.43 ................................................................. 101

Figure 5.2 A plot of a 400 MHz DQF-COSY spectrum of 5 mM KIT in H\(\text{H}_{2}\)O/D\(\text{H}_{2}\)O (9/1, v/v) at 37 °C and pH 2.43. The spectrum was recorded in a digital resolution of 2.2 Hz with 512 \(t_1\) increments. A 1K x 1K real matrix was created by zero-filling once in the \(t_1\) dimension. Prior to the double Fourier transform, shifted sine bell window functions were applied to both dimensions. The positive peaks were plotted using 9 contours (dark ones) while the negative peaks using 2 contours. ................................................................. 103

Figure 5.3 Fingerprint region of the DQF-COSY spectrum of KIT. Thirteen cross peaks were assigned ................................................................. 104
Figure 5.4 Aromatic-NH region of the DQF-COSY spectrum of KIT

Figure 5.5 Fingerprint region of long chain amino acids of the DQF-COSY spectrum of KIT

Figure 5.6 Cross peaks between γCH2 and δCH2 of arginine and between δCH2 and εCH2 of lysine in the DQF-COSY spectrum of KIT

Figure 5.7 Fingerprint region for long and medium chain amino acids in the DQF-COSY spectrum of KIT

Figure 5.8 Fingerprint region for short chain amino acids in the DQF-COSY spectrum of KIT. Arrows indicated the cross peaks from amino acids in the short chain group

Figure 5.9 Long chain amino acid fingerprint region of a 400 MHz TOCSY spectrum of 5 mM KIT solution in H2O/D2O (9/1, v/v) at 37 °C and pH 2.43. The spectrum was recorded using a 70 ms MLEV-17 mixing time and a field strength of 8.8 kHz. The digital resolution of FIDs was 2.2 Hz and 512 t1 increments were collected. Prior to double Fourier transform, zero-filling to 1K in the t1 dimension was performed and a shifted sine bell window function was applied to both dimensions. Only the positive peaks were plotted. Arrows indicated the multi-relay cross peaks

Figure 5.10 Aromatic region of the TOCSY spectrum of KIT

Figure 5.11 Aliphatic region of the TOCSY spectrum of KIT

Figure 5.12 Relay cross peak region of the TOCSY spectrum of KIT

Figure 5.13 Fingerprint region of the TOCSY spectrum of KIT

Figure 5.14 Fingerprint regions of 400 MHz NOESY spectra of 5 mM KIT solution in H2O/D2O (9/1, v/v) at 37 °C and pH 2.43. Three different mixing times were used

Figure 5.15 Secondary shift of alpha protons of KIT as a function of the primary sequence. Blank spaces are residues without assignments

Figure 6.1 "Schiffer-Edmundson" helical wheel expressions of helices for KIT, B19, and SH3 and the purported amphipathic helices for C-PEPTIDE and LAP-20

Figure 6.2 The hydrophobic moment profiles for peptides

Figure 6.3 The negative stain electron micrograph of complexes of LAP-20 and DMPC
Figure 6.4. The negative stain electron micrograph of complexes of KIT and DMPC. ................................................................. 125

Figure 6.5. The negative stain electron micrograph of complexes of B19 and DMPC. ................................................................. 126

Figure 6.6. (a) Amide I range of the IR spectrum of LAP-20 in free form in D2O (thin line) plus simulated spectrum (thick line; see (c)); (b) the deconvoluted spectrum using a line width of 15 cm⁻¹ and an enhancement factor of 1.7; (c) Curve fitting of LAP-20 spectrum. The simulated spectrum (summation of component bands) is the top trace. Fitting components are shown as bands at the bottom ................................................................. 128

Figure 6.7. Amide I range of the IR spectra of LAP-20 in complexes with DMPC. The same notations as in Figure 6.6 were used. The deconvolution line width was 15 cm⁻¹ and the enhancement factor was 2.1. ......................................................................................... 130

Figure 6.8. Amide I range of the IR spectra of KIT in free form in D2O. The notations in Figure 6.6 were used except that the deconvolution line width was 16 cm⁻¹ and the enhancement factor was 2.4. ................................................................. 137

Figure 6.9. Amide I range of the IR spectra of KIT in complexes with DMPC. The notations in Figure 6.6 were used except that the deconvolution line width was 19 cm⁻¹ and the enhancement factor was 2.0. ......................................................................................... 38

Figure 6.10. Amide I range of the IR spectra of B19 in free form in D2O. The notations in Figure 6.6 were used except that the deconvolution line width was 20 cm⁻¹ and the enhancement factor was 2.3. See Table 6.5 for the band frequencies and areas. ......................................................................................... 141

Figure 6.11. Amide I range of the IR spectra of B19 in complexes with DMPC. The notations in Figure 6.6 were used except that the deconvolution line width was 18 cm⁻¹ and the enhancement factor was 2.5. ......................................................................................... 142

Figure 6.12. Amide I range of the IR spectra of SH3 in free form in D2O. The notations in Figure 6.6 were used except that the deconvolution line width was 20 cm⁻¹ and the enhancement factor was 2.1. See Table 6.6 for the band frequencies. ......................................................................................... 144

Figure 6.13. Amide I range of the IR spectra of SH3 in complexes with DMPC. The notations in Figure 6.6 were used except the deconvolution line width was 18 cm⁻¹ and the enhancement factor was 2.3. See Table 6.6 for other parameters. ......................................................................................... 145
List of Abbreviations

DMPC: dimyristoylphosphatidylcholine
DPC-d38: dodecylphosphocholine
2-D NMR: two dimensional NMR
TOCSY: total correlation spectroscopy
COSY: correlated spectroscopy
DQF-COSY: double quantum filtered COSY
NOE: nuclear Overhauser effect
NOESY: NOE spectroscopy
HDL: high density lipoprotein
Apo A-I: apolipoprotein A-I
Apo C-I: apolipoprotein C-I
FTIR: fourier transform infra-red
CD: circular dichroism

\[ \text{sinc} = \frac{\sin(x)}{x} \]

LCAT: lecithin cholesterol acyl-transferase
RMS: root mean square
EM: electronic microscope
Contents in the Appendix

Fit program ................................................................. 165
Secondary derivative program ................................... 172
Chapter 1. Introduction

The manner in which proteins and lipids interact is of fundamental importance in understanding biological processes. It is known that proteins play an important role in the integration of biological membranes, in cross-membrane transportation and in communication between the inside and outside of cells, and that complexation with lipids is necessary to impart activity to many proteins. Serum lipoproteins are a class of lipid-protein complexes which have attracted much attention because of their importance in the progression of ischemic heart diseases. Heart disease is the number one killer of people in North America. The human serum lipoproteins HDL, LDL and VLDL are small, approximately spherical particles consisting of a hydrophobic core of triglycerides and cholesteryl esters surrounded by an amphiphilic surface monolayer of cholesterol, phospholipids, and proteins. The apolipoproteins (A, C and E) are a class of lipid-associating proteins which reversibly bind to the surface of the lipid core. The structure and properties of human serum lipoproteins can be obtained from various reviews (Scanu, 1972; Jonas, 1991; Johnson et al., 1991; Miller, 1987; McKeone et al., 1988; Chen et al., 1989). Most of the apolipoproteins, except apo B, are soluble and have the ability to form complexes with lipids. They exist in equilibrium between the membrane bound form and the free form although the association constants are large. The biological functions of the apolipoproteins are related to their association with lipid particles. Some such associations might possibly be the key step in the metabolism of lipoproteins.

Despite the importance of the lipid associating proteins, the structural details of the protein-lipid interactions are poorly understood. The structural and primary sequential differences between soluble proteins and lipid associating proteins were studied in the past. Early studies by Segrest et al., (1974) revealed a correlation between the binding of lipids
and an increase of $\alpha$-helical content of exchangeable apoproteins as measured by circular dichroism (CD). It was also found that, under identical conditions, apoproteins that did not associate to membranes had no changes in circular dichroism spectra. Based on the relationship of the lipid association to the helical transition shown in CD studies, a new theory, the amphipathic helix model, was derived.

1.1 Amphipathic Helix

1.1.1 Amphipathic model

This theory was used to predict helical secondary structures from protein primary sequences, and to explain the interactions between lipids and peptide chains at the interfaces. A protein fragment could be defined as an amphipathic region if it had (a) a relatively large non-polar face which could interact with fatty acid chains of phospholipids, and (b) a polar face consisting of mostly charged amino acids with an arrangement of negative charges at the center and positive charges at the periphery. The special feature of the amphipathic helix was a two-face topological distributions of polar and nonpolar amino acids. The best amphipathic conformation was alleged to be a helix (Figure 1.1). One side of the helix was highly hydrophobic, which was the lipid binding region, while the other side of the helix was more hydrophilic. In a lipid-protein complex, the nonpolar face, consisting of hydrophobic residues, was buried in the hydrocarbon region of lipids and the polar face, composed of polar residues, was directed toward the aqueous phase interacting with charges on phospholipids. According to this theory, two interactions keeping peptides bound to lipid membranes were hydrophobic forces and ionic interactions between charges
on peptides and charges on phospholipids. They were quantified by the hydrophobicity index and the number of ion pairs.

Figure 1.1 A helical wheel expression of an amphipathic helix.

1.1.2 Occurrence of amphipathic helices

A broad search for the amphipathic helix among the known protein sequences was carried out by Segrest et al., (1977). The results demonstrated that the amphipathic sequence was common among apoproteins as well as other proteins. Amphipathic helices were classified into seven classes according to mean unit hydrophobicities of the apolar face, the charged residue densities and the topographic distributions of charges (Table 1.1. Segrest et al., 1990). It was also found that the amphipathic domain might be involved in other interactions such as protein-protein interactions (class G) and protein-DNA interactions besides the original proposed protein-lipid interactions (class M).
Table 1.1 Seven classes of amphipathic helices and their different properties (From Segrest et al., 1990).

<table>
<thead>
<tr>
<th>PROPERTY</th>
<th>A</th>
<th>H</th>
<th>L</th>
<th>G</th>
<th>K</th>
<th>C</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apolipoproteins</td>
<td>0.42</td>
<td>0.54</td>
<td>0.37</td>
<td>0.72</td>
<td>0.42</td>
<td>0.38</td>
<td>0.22</td>
</tr>
<tr>
<td>Poly-peptide hormones</td>
<td>0.73</td>
<td>0.57</td>
<td>0.74</td>
<td>0.64</td>
<td>0.55</td>
<td>0.80</td>
<td>0.74</td>
</tr>
<tr>
<td>&quot;Lytic&quot; poly-peptides</td>
<td>1.9</td>
<td>2.4</td>
<td>1.6</td>
<td>1.3</td>
<td>3.0</td>
<td>2.1</td>
<td>0.09</td>
</tr>
<tr>
<td>Globular proteins</td>
<td>2.0</td>
<td>0.8</td>
<td>30</td>
<td>5.7</td>
<td>0.8</td>
<td>1.3</td>
<td>0.6</td>
</tr>
<tr>
<td>Calmodulin-Coiled-coil proteins</td>
<td>2.0</td>
<td>0.5</td>
<td>0.4</td>
<td>1.3</td>
<td>0.2</td>
<td>2.5</td>
<td>0.25</td>
</tr>
<tr>
<td>Transmembrane proteins</td>
<td>3.9</td>
<td>2.9</td>
<td>2.0</td>
<td>2.6</td>
<td>3.2</td>
<td>4.6</td>
<td>0.3</td>
</tr>
<tr>
<td>Mean +/− charge ratio</td>
<td>0.9</td>
<td>4.8</td>
<td>4.0</td>
<td>1.1</td>
<td>15</td>
<td>0.8</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Mean localization of charged residues by quadrant:
- Positive: Non-polar face
- Polar face
- Negative: Non-polar face
- Polar face

Mean angle subtended by polar face:
- ≥ 180
- ≤ 100
- ≤ 100
- ≥ 180
- ≥ 180
- ≥ 320
- ≤ 60
1.1.3 Amphipathic peptides

Because of the lack of a method to observe directly the amphipathic domain in a protein molecule, studies of synthetic peptide analogs of the amphipathic helix became the only way to test the model. Three typical synthetic peptides were studied as models.

1.1.3.1 Fragments from native proteins

A gene similarity found in five major human apolipoproteins, apo A-I, apo E, apo A-II, apo C-II and apo C-III, included a structure of four exons and three introns. The most striking homology of exchangable apolipoproteins was a repeating segment of 22 amino acids which had a strong amphipathic helix-forming potential (Segrest et al., 1990, Leblond and Marcel 1991). One peptide, corresponding to residues 121 to 164 of apo A-I, was synthesized. It had the ability to bind to unilamellar phospholipid vesicles and to phospholipid-cholesterol mixed vesicles. In biological assay, it could activate the enzyme lecithin:cholesterol acyltransferase (LCAT) in cholesterol esterification and phospholipid hydrolysis reactions as efficient as 30% of the activity of apo A-I (Fukushirna et al., 1980). A correlation of the biological activity with the α-helical content of the peptide in 50% trifluoroethanol was suggested (Fukushima et al., 1980).

1.1.3.2 Arbitrary synthetic peptides

The arbitrarily designs of amphipathic peptides were directed by rules specifying the lipid associating properties (Pownall et al., 1980). They were summarized as (i) the peptide must have a potential to form α-helix; (ii) the peptide must have 20 (or more) residues in length, which is designated as the critical amphipathic length; (iii) the peptide must have a strong hydrophobicity; (iv) the peptide must follow the amphipathic model. Two families of such synthetic peptides were designated LAP (for lipid associating peptide) and 18A.
1.1.3.2.1 LAP

The LAP family had three members, LAP-16, LAP-20 and LAP-24. All of them were designed to possess the two face amphipathic feature. Experimentally, they were found to bind to dimyristoyl phosphatidylcholine (DMPC) membranes which was verified by a blue-shift in fluorescence experiments. The CD results showed that the α-helical content increased for all three peptides upon binding to DMPC (Pownall et al., 1984).

LAP-20, designed by Pownall and coworkers (Pownall, et al., 1980), had 65% the activity of Apo A-I as an activator of lecithin:cholesterol acyltransferase (LCAT). Both the amphipathic helical model and the hydrophobic moment profile predicted that LAP-20 could form an amphipathic helix in complexes with DMPC (Figure 1.2). The LAP-20/DMPC complex formation was verified by change in turbidity and a blue-shift in the tryptophan fluorescence (Pownall, et al., 1980). The LAP-20/DMPC complexes were demonstrated on the electron micrographs as stacks of discoidal particles (termed Rouleaux).

LAP-20 was proposed to contain less helical structure (10%) in solution but contained more helix (90%) when bound to a lipid bilayer in CD studies (Pownall, et al., 1980).
Figure 1.2 Helical expressions of amphipathic peptides.
The 18A peptide series were designed by Segrest and his colleagues (Kanellis et al., 1980; Anantharamaiah et al., 1985). The peptide, 18A, designed to mimic native amphipathic sequences of apo A-I, demonstrated a good lipid affinity and formed stable complexes with DMPC. A covalently linked dimer of 18A, 18A-pro-18A, had the highest homology to the native apo A-I dimer in lipid binding experiments. The third peptide in this family, 18R, had a reversed amphipathic helix with positive and negative charges exchanged. As expected, peptide 18R had the lowest affinity for DMPC and formed a less stable complex with DMPC. In CD studies of those peptides, peptide 18A showed a doubling of the \( \alpha \)-helical content, from 15\% to 30\%, upon binding to DMPC. Peptide 18A-pro-18A had 49\% and 53\% \( \alpha \)-helical content in the solution and in complexes with DMPC, respectively. Peptide 18R showed no significant \( \alpha \)-helicity. It was concluded that omission of the hydrophobic interaction of lysines with the non-polar face caused the loss of lipid affinity of peptide 18R.

1.2 Problems and approaches

1.2.1 Circular dichroism (CD)

Circular dichroism has been used widely to determine the conformations of proteins. From typical CD studies in the literature, one could find that the \( \alpha \)-helix content of proteins was reported while other secondary structures, such as \( \beta \)-sheets and turns, were ignored. The helix content is overestimated from the CD data (Yang et al., 1991). Besides, when dealing with complexes of proteins and lipids, lipids could cause two special
experimental problems. (a) The low signal to noise ratio due to light scattering from lipid particles could interfere with the measurements of ellipticities (Yang et al., 1991, Fukushima et al., 1980). (b) The most used lipid in studying of protein-lipid interactions was phosphatidylcholine (PC). It appears that not enough attention has been paid to the chiral carbon atom of a PC molecule in measurements of ellipticities (Fukushima et al., 1980). It can be concluded that CD is not a reliable method to apply to protein-lipid mixtures. Since CD results have provided the major experimental support for the amphipathic helical model, the existence of the predicted amphipathic helix became questionable. Recent FTIR studies in this laboratory revealed a major β-sheet conformation of high density lipoprotein (HDL, Yang et al., 1991) which contradicted the early finding of α-helices in HDL based on CD studies (Surewicz et al., 1986; Lux et al., 1972).

1.2.2 Hydrophobic moment

The hydrophobic moment method was introduced to quantify the amphiphilicity of a periodic segment (Eisenberg et al., 1984, see Chapter 2 for details). It was a more dynamical method for predicting protein secondary structures in the sense that it related conformations to the amphiphilicities. A large hydrophobic moment meant a strong amphiphilicity for a peptide chain. The periodicity of the hydrophobicity matches the periodicity of the structure. From hydrophobic moment profiles, conformations with strong amphiphilicities could be predicted, for example, the conformations of LAP-20 in complexes with lipids (Figure 1.3).
According to the hydrophobic moment prediction, a periodicity of 100° corresponds to a helical structure while a 160°-180° periodicity is predicted to be a β-sheet structure. A standard α-helix contains 3.6 residues each turn, which means a 100° rotation between adjacent residues. So, α-helices have a periodicity of 100°. A β-sheet is an extended structure with a zig-zag arrangement of the peptide chain. The periodicity of β-structure is 160°-180°. The value can be less than 180° predicted for the all-trans conformation since β-sheet have ripples and twists.

As shown in Figure 1.3, LAP-20 shows both an α helical peak and a β-structure peak. This means that, if the amphiphicity is the criteria to select secondary structure, β-sheet would be as good as a helix for LAP-20 secondary structure.

1.2.3 Fourier transform infrared (FTIR)

Infrared (IR) spectroscopy has been used to study secondary structures of polypeptides and proteins in solution. In the past, the method was little used because of factors such as low sensitivity of instruments, strong absorbance from water and difficulties in extracting structural information from an IR spectrum. With the advent of
Fourier transform infrared, greatly improved sensitivity of FTIR measurements (signal-to-noise ratio) was obtained. FTIR became the method of choice to estimate secondary structures of proteins (Surewicz et al., 1990; Dong et al., 1990; Wantyghem et al., 1990; Muga et al., 1990). The technique is of particular interest in elucidation structures of membrane or lipid-associated proteins (Surewicz et al., 1987; Muga et al., 1991).

Compared to CD, FTIR has some advantages such as no light scattering affecting infrared measurements of protein-lipid complexes. Unlike CD, FTIR measurements do not emphasize any particular secondary structure; α-helices, β-sheets, random coils and turns are all quantified in FTIR measurements (Sarver, Jr. and Krueger, 1991). However, FTIR has some drawbacks in protein structural studies. Like CD, it is an experimental method and gives an average structure of a protein at relatively high concentrations. Efforts are continuing to find the detailed correspondence between various bands and specific types of protein structures.

1.2.4 Nuclear magnetic resonance experiments

NMR spectroscopy and x-ray crystallography are the two techniques mainly used to investigate detailed 3D conformations of biological macromolecules. NMR has some advantages over x-ray crystallography. For example, a NMR experiment is independent of the formation of the single crystal, which is the limitation of the x-ray method, and can be carried out in a solution which is closer to the physiological condition than a crystal. Furthermore, some dynamic processes are amenable to NMR studies but not x-ray studies. It was hoped that the NMR method would provide complementary data to those from x-ray experiments and that the application of the NMR method to biological systems would give a better elucidation of the relationship between structures and functions of proteins. NMR methods were improved greatly by the introduction of two dimensional (2-D) and three
dimensional (3-D) NMR experiments which could be applied to study large biomacromolecules.

2-D NMR methods rely on two main steps, (a) proton assignments for a macromolecule; (b) $^1$H-$^1$H distance measurements in nuclear Overhauser effect (NOE) experiments. An empirical structure can be built up by coordinating distance constraints into the conformation of a protein (Wüthrich 1986). Distance geometry calculations based on the empirical model should give a precise secondary structure of a protein (Kline et al., 1988). It has been concluded that the empirical NMR model derived from NOE patterns is more detailed and precise than the secondary structure obtained by prediction algorithms or by spectroscopic techniques (Wüthrich 1989). It has been successful in determining protein conformations using NMR techniques (Reinerowski et al., 1990; Eberle et al., 1990; Mammi and Peggion 1990; Skelton et al., 1990; Omichinski et al., 1992; Phillips et al., 1991; Saudek and Pelton 1990; Williamson et al., 1985; Oas et al., 1990; McIntosh et al., 1990) and it is anticipated that large proteins with molecular weights in the range of 30,000 to 40,000 would be amenable to heteronuclear multidimensional NMR studies in the near future (Wüthrich 1989).

Though NMR is powerful, a complex of a protein or a peptide with phosphatidylcholine (PC) unilamellar vesicles is too large to subject to NMR studies. The diameter of a unilamellar vesicle is about 300 Å. The smallest discoidal complex has a size of 200 Å in diameter. In early NMR experiments with native membranes or reconstituted protein-phospholipid unilamellar vesicles, no protein $^1$H signals were observed (Brown, et al., 1975; Austin, et al., 1975; Brown, et al., 1977; Dufourcq, et al., 1976; Feigenson et al., 1977). A logical explanation was that the size of the protein-lipid vesicles was so large that the correlation time of complexes was too long. The consequence of slow motions was the line broadening effect. All signals from such a slow motion system were buried in the
baseline. To overcome the line broadening effect, one must use either heteronuclear multidimensional NMR techniques or a small complex.

1.2.5 Micelles

Since micelles of detergent molecules are smaller than vesicles of phospholipids, a micelle system was established to replace native membranes and artificial vesicles (Wider et al., 1982; Kohda & Inagaki 1992), and characterized (Dill & Flory 1981; Lauterwein et al., 1979). For example, micelles were formed by mixing of melittin with perdeuterated dodecylphosphocholine (DPC, Table 1.2).

The structures of melittin and mEGF in complexes with perdeuterated DPC were determined using 2-D NMR techniques (Lauterwein et al., 1979; Kohda and Inagaki 1992).

Table 1.2 Characteristics of dodecylphosphocholine (DPC) micelles (Lauterwein et al., 1979).

<table>
<thead>
<tr>
<th>Structural parameter</th>
<th>Experimental basis</th>
<th>DPC/micelle</th>
<th>peptide/DPC molar ratio/micelle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stoichiometry</td>
<td>particle weight</td>
<td>56 molecules</td>
<td>1.1/32</td>
</tr>
<tr>
<td>Diameter of equivalent sphere</td>
<td>Diffusion constant</td>
<td>47 Å</td>
<td>49 Å</td>
</tr>
<tr>
<td></td>
<td>Diffusion constant</td>
<td>55 Å</td>
<td>56 Å</td>
</tr>
<tr>
<td></td>
<td>Particle weight and volume</td>
<td>39 Å</td>
<td>35 Å</td>
</tr>
<tr>
<td>Maximum thickness of hydration shell</td>
<td>Diffusion constant</td>
<td>6 Å</td>
<td></td>
</tr>
<tr>
<td>Maximum asymmetry of prolate ellipsoid</td>
<td>Diffusion constant</td>
<td>a/b=6</td>
<td></td>
</tr>
</tbody>
</table>

A micelle has a hydrophilic layer of polar heads of detergent molecules and a hydrophobic core of acyl chains (Figure 1.4). This structure resembles native membranes. So the behavior of proteins on micelles can be studied as a model of interactions between
proteins and native membranes (Lauterwein et al., 1979; Wider et al., 1982; Mammi and Peggion, 1990; Kohda and Inagaki, 1992).

![A micelle model](image)

Figure 1.4 A micelle model (Drill and Flory, 1981).

1.3 Research objectives

One of our objectives was to determine the 3-D structure of an amphipathic peptide and to investigate interactions between peptides and lipids.

Our main approach was 2-D NMR. It has been proven to be powerful in structural studies. From NOE peak patterns, we were able to characterize the secondary structure of a peptide. There are obvious differences between spectra of a random coil and spectra of an ordered structure. If it is a helix, strong NH-NH cross peak patterns and some special medium range NOE patterns are expected. β-sheet structures would display a strong αH-NH NOE pattern.
The DPC micelle system was adopted to study the conformational changes upon the binding of peptides to micelles. Since the similarities of micelles to biological membranes were shown by various authors (see above), we believed that it was reasonable to test the amphipathic model in a micelle system. According to the amphipathic helix theory, an amphipathic peptide will be in random coils, or less structured conformations, in solution and be a helix when bound to membranes or micelles. NMR experiments were performed on peptides in solution and bound to micelles, to find the spectral differences indicating the structural changes from a random coil to an ordered structure.

Another aspect of the current study was a comparison of CD and FTIR as two methods for predicting secondary structure of amphipathic peptides. Recent studies on apo A-I in this lab revealed a large amount of β-sheet existing in complexes of apo A-I with DMPC and in HDL₃ (Yang et al., 1991) which contradicted the earlier model featured by nine helices which was based on CD results (Marcel et al., 1991). It is known that CD exaggerates α-helices to such an extent that the amphipathic model would have been a misleading result. Since CD is not a reliable method to study lipid-protein complexes, FTIR studies were undertaken on the protein-lipid complexes to test the existence of α-helices proposed by CD studies.
Chapter 2. Theories

2.1. Nuclear magnetic resonance spectroscopy

2.1.1 Principles

2.1.1.1 Bloch equations

In a static magnetic field, a nucleus of spin I has $2I+1$ energy levels. For example, nuclei of spin 1/2 reside at two energy levels, indicated by $\alpha$ and $\beta$, in a static magnetic field $H_0$ at thermal equilibrium. The population at the low energy level is greater than that at the high level, according to the Boltzmann distribution, so there is a net magnetization $M_0$. When a radio-frequency field (rf) with an angular frequency of $\omega$, which is equivalent to a second magnetic field $H_1$, is applied to the system, transitions between two energy levels will occur if $\omega=\omega_0$. The $\omega_0$ is called the Larmor frequency, which has a value of $\gamma H_0$. The $\gamma$ is the gyromagnetic ratio of the nucleus. After the $H_1$ is turned off, the signals are detected and the system returns back to the thermal equilibrium state. This recovery process is called the relaxation.

Vector descriptions of magnetizations give a more visible picture. For example, at the thermal equilibrium, magnetizations are represented by vectors aligned along the direction of the magnetic field $H_0$ (z axis). The interactions between magnetizations and $H_0$ are displayed as a precessing movement of the vectors about $H_0$ (Figure 2.1.a). The effect of a second rf field ($H_1$) on magnetizations is indicated by rotations of the vectors away from the z axis (Figure 2.1.b). The components of the vectors in the x-y plane are detected.
The spin-lattice relaxation process is viewed as the vectors rotating back to the z axis from the x-y plane (Derome 1988).

Figure 2.1 Vectors used to mimic the behavior of a magnetization under magnetic fields $H_0$ and $H_1$ which represent main magnetic field and the second radio frequency, respectively.

Mathematical descriptions of the interactions of the magnetic moments with $H_0$ and $H_1$, and of relaxations, are given by the Bloch equations [2.1.1], [2.1.2] and [2.1.3].

$$\frac{dM_x}{dt} = \gamma \left[M_y H_0 + M_z H_1 \sin(\omega t)\right] - \frac{M_x}{T_2} \quad [2.1.1]$$

$$\frac{dM_y}{dt} = \gamma \left[M_z H_1 \cos(\omega t) - M_x H_0\right] - \frac{M_y}{T_2} \quad [2.1.2]$$

$$\frac{dM_z}{dt} = \gamma \left[M_x H_1 \sin(\omega t) + M_y H_1 \cos(\omega t)\right] - \frac{(M_z - M_0)}{T_1} \quad [2.1.3]$$

where $T_1$ (spin-lattice relaxation time) and $T_2$ (spin-spin relaxation time) are two characteristic times for the magnetization to return to equilibrium. $M_x, M_y, M_z$ are the x, y, z components of the magnetization $M$. $M_0$ is the magnetization $M$ at thermal equilibrium.

The relaxation is particularly interesting in NMR studies.
2.1.1.2 Relaxation mechanisms (Noggle and Schirmer 1971; Farrar and Becker 1971; Abragam 1961)

Relaxations can be initiated by a number of different physical mechanisms. In general, any mechanism which gives rise to a fluctuating magnetic field at a nucleus is a possible relaxation mechanism. Five important relaxation mechanisms are i) intra-molecular dipole-dipole interactions (DD), ii) electric quadrupole interactions (Q), iii) scalar coupling interactions (SC), iv) spin-rotation interactions (SR), v) chemical shift anisotropy interactions (SA).

For any given spin system, the observed relaxation rates will result from contributions of various different relaxation mechanisms, and can be expressed as a summation of rates from all mechanisms.

\[
\frac{1}{T_1} = \frac{1}{T_1(DD)} + \frac{1}{T_1(SA)} + \frac{1}{T_1(SC)} + \frac{1}{T_1(SR)} + \frac{1}{T_1(Q)} \hspace{1cm} [2.1.4]
\]

\[
\frac{1}{T_2} = \frac{1}{T_2(DD)} + \frac{1}{T_2(SA)} + \frac{1}{T_2(SC)} + \frac{1}{T_2(SR)} + \frac{1}{T_2(Q)} \hspace{1cm} [2.1.5]
\]

2.1.1.2.1 Dipole-dipole relaxation

This is the dominant mechanism in liquid for spin 1/2 nuclei. When two nuclei I and S are placed in a static magnetic field $H_0$, they experience a total magnetic field $H_{tot} = H_0 + H_{loc}$, where the $H_{loc}$ is a time-dependent field from neighboring spins. The $H_{loc}$ is a function of $r_{is}$ and $\theta_{is}$, where $r_{is}$ is the distance between the two dipoles, $\theta_{is}$ is the angle between $r_{is}$ and $H_0$. If the two dipoles are within a rigid molecule, $H_{loc}$ is a function of angle $\theta_{is}$.

The quantitative expressions of dipolar relaxation rates for two 1/2 spins are
where $\tau_c$ is the correlation time for the reorientation of the vector $r$, $\gamma$ is the gyromagnetic ratio of the nuclei.

2.1.1.2.2 Electric quadrupole relaxation

Any nucleus with the spin $I$ greater than $1/2$ has an electric quadrupole moment $eQ$. The interaction of the nuclear quadrupole moment with the electric field gradient $eq$ at the nucleus provides a very efficient relaxation mechanism. The relaxation rates, in extreme narrowing, are

$$\frac{1}{T_1(Q)} = \frac{1}{T_2(Q)} = \frac{3}{40} \frac{(2I+3)}{I^2(I-1)} \frac{\eta^2}{3} \frac{e^2 Q q}{\hbar} \frac{4\pi^2}{\hbar^2} \tau_c$$  \hspace{1cm} [2.1.7]$$

where $e$ is the electron charge, $\tau_c$ is the correlation time, and $\eta$ is the field gradient asymmetry which describes deviations of the nuclear environment from axial symmetry.

2.1.1.2.3 Spin-rotation relaxation

The interaction of a nuclear magnetic moment and the rotational magnetic moment of the molecule in which the nucleus is situated provides a direct mechanism for transferring nuclear spin energies to the nucleus.

$$\frac{1}{T_{1(SR)}} = \frac{1}{T_{2(SR)}} = \frac{2}{3} \frac{(I \hbar T)}{\hbar^2} \frac{4\pi^2}{\hbar^2} C^2 \tau_{SR}$$  \hspace{1cm} [2.1.8]$$

$C^2$ is the squared average of the SR tensor, $\tau_{SR}$ the spin-rotation correlation time, $I$ the moment of inertia, $k$ the Boltzmann constant, $T$ the absolute temperature.
2.1.1.2.4 Chemical shift anisotropy relaxation

Chemical shifts originate from shielding effects. In liquid state, the shielding factor $s$ is an average of components along $x$, $y$, $z$ axes, but on a shorter time scale it undergoes fluctuations in the local magnetic field. If $\sigma_x$, $\sigma_y$, $\sigma_z$ are not equal, this provides a relaxation mechanism.

\[
\frac{1}{T_1(SA)} = \frac{1}{15} \gamma^2 H_0^2 \left( \sigma_{//} - \sigma^\wedge \right)^2 \frac{2 \tau_c}{(1 + \omega^2 \tau_c^2)} \tag{2.1.9}
\]

\[
\frac{1}{T_2(SA)} = \frac{1}{90} \gamma^2 H_0^2 \left( \sigma_{//} - \sigma^\wedge \right)^2 \left[ \frac{6 \tau_c}{(1 + \omega^2 \tau_c^2)} + 8 \tau_c \right] \tag{2.1.10}
\]

$\sigma_{//}$, $\sigma^\wedge$ refer to the shielding factors along and perpendicular to the symmetry axis of an axially symmetric system, respectively (London 1989).

2.1.1.2.5 Scalar coupling relaxation

When scalar couplings between spins $I$ and $S$ are interrupted by other time dependent processes, they can be a relaxation mechanism. If the interrupting process is a chemical exchange, it is called "scalar coupling of the first kind". For spin $I$, the relaxation rates are given in [2.1.11] and [2.1.12].

\[
\frac{1}{T_1(SC)} = \frac{2}{3} S(S+1) A^2 \frac{\tau_c}{(1 + (\omega_I - \omega_S)^2 \tau_c^2)} \tag{2.1.11}
\]

\[
\frac{1}{T_2(SC)} = \frac{1}{3} S(S+1) A^2 \left[ \tau_e + \frac{\tau_e}{(1 + (\omega_I - \omega_S)^2 \tau_c^2)} \right] \tag{2.1.12}
\]

where $A$ is the coupling constant which becomes a function of time because of chemical exchange, $\tau_e$ the effective correlation time for the resulting exchange process, $\omega_I$ and $\omega_S$ are Larmor frequencies of the two nuclei $I$ and $S$. 
If the scalar couplings are modulated by a rapid relaxation of the spin S, the splitting of S on the spin I will not be observed. This is termed "scalar coupling of the second kind".

\[
\frac{1}{T_1^{(SC)}} = \frac{8}{3} \pi^2 J^2 S(S+1) \frac{T_{2s}}{(1+(\omega_I-\omega_S)^2 T_{2s}^2)} \tag{2.1.13}
\]

\[
\frac{1}{T_2^{(SC)}} = \frac{4}{3} \pi^2 J^2 S(S+1) \left[ T_{2s} + \frac{T_{2s}}{(1+(\omega_I-\omega_S)^2 T_{2s}^2)} \right] \tag{2.1.14}
\]

where \( T_{2s} \) is the spin-spin relaxation time of spin S. \( J \) is the scalar coupling constant.

Some nomenclature used by Noggle and Schirmer (1971) will be used hereafter.

2.1.1.3 The nuclear Overhauser effect (NOE)

In a multispin system, the relaxation of spin I is a summation of all mechanisms as mentioned above.

\[
\frac{d\langle I_{zi} \rangle}{dt} = -\rho_i (\langle I_{zi} \rangle - I_{0i}) - \sum_{j} \sigma_{ij} (\langle I_{zj} \rangle - I_{0j}) \tag{2.1.15}
\]

\( \langle I_{zi} \rangle \) represents the z component of the magnetic moment of the nucleus i. \( I_{0i} \) is the value of \( \langle I_{zi} \rangle \) at thermal equilibrium, \( \rho_i \) is the total relaxation rate constant for the spin i, \( \sigma_{ij} \) is the cross-relaxation rate constant contributed by spin j to spin i.

Cross-relaxations can be achieved by interactions of the magnetic moments of two spins through space or by chemical exchange processes. The interactions through space are called dipolar couplings which are more interesting to us than the chemical exchanges. Because of the cross relaxations, two spins do not relax freely, but influence each other. The deviation of a spin from equilibrium can be transferred to the other spin through the dipolar couplings and vice versa. I will use a two-spin system to describe the dipolar couplings in some detail. The energy levels of a two spin system are depicted in Figure 2.2. \( W^I \) and \( W^S \) are transition probabilities for I and S spins. When the spin I is polarized
by a rf at the Larmor frequency of the spin I, the populations of I are equalized between the state 1 (or 3) and the state 2 (or 4). The population differences of the spin I from those in the Boltzmann equilibrium generate a magnetization $-I_z$. There is no effect on population distributions of the spin S in absence of the transitions $W_2$ and $W_0$. Since the population differences of the spin S can be rearranged between the state 1 and the state 4 by $W_2$ and between the state 2 and the state 3 by $W_0$, the population differences of the spin S are changed by a factor of $W_2 - W_0$, when pathways $W_2$ and $W_0$ are introduced. The magnetization is transferred from the spin I to the spin S (Figure 2.2, Kessler et al., 1988). The factor of $W_2 - W_0$ is the cross relation rate which is a function of frequencies and the distance between the two spins. It can be calculated in a two like-spin system by the equation [2.16].

\[
\text{Cross Relation Rate} = \frac{1}{10} \frac{\gamma^4 h^2}{r^6 \pi^2} \left[ \frac{6 \tau_c}{(1 + 4 \omega^2 \tau_c^2)} - \tau_c \right]
\]

where $\gamma$ is the gyromagnetic ratio, $\omega$ is the Larmor frequency, and $\tau_c$ is the correlation time.

Figure 2.2 The energy diagram for a two-spin system. The $W$s are the transition probabilities. The subscripts indicate the number of spin changes. The I subscripts specify the relaxations of the spin I while the S represents those for the spin S. The energy levels are labelled by the numbers. (Solomon, 1955, Harris, 1983).
2.1.2 Classical approach to 2-D NMR

In order to get an explicit picture of a 2-D NMR experiment, I will use vector representations to demonstrate basic 2-D experiments.

2.1.2.1 Correlation spectroscopy (COSY)

The pulse sequence of a COSY experiment is \((\pi/2)_y - t_1 - (\pi/2)_y - t_2\) (Aue et al., 1976, Nagayama et al., 1980, Bax and Freeman 1981) (Figure 2.3). The \(t_1\) and \(t_2\) are two time intervals, \((\pi/2)_y\), is a 90° pulse along the \(y\) axis. Initially, magnetization vectors lie along the \(z\) axis at thermal equilibrium. The first pulse flips the vectors down to the \(x\)-\(y\) plane along the \(-x\) axis. During \(t_1\), the vectors (chemical shift +\(J/2\) and chemical shift -\(J/2\)) are fanned out due to spin-spin interactions (\(J\) couplings). The projections of the vectors on the \(y\) axis are modulated by sine functions of \(t_1\). The second pulse flips all vectors out of the \(x\)-\(y\) plane except the components on the \(y\) axis. The \(y\) components are detected in \(t_2\) as functions of \(t_1\) and \(t_2\). Double Fourier transformations of the data in the time domain give a spectrum in frequency domains.

\[\begin{array}{c}
90^\circ_y & t_1 & 90^\circ_y & t_2 \\
\end{array}\]

Figure 2.3 The pulse sequence for a COSY experiment. The \(t_1\) and \(t_2\) are two time variables. The \(y\) subscripts indicate the pulses along \(y\) axis.

2.1.2.2 Nuclear Overhauser effect spectroscopy (NOESY)

The pulse sequence in a NOESY experiment is \((\pi/2)_y - t_1 - (\pi/2)_y - \tau - (\pi/2)_y - t_2\) which is similar to a COSY experiment (Jeener et al., 1979, Kumar et al., 1980, Bodenhausen et al., 1984) (Figure 2.4) except for insertion of a mixing time \(\tau\) and a pulse before the
acquisition time $t_2$. Following the first pulse and the $t_1$ duration, the $x$ components of coherences are flipped by the second pulse to the $z$ axis. Cross-relaxations take place during time delay $\tau$ among the vectors along the $z$ axis by transformations of coherences between the spin I and the spin S. The resulting coherences, which are regulated by the dipolar couplings, are rotated back to the $x$ axis in the $x$-$y$ plane by the third pulse and detected in $t_2$. The spectrum is obtained by a double Fourier transform.

![Pulse sequence for a NOESY experiment](image)

**Figure 2.4** The pulse sequence for a NOESY experiment. The $\tau$ is called the mixing time. Referring to Figure 2.3 for meanings of other symbols.

2.1.3 Product operator formalism (Kessler et al., 1988)

2.1.3.1 The product operator formalism

The operator formalism was first introduced by Sorensen and his colleagues (Sorensen et al., 1983). It was conducted by quantum mechanics calculations and gives a clear picture similar to that from the magnetization vector descriptions. It has been successfully used to give a detailed explanation of a correlation spectroscopy (COSY) experiment (Atta-ur-Rahman 1989).

2.1.3.2 Rules

The Cartesian operators are employed to represent magnetizations. $I_x$ is associated with an $x$ magnetization, $I_y$ with a $y$ magnetization, $I_z$ with a $z$ magnetization. The Cartesian operators ($I_x, I_y, I_z$) act just like magnetizations when interactions (pulses, chemical shifts or $J$ couplings) take place. An NMR experiment is simplified to a series of expressions.
The starting state of a spin is connected to the end state by an arrow. The terms above the arrows are the operators. $\beta \varphi$ represents a pulse with a flip angle of $\beta$ about an axis in the $x$-$y$ plane which forms an angle $\varphi$ with the $x$ axis. $\Omega I_z t$ represents the evolution of chemical shift $\Omega$ of a nucleus I during the time $t$. $2\pi J I_z S_z t$ represents the evolution of couplings between nuclei I and S during the time $t$.

The basic rules for using the operators are listed below:

a). Each operator in the product operator is transformed separately during the evolution time.

b). Two transverse operators or two longitudinal operators do not show mutual couplings, couplings being observed between products of transverse and longitudinal operators.

c). The influence of chemical shifts and coupling frequencies is simultaneously observed.

d). Only transverse operators are observed after $t_2$.

2.1.3.3 COSY

COSY is the simplest 2-D experiment with two $90^\circ$ pulses (Figure 2.3). The first pulse generates magnetizations in the $x,y$ plane which are frequency labelled in the evolution time $t_1$. The second pulse selects only parts of the magnetizations in the $x,y$ plane which are detected in $t_2$.

In a COSY experiment, the diagonal peaks are generated by chemical shifts as in expression (A). Cross peaks represent J couplings as in expression (B).

\[ I_z \xrightarrow{90^\circ_z} I_x \xrightarrow{t_1} I_y \xrightarrow{90^\circ_y} I_y \xrightarrow{t_2} I_x, I_y \quad (A) \]

\[ I_z \xrightarrow{90^\circ_y} I_x \xrightarrow{t_1} -2I_x S_z \xrightarrow{90^\circ_y} 2I_z S_x \xrightarrow{t_2} S_x, S_y \quad (B) \]
The complete product operator expression of a COSY experiment is given in Figure 2.5.

![Diagram of COSY experiment](image)

Figure 2.5 The product operator presentation of a COSY experiment for the spin I in a two spin (S,I) system. On the right hand side, the operators represent pulses and interactions. The bottom row shows magnetizations detected. The (od) and the (oc) represent diagonal and cross peaks, respectively.

2.1.3.4 Other COSY-type experiments

2.1.3.4.1 Double quantum filtered COSY (DQF-COSY)

Double quantum filtered COSY is basically a COSY type experiment which has been employed to study J-couplings (Rance et al., 1983). DQF-COSY has three 90°
pulses. Principally, the third pulse converts the double quantum coherences generated by
the second pulse to single quantum coherences which are detected in the $t_2$ time period.
Because the anti-phase magnetizations for diagonal peaks are converted to double quantum
coherences first and then back to anti-phase magnetizations, both cross peaks and diagonal
peaks have anti-phase character and can be phased to have simultaneously pure 2-D
absorption lineshapes. Compared to regular COSY spectra which have dispersion modes
for diagonal peaks, DQF-COSY spectra have a narrow diagonal.

2.1.3.4.2 Total correlation spectroscopy (TOCSY)

Another useful COSY type experiment is called TOCSY (Bax and Davis
1985), in which a spin-lock field is applied to the system which causes, so called
isotropic mixing of spins. During the spin-lock time, an oscillatory transfer of
coherences throughout the coupled spin systems occurs. As a result of such a
"mixing", all pairs of spins in a coupling network, no matter whether directly or
indirectly coupled, will show cross peaks. Practically, a TOCSY experiment with a
short mixing time is a substitute of a single-relay COSY experiment, a TOCSY
experiment with a long mixing time could serve as a multi-relay experiment (Ernst et
al., 1987, Bax, 1989).

2.1.3.5 NOESY

NOESY is a basic experiment with three $90^\circ$ pulses and three time variables
(Figure 2.4). For a system of two spins $I$ and $S$, the product operator description of
a NOESY is simple. The first $90^\circ$ pulse flips magnetizations $I_z$ and $S_z$ from the $z$
axis to the $x$ axis. During the evolution time $t_1$, the magnetic moments precess freely
in the $x,y$ plane about the $z$ axis by an angle of $\omega t_1$. So $I_x=I_0\cos(\omega t_1)$ and
$S_x=S_0\cos(\omega s t_1)$. The second pulse converts $I_x$ and $S_x$ to longitudinal components,
\[ I_z = -I_0 \cos(\omega_1 t_1) \] and \[ S_z = -S_0 \cos(\omega_S t_1) \], respectively (neglecting \( T_2 \) relaxations).

Magnetization relaxations take place in the mixing time \( \tau \), including \( T_1 \) relaxations and cross relaxations between the spins \( I \) and \( S \). The magnetization transfers between the spin \( I \) and the spin \( S \) are allowed by the cross relaxations. The consequences of such transfers are that a part of the \( S_z \) becomes \( I_z \) and vice versa. The \( S_z \) and the \( I_z \) are restored along the x axis by the final 90° pulse for detection.

Four signals are detected in the second time variable \( t_2 \). The non-transferred magnetizations give the diagonal peaks at \((\Omega_1, \Omega_2) = (\Omega_I, \Omega_I)\) and at \((\Omega_S, \Omega_S)\) while the transferred magnetizations show the cross peaks at \((\Omega_1, \Omega_2) = (\Omega_I, \Omega_S)\) and at \((\Omega_S, \Omega_I)\).

With the rules outlined by Kessler et al., (1988), the product operators in a NOESY experiment construct Figure 2.6 for spin \( I \) in a non-scalar coupled two-spin system (\( S, I \)).
Figure 2.6 Evolution of product operators in a NOESY experiment. The (od) and (oc) represent diagonal and cross peaks, respectively.

The developments of the cross and the diagonal peaks are given in expressions (C) and (D), respectively, without considerations of the intensity modulations.

To construct the spectrum, it is necessary to use the whole expression of the free induction delay (FID) in (E) and (F).
The \( f(\tau) \) and \( g(\tau) \) are two relaxation functions of mixing time \( \tau \) which will be discussed in some detail in the next section. The magnetization transfer from \( I \) to \( S \) is involved in the formation of \( S \) in (E) but not for \( I \) in (F), so the double Fourier transform of (E) and (F) gives one cross peak at \((\Omega_1, \Omega_3)\) and a diagonal peak at \((\Omega_i, \Omega_i)\) in a NOESY spectrum.

Expressions (E) and (F) suggest that NOE cross peaks and diagonal peaks are both \( \tau \)-dependent signals. Practically, it has been observed that NOE cross peak intensity increases from 0 at \( \tau=0 \) to a maximum, then decreases to 0 again at \( \tau=\infty \) (Kumar et al., 1981). But the diagonal peak intensity decreases all the way to 0 at \( \tau=\infty \) (Kumar et al., 1981). The detailed functions are given later in equations [2.1.23] and [2.1.26].

NOE buildup rates have been widely used to compute internuclear distances (Wüthrich 1986). The NOE buildup rate is the changing rate of the intensity of the NOE cross peaks which can been obtained from a graph of NOE cross peak intensities versus the mixing time \( \tau \). The initial changing rate is inversely proportional to the distance to the power of six (Wüthrich 1986) which will be proven in the following section.

2.1.4 The relaxation processes and distance measurements

In a 2-dimensional NOE experiment, as discussed previously, the first two pulses generate longitudinal components \( I_z = I_0 \cos(\omega_1 t_1) \) and \( S_z = -S_0 \cos(\omega_3 t_1) \). During the mixing time \( \tau \), both spin-lattice and cross relaxation take place. The relaxation process
during mixing time \( \tau \) (hereafter \( t \) is used for convenience) is characterized by equations [2.1.17] and [2.1.18] for a two-spin system.

\[
\frac{d<\mathbf{I}_z>}{dt} = -\rho_I [<\mathbf{I}_z>-\mathbf{I}_0] - \sigma [<\mathbf{S}_z>-\mathbf{S}_0] \tag{2.1.17}
\]

\[
\frac{d<\mathbf{S}_z>}{dt} = -\rho_S [<\mathbf{S}_z>-\mathbf{S}_0] - \sigma [<\mathbf{I}_z>-\mathbf{I}_0] \tag{2.1.18}
\]

where \( \rho_I = W_0 + 2W_{1I} + W_2 \)

\( \rho_S = W_0 + 2W_{1S} + W_2 \)

\( \sigma = W_2 - W_0 \) (Solomon 1955, Harris 1983).

\( \mathbf{I}_z \) is the z component of the magnetization of the spin I; \( \mathbf{I}_0 \) is \( \mathbf{I}_z \) at the thermal equilibrium, \( \rho_I \) is the overall spin-lattice relaxation rate constant for the spin I. A similar definition applies to the spin S. The \( \sigma \) is the cross relaxation rate constant between the spin I and spin S.

For a two-like-spin system,

\[
\rho_I = \rho_S = \frac{1}{10} \frac{\gamma^2 h^2}{4\pi^2} \frac{\tau_c}{r^6} \left[ 1 + \frac{3}{1+\omega^2\tau_c^2} + \frac{6}{1+4\omega^2\tau_c^2} \right] \tag{2.1.19}
\]

\[
\sigma = \frac{1}{10} \frac{\gamma^2 h^2}{4\pi^2} \frac{\tau_c}{r^6} \left[ \frac{6}{1+4\omega^2\tau_c^2} - 1 \right] \tag{2.1.20}
\]

where \( \gamma \) is the gyromagnetic ratio of proton; \( h \) is the Planck constant; \( \tau_c \) is the correlation time; \( r \) is the distance between the two protons (Noggle and Schirmer 1971, Kessler et al., 1988).

After Laplace transformation (Rainville and Bedient 1974) and putting in the initial conditions of \( \mathbf{I}(t=0)=\mathbf{I}_i \) and \( \mathbf{S}(t=0)=\mathbf{S}_i \), one obtains the solutions of the equations [2.1.17] and [2.1.18] in the equations [2.1.21] and [2.1.22]. The variable \( t \) is the mixing time.
As stated above, \( I_i = -I_0 \cos(\omega_1 t_1) \) and \( S_i = -S_0 \cos(\omega_S t_1) \) in a NOESY experiment, \( \omega \) is the Larmor frequency, and \( t_1 \) is length of the first time variable. These are the frequency-labelled signals.

The first term in equation [2.1.21] represents the magnetization of the spin \( I \) which originates from the spin \( I \). Therefore, it gives a diagonal peak in a NOESY spectrum. The second term in [2.1.21] characterizes a magnetization of the spin \( I \) which is transfered from the spin \( S \). So it shows a cross peak. Among the signals, those, not frequency-labelled, give axis peaks in a NOESY spectrum. Equation [2.1.21] and equation [2.1.22] behave similarly.

The intensities of the diagonal and the cross peaks are expressed in the equations [2.1.23]-[2.1.26].

\[
I_{z\text{diagonal}}(t) = \frac{I_0 - I_i}{2} (e^{\sigma t} + e^{-\sigma t}) e^{-\rho t} \tag{2.1.23}
\]

\[
S_{z\text{diagonal}}(t) = \frac{S_0 - S_i}{2} (e^{\sigma t} + e^{-\sigma t}) e^{-\rho t} \tag{2.1.24}
\]

\[
I_{z\text{cross}}(t) = \frac{S_0}{2} (e^{\sigma t} - e^{-\sigma t}) e^{-\rho t} \tag{2.1.25}
\]

\[
S_{z\text{cross}}(t) = \frac{I_0}{2} (e^{\sigma t} - e^{-\sigma t}) e^{-\rho t} \tag{2.1.26}
\]

From the equations [2.1.25] and [2.1.26], the intensities of the cross peaks are proven to experience the changes from 0 at \( t=0 \) to a maximum at \( t = \frac{1}{2\sigma} \ln \frac{\rho + \sigma}{\rho - \sigma} \) and approach to 0 at \( t=\infty \). The \( I_{z\text{diagonal}}(t) \) and \( S_{z\text{cross}}(t) \) correspond to \( g(t) \) and \( f(t) \) in expressions (E) and (F), respectively.

-32-
Since $\sigma$ is a factor related to distance and the correlation time $\tau_c$, the intensity change of cross peaks depends on the distance and the correlation time $\tau_c$ by equation [2.1.20]. The more explicit relationship between intensities of cross peaks and the distances is expressed in the equations [2.1.27] and [2.1.28].

\[ I_{\text{cross}}'(t) = -\frac{\rho S_0}{2} (e^{\sigma t} - e^{-\sigma t}) e^{-\rho t} + \frac{\sigma S_0}{2} (e^{\sigma t} + e^{-\sigma t}) e^{-\rho t} \]  

[2.1.27]

\[ S_{\text{cross}}'(t) = -\frac{\rho I_0}{2} (e^{\sigma t} - e^{-\sigma t}) e^{-\rho t} + \frac{\sigma I_0}{2} (e^{\sigma t} + e^{-\sigma t}) e^{-\rho t} \]  

[2.1.28]

At time $t = 0$, if $\omega^2 \tau_2^{-2} \ll 1$ in equation [2.1.20], $I_{\text{cross}}'$ and $S_{\text{cross}}'$ are functions of internuclear distance $r$. It is expected from equations [2.1.29] and [2.1.30] that the intensity changes of the NOE cross peaks at time $t=0$ are inversely proportional to the distance to the power of six. Thus, the initial rate approximation is carried on to calculate the internuclear distance (Wüthrich 1986, Kumar et al., 1981).

The similar calculation can be used to predict the changes of the diagonal peak. The differentials of equations [2.1.23] and [2.1.24] are calculated.

\[ I_{\text{diagonal}}'(t) = -\frac{\rho I_0}{2} (e^{\sigma t} + e^{-\sigma t}) e^{-\rho t} + \frac{\sigma I_0}{2} (e^{\sigma t} - e^{-\sigma t}) e^{-\rho t} \]

\[ S_{\text{diagonal}}'(t) = -\frac{\rho S_0}{2} (e^{\sigma t} + e^{-\sigma t}) e^{-\rho t} + \frac{\sigma S_0}{2} (e^{\sigma t} - e^{-\sigma t}) e^{-\rho t} \]

After rearranging the above equations,
\[ I_{\text{cross}}'(t) = (-\rho + \sigma) \frac{I_0}{2} e^{\sigma t - \rho t} - (\rho + \sigma) \frac{I_0}{2} e^{-\sigma t + \rho t} \]  
\[ S_{\text{cross}}'(t) = (-\rho + \sigma) \frac{S_0}{2} e^{\sigma t - \rho t} - (\rho + \sigma) \frac{S_0}{2} e^{-\sigma t + \rho t} \]  

Since \( \rho = W_0 + 2W_{11} + W_2 > \sigma = W_2 - W_0 \), \( I_{\text{cross}}'(t) \) and \( S_{\text{cross}}'(t) \) are always negative. This means that the diagonal peaks decrease during the mixing time \( t \). The decrease has been observed in NOESY experiments (Kumar et al., 1981).

2.1.5 Phase cycling

The central concept to understand this section is coherence. In a quantum mechanical model, a coherence \( \rho_{rs} \) is an off-diagonal element of a density matrix.

\[ \rho_{rs} = < r | \rho(t) | s > \]

where \( \rho(t) \) is the density operator and \( < r | \) and \( | s > \) are eigenstates. In another words, a coherence is related to a transition between two eigenstates. If \( M_r \) and \( M_s \) are the magnetic quantum numbers of eigenstates \( < r | \) and \( | s > \), the coherence order is given by \( P = M_r - M_s \). For an allowed transition with \( M_r - M_s = \pm 1 \), the coherence \( \rho_{rs} \) is related to the transverse magnetization, another coherence with \( \rho \neq \pm 1 \) does not give observable magnetizations (Ernst et al., 1987). Coherence orders can be changed by pulses, in NMR experiments, from 0 in a thermal equilibrium to -1 which is the end coherence order to be detected by quadrature detection. Different NMR experiment has different way to change coherence orders, which is called coherence pathway. For example, in a COSY experiment, coherence pathway is from 0 to +1 to -1, but in a NOESY, the pathway is from 0 to +1 to 0 to -1.

Besides the width and the intensity of a pulse, the phase of a pulse is important, especially in a multipulse experiment. The way to systematically change the phases of
pulses and the receiver is specified by phase cycling programs. The phase cycling helps to suppress artifacts due to imperfect pulses and it also serves to select wanted coherence pathways from other possible pathways. The coherences selected by a pulse with a phase cycle of $\varphi$, where $\varphi = 2\pi/N$ and $N$ is an integer, are indicated by $\Delta p \pm N$, where $\Delta p$ is the minimum coherence difference (Bodenhausen et al., 1984). It is possible to combine the phase cycling into operator analyses of COSY and NOESY experiments. To be simple and straightforward, I will use Figure 2.7 to show the use of phase cycling.

Figure 2.7 is a pulse scheme of a NOESY experiment along with coherence transfer pathways. The first pulse generates +1, -1 coherences no matter what phase it has. The second pulse is used to select +1 or -1 coherence changes and suppress others. If we choose $\Delta p = +1$, all coherences we want are +1 and -1. To fit to the formula $\Delta p \pm N$, the $N$ must be 2. So the phase increment is $\pi$. It needs a minimal 180° phase cycle. The -1 coherence is chosen from 1, 0, and -2 by the third pulse with a minimal cycle of 90°. So the basic phase cycling for NOESY experiments includes 8 combinations of phases of the second and third pulses. If we define the X axis is 0, the Y, -X, and -Y correspond to 1, 2, and 3, respectively.
2.1.6 Protein 3-D structure and 2-D NMR spectra

Basically, two strategies were developed to solve protein structures using 2-D NMR. The similarities between the two methods are using COSY spectra to assign chemical shifts to protons and using NOESY experiments to estimate relative distances between protons.

2.1.6.1 Sequential assignment strategy

A complete structural elucidation of protein conformations by NMR studies involves two steps. The first step is NMR resonance assignments, the second one is called "structure determination by distance geometry" (Wüthrich, 1986).

2.1.6.1.1 Side chain spin-system and sequential assignments

Protons in a protein are divided into four groups according to their chemical shifts: amide protons, aromatic protons, α-protons, and side chain protons (Gross and Kalbitzer 1988). All side chain protons form a spin-system which shows a certain cross peak pattern in COSY experiments. Twenty amino acids have twenty different side chains (Figure 2.8) resulting in several major COSY cross peak patterns (Figure 2.9). The identification of spin-systems is achieved by recognizing cross peak patterns among numbers of cross peaks in a COSY spectrum. From scalar couplings, the backbone protons (α-protons and amide protons) are connected to the spin-system through the CβH-CαH-NH cross peaks in COSY spectra and assigned to a type of amino acids (Chazin et al., 1988a). NOESY shows the sequential connections by NOE cross peaks between CαH(i)-NH(i+1) and NH(i)-NH(i+1). Using information from both COSY and NOESY spectra, one can link single amino acid residues together into a fragment (Wagner et al., 1981, Figure 2.10). By
comparison of these fragments with the primary sequence of the protein, one would assign
back-bone and side chain protons in the protein to chemical shift values.

H
Gly, G
AX
Ala, A
A3X
Ser, S
AMX
Cys, C
AMX
CHI
CHI
\%
\%
\%
Val, V
A3B3MX
Thr, T
A3MX
Asp, D
AMX
Asn, N
AMX
Leu, L
A3B3MPX
Ile, I
A3MPTX(B3)X
Lys, K
A2(F2T2)MPXc
Arg, R
A2(T2)MPXc

Figure 2.8 Side chains of 20 amino acids. (From Wüthrich 1986).

Figure 2.9 Spin system of 20 amino acids. (From Wüthrich 1986).
2.1.6.1.2 Conformation determinations

The well documented secondary structures of a protein are: α-helices, β-sheets, tight turns, and random coils (Table 2.1). Based on the relationship between NOE cross peak patterns and those secondary structures (Wüthrich et al., 1984), we are able to recognize each structure from the NOE patterns. In the sequential backbone-backbone NOEs, the extended regions of strong CaH(i)-NH(i+1) NOEs indicate that the polypeptide adopts an extended conformation such as that found in a β-sheet. On the other hand, an extended run of strong NH(i)-NH(i+1) NOEs may indicate a helical structure. Also, from the medium and long range backbone-backbone NOEs, both parallel and antiparallel β-sheet structures are associated with long-range CaH(i)-NH(j) and NH(i)-NH(j) NOEs.
where residues \(i\) and \(j\) are usually more than five residues apart in the primary sequence. Antiparallel \(\beta\)-sheet structures are also characterized by \(\text{CaH}(i)-\text{CaH}(j)\) NOEs. (Driscoll et al., 1990, Williamson et al., 1984, Arrowsmith et al., 1990, Braun et al., 1983, Chazin et al., 1988b, Khan et al., 1990, Gao et al., 1990, Saudek and Pelton 1990, Eberle et al., 1990). \(\alpha\)-helices are identified by the presence of medium range NOEs (Table 2.1).

![Table 2.1 Sequential and medium \(^1\text{H}-\text{H}\) NOE in common secondary structures of proteins. (From Wüthrich 1986)]

Since NOE cross peaks in NOESY spectra originate from cross relaxation, the intensities of cross peaks depend on the mixing time \(\tau_m\). If we do a series of NOESY experiments with increasing delay \(\tau_m\) we would expect the intensities of NOE peaks to be functions of \(\tau_m\). This is a conventional method to distinguish NOE cross peaks from artifacts. Each real NOE cross peak, indicating distance between two nuclei, provides a constraint (\(\leq 5\ \text{Å}\)). To obtain a 3-dimensional structure of a protein, the distance information obtained from NOEs is used as input for distance geometry calculations. The precision of the model can be improved upon by increasing the number of constraints.
2.1.6.2 Main-chain-directed strategy

Practically, assignments of side chain protons are complicated due to chemical shift degeneracies and peak overlappings. Besides in a structured protein, the spin systems may be altered by the local environment in the protein and the intensities of cross peaks may vary due to some other factors, such as relaxations and phase of lines. A new strategy has been developed by Englander and Wand (1987). It is called a main-chain-directed strategy. Unlike the sequential strategy, the main chain directed approach is focused on the NH-αH-βH connections. Instead of working on the crowded side chain regions, one can concentrate on the NH-αH region in a COSY spectrum and NH-αH-βH region in a relay-COSY spectrum. This is relatively easy to do, compared to the assignments of the side chains, due to the simple and linear connections of the NH-αH-βH subspin system.

A series of analyses of known proteins was carried out (Englander and Wand 1987). A library of NOE peak patterns of main chains (NH-αH-βH) of amino acids for commonly observed secondary structures was established. In a helix, strong cross peaks between NH to βHs of same and adjacent residues should be observed besides well known NH-NH cross peaks. In an antiparallel β-sheet, intrastrand NOE should be strong between αHs and NHs. In addition, an "inner loop" is defined as α\textsubscript{H}\textsubscript{i}-NH\textsubscript{i+1}-α\textsubscript{H}\textsubscript{j}-NH\textsubscript{j+1} and an "outer loop" as (αH-NH)\textsubscript{i}-(αH-NH)\textsubscript{i+1}-NH\textsubscript{j+1}-(αH-NH)\textsubscript{i}. NOEs from both "inner loop" and "outer loop" should be observable for an antiparallel β-sheet. In a parallel β-sheet, a single loop of (NH-αH)\textsubscript{i}-(NH-αH)\textsubscript{i+1}-NH\textsubscript{j+1}-α\textsubscript{H}\textsubscript{j}-(NH-αH)\textsubscript{i} is expected besides the inter- and intrastrand NOEs.
2.2 Fourier Transform Infrared Spectrum

2.2.1 Interferogram

The light paths of an infrared spectrometer are depicted in Figure 2.11.a. Two light beams reflected from mirror 1 and mirror 2, respectively, will interfere on the detector D. When the difference between two light paths \( \Delta d (=d_2-d_1) \) equals multiplications of the wave length of the light, a constructive interferogram is obtained. On the other hand, a half wavelength difference will result in a destructive interferogram. The signal detected is a function of the position of the mirror 2 as in Figure 2.11.b. A interferogram is obtained by digitization of the signal from the detector D. The resulted interferogram is a \( \cos \) wave function without amplitude modulations. Theoretically, a spectrum is a result of Fourier transformation of such an interferogram.

Practically, a interferogram is truncated due to three instrumental factors: the dynamic range of the digitization, the S/N ratio of the spectrum and the instrumental resolution used to record the spectrum (Kauppinen et al., 1981a). A truncated interferogram can be decomposed into a \( \cos \) wave function with an amplitude modulation of a boxcar function. Fourier transform of such a interferogram yields a spectrum with large side-lobes because of the truncation effect.

Two of the key steps in the process of FTIR spectra and extracting information from a FTIR spectrum are the apodization and Fourier self-deconvolution.
Figure 2.11 The light paths of an infrared spectrometer (a) and the signal detected from the detector D (b). The S is a light source. D is a detector. The mirror 2 is a movable mirror. \( \Delta d \) is the distance difference between the two light paths \( d_1 \) and \( d_2 \).

2.2.2 Apodization

The convolution theorem states that, mathematically, Fourier transform of a product of two functions is equal to the convolution of Fourier transform of the two functions. An IR spectrum, \( E(u) \), in the frequency domain is a result of Fourier transformation of an interferogram, \( I(x) \), in the distance domain which originally is a product of a wave function of infinity extent, \( w(x) \), and a boxcar function, \( b(x) \).

\[
E(u) = F\{I(x)\} = F\{w(x) \cdot b(x)\} \tag{2.2.1}
\]

\( F \) indicates a Fourier transform. Therefore, the Fourier transform of an interferogram equals the convolution of Fourier transform of the wave function and the boxcar function. If \( \delta(x) = F\{w(x)\}; B(u) = F\{b(x)\} \), then

\[
E(u) = F\{w(x) \cdot b(x)\} = \int \delta(u') \cdot B(u-u') \, du' \tag{2.2.2}
\]

The Fourier transformation of a wave function gives a \( \delta \) function in the frequency domain which is zero everywhere except at the value of the wave frequency while the Fourier transform of the boxcar function produces a \( \text{sinc} \) function in the frequency domain.
which is a function with a non-zero value at zero and side-lobes in either sides of the peak. An IR spectrum is constituted by the convolution of the $\delta$ function and the sinc function. It has a peak at the wave frequency and the shape of the sinc function (Green and Reedy 1978; Kauppinen et al., 1981a). In an apodization process, a boxcar function is replaced by an apodization function, $s(x)$, in a convolution calculation in order to obtain an interferogram which has small truncation effects in the distance domain and minimal side lobes in the frequency domain after Fourier transformations.

Several apodization functions have been studied (Kauppinen et al., 1981b). The commonly used functions are: trapezoidal, triangular, triangular square, bessel, cos, sinc square, and gaussian (Kauppinen et al., 1981b). The default apodization function on the Bruker IFS 85 FTIR spectrometer is the Blackmann-Harris three-term function which has a form $I(\text{new}) = I(\text{old})(A + B \cos(\theta) + C \cos(2\theta) + D \cos(3\theta))$ (Section 8, Bruker IFS Users Manual).

2.2.3 Self-deconvolution

An infrared spectrum is often composed of more than one component bands and it is not practically applicable to identify each band directly from the original spectrum. A special procedure, which is called self-deconvolution, is required to find all bands and their positions. As stated above, an interferogram can be a product of a wave function and a line-shape function (apodization function). A spectrum is a convolution of Fourier transform of the wave function and the line-shape function (equation [2.2.1]). The wave function, after the Fourier transform from the distance domain to the frequency domain, specifies the frequency feature of the spectrum ($w(u)$ in equation [2.2.2]) while the line-shape function defines the shape of the spectrum ($B(u)$ in equation [2.2.2]). A fast decaying line-shape function gives a broad line after Fourier transformation while a slow decay line-shape function generates a sharp line. The self-deconvolution algorithm changes a fast decay line-
shape function to a slow decay line-shape function to achieve the line narrowing effect (Surewicz and Mantsch 1988). Doing an inverse-Fourier transform (F⁻¹) on both sides of equation [2.2.1], and replacing \( b(x) \) by \( s(x) \), one can obtain equation [2.2.3].

\[
F^{-1} \{ E(u) \} = w(x) \cdot s(x) \tag{2.2.3}
\]

To recover the pure wave function, \( 1/s(x) \) must be multiplied to both sides of equation [2.2.3]. A new shape function with a slower decay rate is multiplied to the recovered pure wave function to get a new interferogram \( I'(x) \).

\[
I'(x) = F^{-1} \{ E(u) \} \cdot \frac{s'(x)}{s(x)} \tag{2.2.4}
\]

A new spectrum with narrower lines can be obtained by Fourier transform of equation [2.2.4].

\[
E'(u) = F \{ I'(x) \} \tag{2.2.5}
\]

As a result of such an operation, a spectrum in the frequency domain composed of sharp lines can be achieved. When an analysis of an IR spectrum becomes difficult due to a severe overlapping of lines, self-deconvolution is used to reveal the components under the broad peaks (Kauppinen et al., 1981a, b).

Practically, the line-shape function is assumed to be an exponential decay function. The Fourier self-deconvolution consists of four steps. Firstly, an interferogram is computed using an inverse Fourier transform; secondly, an exponentially decreasing function with a half-width at half-height (HWHH) of \( \gamma \) is multiplied with the interferogram to neutralize the decay effect and get a pure wave function; thirdly, a new exponentially decaying function with a new HWHH of \( \gamma = k\gamma' \) is multiplied with the pure wave function, where \( \gamma \) and \( k \) are named HAHH and resolution enhancement factor, respectively; lastly, the Fourier transform is performed on the product of the pure wave function and the
exponential decay function (Surewicz and Mantsch, 1988). The resulting spectrum has a new line width $\gamma$ (Mantsch et al., 1988).

2.2.4. The amide I band and the secondary structure of proteins

2.2.4.1 Fitting

For a given spectrum, all components contained in it can be revealed using Fourier self-deconvolution. Practically, a Gaussian line shape (Byler and Susi, 1986) or a combination of Gaussian and Lorentzian line shape is assumed for all component bands. The numbers and positions of bands were obtained from deconvoluted spectra. The band positions, intensities and band widths were input into a fitting program written by myself and the intensities and the widths of bands were iterated in typical fitting processes. The root mean square of differences between the original spectrum and the sum of component bands is calculated as an indicator of the fitting quality. The area of each component band is computed.

2.2.4.2 Band assignments

Three regions in an IR spectrum of proteins, amide I, amide II, and amide III, have been studied (Koenig and Tabb 1980, Wasacz et al., 1987, Kaiden et al., 1987). The amide I, between 1600 cm$^{-1}$-1700 cm$^{-1}$, represents mainly the C=O stretching vibrations of the amide groups coupled to the in-plane NH bending and CN stretching modes. The frequency of the vibration depends on the nature of the hydrogen bonds connecting C=O and NH which are determined by the secondary structure of a protein. The relationship between the position of the amide I bands and the type of secondary structures has been studied using IR spectra of homopolypeptides (Susi et al., 1967, Timasheff et al., 1967, Dwivedi and Krimm, 1984, Sengupta and Krimm, 1985). The assignments of specific conformations to experimentally observed amide I component bands have been established.
using normal coordinate calculations (Krimm and Bandekar, 1986) and the analysis of the infrared spectra of proteins with known three-dimensional structures (Byler and Susi, 1986, Surewicz and Mantsch, 1988). Commonly, α-helix structures are represented by bands between 1650 cm⁻¹ and 1658 cm⁻¹, which is consistent with theoretical calculations (Krimm and Bandekar, 1986). An "unordered" conformation is often associated with bands around 1644 cm⁻¹ in IR spectra recorded in deuterium oxide. The turns are believed to show a highly characteristic band around 1665 cm⁻¹ as well as all bands from 1670 cm⁻¹ to 1695 cm⁻¹ except those bands around 1675 cm⁻¹. Infrared bands between 1620 cm⁻¹ and 1640 cm⁻¹ are assigned to β-sheets along with bands around 1675 cm⁻¹. The low wavenumber bands (below 1620 cm⁻¹) have been studied as indicators of intermolecular β-structures (Carrier et al., 1990, Surewicz et al., 1987, Muga et al., 1990) and side chain absorptions (Byler and Susi, 1986).

2.2.4.3 Structural studies

Under the assumption that the area of the fitted component bands is directly related to the relative populations of the conformational structures represented by these components (Byler and Susi 1986, Surewicz and Mantsch 1988), the secondary structural of the protein is expressed as a percentage of the areas associated with that structure.
Chapter 3. Materials and Methods

3.1 Materials

C-PEPTIDE and LAP-20 were purchased from the University of Victoria Microsequencing Lab. KIT, SH3, and B19 were kindly provided by Dr. Ian Clark-Lewis, University of British Columbia. The sequences of the peptides are given in Table 3.1. Dimyristoylphosphatidylcholine (DMPC) was purchased from Sigma Chemical Co.. Deuterium Oxide (99% atom D) was purchased from Isotec, Inc.. Deuterated dodecylphosphocholine (DPC-d38) was a kind gift from Dr. K. Wüthrich, Eidenössische Technische Hochschule.

Table 3.1 Sequences of the peptides

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-PEPTIDE</td>
<td>Ser&lt;sup&gt;1&lt;/sup&gt;-Ala-Lys-Met-Arg&lt;sup&gt;5&lt;/sup&gt;-Glu-Trp-Phe-Ser-Glu&lt;sup&gt;10&lt;/sup&gt;-Thr-Phe-Gln-Lys-Val&lt;sup&gt;15&lt;/sup&gt;-Lys-Glu-Lys-Leu-COOH</td>
</tr>
<tr>
<td>LAP-20</td>
<td>Val&lt;sup&gt;1&lt;/sup&gt;-Ser-Ser-Leu-Leu&lt;sup&gt;5&lt;/sup&gt;-Ser-Ser-Leu-Lys-Glu&lt;sup&gt;10&lt;/sup&gt;-Tyr-Trp-Ser-Ser-Leu&lt;sup&gt;15&lt;/sup&gt;-Lys-Glu-Ser-Phe-Ser&lt;sup&gt;20&lt;/sup&gt;-COOH</td>
</tr>
<tr>
<td>KIT</td>
<td>Cys&lt;sup&gt;1&lt;/sup&gt;-Ala-ALA-Gln-Arg&lt;sup&gt;5&lt;/sup&gt;-Asp-Gly-Thr-Trp-Leu&lt;sup&gt;10&lt;/sup&gt;-His-Ser-Asp-Lys-Phe&lt;sup&gt;15&lt;/sup&gt;-Thr-Leu-Lys-Val-Arg&lt;sup&gt;20&lt;/sup&gt;-Glul-Ala-Ile-Lys-Ala&lt;sup&gt;25&lt;/sup&gt;-Ile-Pro-NH&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
<tr>
<td>SH3</td>
<td>Ala&lt;sup&gt;1&lt;/sup&gt;-Leu-Tyr-Asp-Tyr&lt;sup&gt;5&lt;/sup&gt;-Glu-Ser-Trp-Thr-Glu&lt;sup&gt;10&lt;/sup&gt;-Thr-Asp-Leu-Ser-Phe&lt;sup&gt;15&lt;/sup&gt;-Lys-Lys-Gly-Glu-Arg&lt;sup&gt;20&lt;/sup&gt;-Leu-Gly-Gly-Cys-NH&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
<tr>
<td>B19</td>
<td>Pro&lt;sup&gt;1&lt;/sup&gt;-Asn-Thr-Lys-Asp&lt;sup&gt;5&lt;/sup&gt;-Ile-Asp-Asn-Val-Glu&lt;sup&gt;10&lt;/sup&gt;-Phe-Lys-Tyr-Leu-Thr&lt;sup&gt;15&lt;/sup&gt;-Arg-Tyr-Glu-Gln-His&lt;sup&gt;20&lt;/sup&gt;-Val-Ile-Arg-Met-Leu&lt;sup&gt;25&lt;/sup&gt;-Arg-Leu-Cys-NH&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
</tbody>
</table>
3.2 Methods

3.2.1 NMR Spectra

All the NMR sample solutions were filtered through a 2 mm filter into a NMR tube and repeatedly freeze-thawed under nitrogen flow.

All spectra was collected at 37 °C, on a Bruker 400 MHz NMR spectrometer equipped with a AMX32 computer. All experiments were recorded in the pure-phase absorption mode (Ernst et al., 1987), using either time proportional phase increments (Marion and Wüthrich 1983) or the method reported by States et al., (1982). Typically, a presaturation soft pulse was used to suppress the strong water signal in all experiments and a total of 32-196 transients and 240-512 \( t_1 \) increments of 2K data points were recorded. Digital resolutions of free induction decays (FID) were 1.9-2.2 Hz. A normal processing procedure was followed. Spectra were zero-filled to 1K in the \( t_1 \) dimension. Prior to Fourier transformation, FIDs were apodized using a shifted sine-bell or squared sine bell weighting function in both \( t_1 \) and \( t_2 \) dimensions. Double Fourier transforms were performed. The final spectra were stored in a 1K x 1K real data matrix.

3.2.1.1 C-PEPTIDE

A total of 14.6 mg of free C-PEPTIDE was dissolved in 0.4 ml H\(_2\)O (9:1, v/v). The pH of the solution was 2.5. The sample was refereed to as "C-PEPTIDE alone in solution". For the experiments of micelles, 16.4 mg of C-PEPTIDE was dissolved in 0.4 ml H\(_2\)O (9:1, v/v), then 92 mg of perdeuterated dodecylphosphocholine (DPC-d38) was added to the solution. The molar ratio of peptide to detergent was approximately 1 to 5. The pH of the complex solution was adjusted to 2.6 by addition of deuterated acetic acid. The sample
was termed "C-PEPTIDE in complexes". For the C-PEPTIDE alone in solution, a TOCSY experiment with a spin-lock time of 69 ms and a radio frequency (rf) strength of 10.1 kHz, and three NOESYs with mixing times of 150, 350, 500 ms were recorded. For C-PEPTIDE in complexes, a DQF-COSY experiment, a TOCSY experiment with a spin-lock time of 70 ms and three NOESY experiments with mixing times of 75, 150, 300 ms were collected. The rf strength in the TOCSY experiments was 9.9 KHz as suggested by Bruker's consultants. A 2 s presaturation soft pulse was used to suppress the water signal in the experiments. Normal processing procedure was followed.

### 3.2.1.2 KIT

15 mg KIT was dissolved in 0.5 ml H₂O (9:1, v/v). Two TOCSY spectra were recorded with spin-lock times of 100 ms and 70 ms. The radio frequency (rf) field strengths were 7.7 KHz and 8.8 KHz, respectively. A DQF-COSY experiment was recorded under the typical conditions. Four NOESY experiments were collected with mixing times of 100 ms, 300 ms, 450 ms, and 550 ms. Spectra were processed following the normal processing procedure.

### 3.2.2 Assignment strategies

The general strategy introduced by Wüthrich (1986) was followed. The two crucial steps were the spin system assignments and the sequence-specific assignments. The spin system for each amino acid was identified using DQF-COSY and TOCSY spectra in H₂O (9:1, v/v). The sequence-specific assignments were achieved combining the information from NOESY and TOCSY spectra.
3.2.2.1 Spin systems

Amino acids were divided into three categories according to the length of side chains. The long side chain group included Lys, Leu, Ile, Arg, and Pro. The medium side chain group included Glu, Gln, Met, Val, and Thr. The rest of the amino acids belonged to the short chain group. They were Gly, Phe, Tyr, Trp, His, Asp, Asn, Cys, Val, and Ser.

For the amino acids in the first group three or more relay steps were observed. Isoleucines showed five peak patterns corresponding to all of the side chain protons. Leucines had one peak for the two methyl groups and one overlapping peak for the $\beta$ and the $\gamma$ protons. Lysines were identified as a four peak pattern. Arginines, with a three peak pattern, were assigned in a manner similar to the lysines.

Valines were recognized by the relays from the $\alpha$ proton to the $\beta$ methylene and further to separated methyl protons in a TOCSY spectrum. Threonines showed a unique pattern with two peaks in a group in the $\alpha$ proton range and one peak in the methylene range in a TOCSY spectrum. The methionine, glutamate and glutamine were recognized by the relays from $\alpha$-protons to $\beta$-protons. Additional information from NOESY spectra were needed to distinguish among the three.

Alanines were identified by a single $\alpha$ proton signal with a long relay to methyl protons along the diagonal. Glycines could be distinguished by a pattern of two $\alpha$ protons. The rest of the amino acids in the third group were difficult to assign from TOCSY spectra alone. They were assigned using the information from NOESY, TOCSY and DQF-COSY spectra.
3.2.2.2 Structural studies

Two protein stereo-structures were characterized by NOE patterns. Strong NH-NH NOEs indicated α-helices while strong αH-NH NOEs indicated β-sheets. Redundant shifts and shifts close to their canonical values were taken as evidences of unstructured peptide. This usually resulted in a crowded spectrum. The two features of a crowded spectrum were severe overlaps in the αH-NH region and overlaps for amino acids with the same molecular structure.

3.2.3 Proton exchange experiments

C-PEPTIDE solution in H$_2$O:D$_2$O (9:1, v/v) and C-PEPTIDE with DPC-d38 in H$_2$O:D$_2$O (9:1, v/v) were lyophilized and redissolved in 0.4 ml D$_2$O (99%), separately. The first 1-D NMR spectra were recorded on 400 MHz Bruker NMR spectrometer within 7 minutes after the addition of D$_2$O. Sixteen scans and 32 scans were collected for C-PEPTIDE in solution and C-PEPTIDE in complexes, respectively. The spectra, with a 10 ppm sweep width, were stored in 32 K complex data points. For C-PEPTIDE in solution, the time course was 7, 16, 42, and 90 minutes, while for C-PEPTIDE in complexes the spectra were taken at 7, 25, 46, and 108 minutes after the addition of D$_2$O.

3.2.4 Secondary structure predictions

3.2.4.1 Secondary shift of peptides

Secondary shifts were calculated as the difference between standards and experimental measurements. The canonical values of amino acids (Table 3.1) were used as standards which were obtained from unstructured peptides (Wüthrich 1986). Upfield secondary shifts indicated a helical structure (positive secondary shifts) while downfield
secondary shifts implied a β-like structure (negative secondary shifts). Random changes of the secondary shifts were deemed to be an unstructured conformation.

Secondary shifts had a second use when studying the conformation of C-PEPTIDE. Comparison of the chemical shifts of C-PEPTIDE in solution with that of C-PEPTIDE in complexes was used to determine the binding domain of C-PEPTIDE, i.e. the chemical shift difference of C-PEPTIDE in complexes from chemical shifts of C-PEPTIDE in solution. The binding domain should be the segment with the largest secondary shift value for each residue in the domain.

Table 3.2  Canonical values of chemical shifts of 20 common amino acids (Wüthrich 1986)

<table>
<thead>
<tr>
<th>Residue</th>
<th>NH (ppm)</th>
<th>αH (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>8.25</td>
<td>4.35</td>
</tr>
<tr>
<td>Arginine</td>
<td>8.27</td>
<td>4.38</td>
</tr>
<tr>
<td>Asparagine</td>
<td>8.41</td>
<td>4.76</td>
</tr>
<tr>
<td>Aspartate</td>
<td>8.31</td>
<td>4.69</td>
</tr>
<tr>
<td>Cysteine</td>
<td>8.37</td>
<td>4.29</td>
</tr>
<tr>
<td>Glutamine</td>
<td>8.41</td>
<td>4.37</td>
</tr>
<tr>
<td>Glycine</td>
<td>8.39</td>
<td>3.97</td>
</tr>
<tr>
<td>Histidine</td>
<td>8.41</td>
<td>4.63</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>8.19</td>
<td>4.23</td>
</tr>
<tr>
<td>Leucine</td>
<td>8.42</td>
<td>4.38</td>
</tr>
<tr>
<td>Lysine</td>
<td>8.41</td>
<td>4.36</td>
</tr>
<tr>
<td>Methionine</td>
<td>8.42</td>
<td>4.52</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>8.23</td>
<td>4.66</td>
</tr>
<tr>
<td>Proline</td>
<td>8.09</td>
<td>4.70</td>
</tr>
<tr>
<td>Serine</td>
<td>8.38</td>
<td>4.50</td>
</tr>
<tr>
<td>Threonine</td>
<td>8.24</td>
<td>4.35</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>8.09</td>
<td>4.70</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>8.18</td>
<td>4.60</td>
</tr>
<tr>
<td>Valine</td>
<td>8.44</td>
<td>4.18</td>
</tr>
</tbody>
</table>
3.2.4.2 Hydrophobic moment

The hydrophobicity index of a protein was calculated as a mean hydrophobicity per residue for the non polar "face" in an amino acid sequence. The hydrophobicity of an amino acid was defined originally as the free energy of transfer of the amino acid from ethanol to water. The hydrophobicity scale used by Segrest and colleagues, and this study, was as follows: Trp, 6.5; Phe, 5.0; Ile, 5.0; Tyr, 4.5; Leu, 3.5; Val, 3.0; Met, 2.5; Pro, 1.5; Ala, 1.0; His, 1.0; Thr, 0.5; Gly, 0; Cys, 0; Ser, -0.5; Glu, -1.0; Asn, -1.5 (Segrest and Feldmann 1977).

The hydrophobic moment of an amino acid was defined as a vector in the direction from the α-carbon toward the geometric center of the side chain. The hydrophobic moment of a protein, containing n residues, was the vectorial sum of the hydrophobic moments of the individual amino acid.

\[
\mu = \left( \left[ \sum H_n \sin(\delta n) \right]^2 + \left[ \sum H_n \sin(\delta n) \right]^2 \right)^{1/2}
\]

where \( \mu \) was the length of the hydrophobic moment which equalled the vectorial sum of the components of the hydrophobicity of the individual amino acids \( H_n \). \( \delta \) was defined as the angle (in radians) between successive side chains emerging from the backbone. For an α-helix, \( \delta \) was \((10 \times p/18 \ \text{radians}) \) or 100°, and for a β-structure, \( \delta \) was predicted to be between 160° to 180°. The hydrophobic scale \( H_n \) was derived from the free energy required to transfer amino acids from aqueous to a hydrophobic environment and was as follows: Ile, 0.73; Phe, 0.61; Val, 0.54; Leu, 0.53; Trp, 0.37; Met, 0.26; Ala, 0.25; Gly, 0.16; Cys, 0.04; Tyr, 0.02; Pro, -0.07; Thr, -0.18; Ser, -0.26; His, -0.40; Glu, -0.62; Asn, -0.64; Glu, -0.69; Asp, -0.72; Lys, -1.10; Arg, -1.76 (Eisenberg et al., 1984).
3.2.5 FTIR studies

The FTIR studies were carried out at room temperature on the Bruker IRS 85 FTIR spectrometer equipped with the Aspect 2000 computer. The sample solution was filled into the 25 μm gap between two BaF₂ windows. For each spectrum, 800 scans were collected at room temperature with constantly purging using dried air. A 2 cm⁻¹ resolution was achieved in a range from 4000 cm⁻¹ to 400 cm⁻¹. Additional spectra were recorded, under the identical conditions, for background, water vapor, D₂O solvent, and the DMPC mixture.

3.2.5.1 Sample preparation

LAP-20, KIT, SH3, and B19 were dissolved in 0.5-1.0 ml D₂O, separately. The final concentrations were approximately 12 mg/ml for LAP-20 and B19, 13.5 mg/ml for SH3, and 15.2 mg/ml for KIT. For the experiments of lipid-peptide complexes, a total of 150 mg solid DMPC was mixed with 5 ml D₂O to prepare a multilamellar solution. The peptide solutions were lyopholyzed, after collection of the IR spectra, and the solid peptides were redissolved in 0.5-1.0 ml DMPC-D₂O mixture to achieve the same peptide concentrations as for the peptides alone. The ratio of peptide to DMPC was approximately 1 to 2.5.

3.2.5.2 FTIR spectra

Single beam spectra of samples, the water vapor, the D₂O solvent, and the DMPC solution were determined between 2000 cm⁻¹ to 1000 cm⁻¹ against the single beam spectrum of background using Bruker's standard program ATS88W. Absorption spectra of peptides were calculated as the differences between absorptions of peptides and the absorption of the D₂O solvent. Both of the absorptions of D₂O and DMPC were subtracted from the
absorption spectra of the DMPC-peptide complexes to obtain the spectra of the peptides in complexes. Repeated subtraction of the water vapor peaks using the second derivative of the water vapor spectrum were performed for all samples until a smooth spectral curve were achieved.

3.2.5.3 Deconvolution and curve fitting

The component bands of the spectrum between 1700 cm\(^{-1}\) to 1600 cm\(^{-1}\) were revealed using Bruker's self-Fourier deconvolution program (named FOUDECONV). Typically, a half-linewidth of about 20 cm\(^{-1}\) and an enhancement factor (k) of 2.0-2.3 were used. The second derivative spectra were calculated for a three point window using a Pascal programme that I wrote (see Appendix). The root mean square (RMS) value was calculated as the square root of an average of the square of the differences between original and fitted curves. The linewidths and the intensities of the components were iterated to minimize the RMS when doing the curve fitting (see Section 2.2.4.1).

The component areas were computed. The peptide secondary structure was expressed as the relative percentage of the areas for the components associated with the structure to the total area of the amide I band (i.e. bands associated with side chain absorption were not included).

3.2.5.4 Assignments for the deconvoluted bands

Band assignments were based on the commonly used results (Surewicz and Mantsch, 1988, Byler and Susi, 1986). Two well-defined structures, \(\alpha\)-helix and \(\beta\)-sheet, were observed between 1658 to 1650 cm\(^{-1}\) and 1640 to 1620 cm\(^{-1}\), respectively. Bands between 1648 cm\(^{-1}\) and 1640 cm\(^{-1}\) were assigned to random structures. High-frequency components from 1695 cm\(^{-1}\) to 1660 cm\(^{-1}\) were assigned to turns except those around 1675 cm\(^{-1}\), which were assigned to \(\beta\)-structure. For the bands below 1620 cm\(^{-1}\), the assignments
were either intermolecular β-sheet or side chains depending on other information. Intermolecular β-sheet exists in solution but not in complexes.

3.2.6 Electron Microscopy

The electron micrographs were taken on the Philips 300 electron microscope. Samples were prepared on 200 mesh Formvar carbon-coated grids and stained by 2% ammonium molybdate pH 7.5. Typically, one small drop of the sample solution was allowed to stand on the grid for approximately 1-2 minutes. The grid was carefully dried using Whatman filter paper. Then, the same procedure was repeated with replacement of the sample solution by the stain solution. Finally, the grid was completely dried by air flow. The diameters of the disk-shaped complexes were measured directly on the negatives. The magnifications were 90,000 x.
Chapter 4. C-PEPTIDE

C-PEPTIDE is one of the best amphipathic model peptides according to the criteria of Segrest et al. (1974). It contains nineteen residues (Figure 3.1). About half the residues have either long or aromatic side chains. From the primary sequence, we predicted, using the amphipathic helix model and the hydrophobic moment method, that it would adopt a helical conformation in a complex with lipids (see Chapter 6 for details).

4.1 NMR of C-PEPTIDE in solution

A 1-D spectrum of C-PEPTIDE in solution is shown in Figure 4.1. The spectrum indicates that there is enough spectral resolution and fine structure to yield informative 2-D spectra. C-PEPTIDE in solution was studied using TOCSY and NOESY spectra.

![1-D spectrum](image)

Figure 4.1 1-D spectrum of 5 mM C-PEPTIDE in H2O/D2O (9/1, v/v). The spectrum was recorded at 400 MHz at 37 °C.
4.1.1 Spin Systems

The spin systems of amino acids with long side chains were easily recognized in the TOCSY spectra. Figure 4.2 is a TOCSY spectrum of C-PEPTIDE in solution with the assigned amino acids indicated on the spectrum. One leucine showed a distorted pattern. The methyl groups were correlated to CH and CH$_2$ by one cross peak (Figure 4.2). The connectivity went further down to the $\alpha$H. A valine showed a spin system characterized by two methyls and one methylene (Figure 4.2) in the TOCSY spectrum.

Five amino acids with medium side chains were three glutamate, one glutamine, and one methionine. There were only three such spin patterns in the TOCSY spectrum. One of the three was repeated three times and assigned to glutamates. The second pattern, having chemical shift values close to that of glutamates, was assigned to glutamine. The last one was from the methionine (Figure 4.2). The threonine had a unique pattern with both $\alpha$H and $\beta$H coupled to the CH$_2$ (Figure 4.2). Lysines were all overlapped.

The short chain group had six members, one alanine, one tryptophan, two phenylalanines, and two serines. The alanine was recognized by a long coupling to CH$_3$ (Figure 4.2). Two serines were separated with the rest and had a big downfield shift in the side chain region (Figure 4.3).

The fingerprint region of the TOCSY spectrum is shown in Figure 4.4. The $\delta$CH$_2$ of an arginine was well separated from $\varepsilon$CH$_2$ of lysines. Three lysines degenerated to show only one big relay peak (Figure 4.4). From relay cross peaks from aromatic rings to $\beta$Hs, phenylalanines were identified. Long relay peaks of NHs (Figure 4.5) and couplings between $\alpha$H and aliphatic protons (Figure 4.2) were used to obtain assignments of NHs.
Figure 4.2 An αH relay region of 400 MHz TOCSY spectrum of 5 mM C-PEPTIDE solution in H₂O/D₂O (9/1, v/v) at 37 °C. 450 experiments were recorded with 32 scans and 2K complex data points for each scan. The spectrum width was 3846 Hz. A 1K x 1K real matrix was created by zero-filling once in the t₁ domain. Shifted squared sine bell functions were applied to t₁ and t₂ dimensions before double Fourier transform. Only positive peaks are shown. (V, valine; L, leucine; T, threonine; A, alanine; K, lysine; M, methionine; Q, glutamine; E, glutamate)
Figure 4.3 Side chain region of the TOCSY spectrum of C-PEPTIDE in solution. (F, phenylalanine; W, tryptophan; M, methionine; K, lysine; R, arginine; S, serine)
Figure 4.4 Fingerprint region of the TOCSY spectrum of C-PEPTIDE in solution. (M, methionine; F, phenylalanine; W, tryptophan; K, lysine; R, arginine; S, serine; L, leucine; T, threonine; E, glutamate; Q, glutamine; V, valine)
Figure 4.5 NH relay region of TOCSY spectrum of C-PEPTIDE in solution. (V, valine; L, leucine; T, threonine; A, alanine; K, lysine; R, arginine; E, glutamate; Q, glutamine)
Three residues had aromatic side chains. A tryptophan and only one phenylalanine appeared in the aromatic region (Figure 4.6). The overall assignments for C-PEPTIDE in solution are given in Table 4.1.

4.1.2 Sequential Assignments and Conformation

NOESY experiments were performed with three different mixing times in the range from 150 ms to 500 ms. There were almost no cross peaks in the NOESY spectrum with mixing time of 100 ms. Using longer mixing times, some NOE cross peaks appeared between αHs and NHs of two fragments (Ser9-Glu10-Thr11-Phe12 and Lys14-Val15-Lys16) (Figure 4.7, 4.8). These two segments might be part of a β-sheet structure.

Two unexpected NOE cross peaks arose between leucine 19, lysine 18 and glutamate 17. One of them put a constraint between the αH of leucine 19 and the amide of lysine 18. The other indicated a close distance between the amide of leucine 19 and the aH of glutamate 17. The possible explanation of the constraints was that the leucine bent back and attached to the main chain at the site between the glutamate and the lysine.

The remainder of C-PEPTIDE appeared to have an unstructured conformation in solution.
Figure 4.6 Amide and aromatic region of TOCSY spectrum of C-PEPTIDE in solution. (W, tryptophan; F, phenylalanine)
Table 4.1 Proton assignments of C-PEPTIDE in solution

<table>
<thead>
<tr>
<th>Aromatic 2H</th>
<th>4H</th>
<th>5H</th>
<th>6H</th>
<th>7H</th>
<th>NH</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>7</td>
<td>7.30</td>
<td>7.68</td>
<td>7.28</td>
<td>7.36</td>
</tr>
<tr>
<td>F</td>
<td>1.</td>
<td>7.25</td>
<td>7.40</td>
<td>7.40</td>
<td>7.43</td>
</tr>
<tr>
<td></td>
<td>2.</td>
<td>7.12</td>
<td>7.62</td>
<td>7.62</td>
<td>7.39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>NH</th>
<th>αH</th>
<th>βH</th>
<th>γH</th>
<th>δH</th>
<th>εH</th>
<th>NH</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>5</td>
<td>8.49</td>
<td>4.35</td>
<td>1.80</td>
<td>1.66</td>
<td>3.22</td>
<td>7.14</td>
</tr>
<tr>
<td>K</td>
<td>18</td>
<td>8.37</td>
<td>4.39</td>
<td>1.91</td>
<td>1.55</td>
<td>1.80</td>
<td>3.11</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>19</td>
<td>8.07</td>
<td>4.19</td>
<td>2.18</td>
<td>2.18</td>
<td>1.05</td>
<td>(2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>NH</th>
<th>αH</th>
<th>βH</th>
<th>γH</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>15</td>
<td>8.27</td>
<td>4.46</td>
<td>1.73</td>
</tr>
<tr>
<td>M</td>
<td>4</td>
<td>8.45</td>
<td>4.59</td>
<td>2.25, 2.34</td>
</tr>
<tr>
<td>E</td>
<td>17</td>
<td>8.21</td>
<td>4.37</td>
<td>2.05, 2.14</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8.27</td>
<td>4.47</td>
<td>2.09, 2.20</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>8.33</td>
<td>4.56</td>
<td>2.14</td>
</tr>
<tr>
<td>Q</td>
<td>13</td>
<td>8.33</td>
<td>4.47</td>
<td>2.07</td>
</tr>
<tr>
<td>T</td>
<td>11</td>
<td>8.08</td>
<td>4.33</td>
<td>4.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>NH</th>
<th>αH</th>
<th>βH</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>8.78</td>
<td>4.52</td>
</tr>
<tr>
<td>W</td>
<td>7</td>
<td>7.90</td>
<td>4.64</td>
</tr>
<tr>
<td>F</td>
<td>12</td>
<td>8.15</td>
<td>4.70</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>8.12</td>
<td>4.71</td>
</tr>
<tr>
<td>S</td>
<td>9</td>
<td>8.10</td>
<td>4.43</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>8.27</td>
<td>4.52</td>
</tr>
</tbody>
</table>
Figure 4.7 Fingerprint region of 400 MHz NOESY spectrum of 5 mM C-PEPTIDE in solution in H2O/D2O (9/1, v/v) at 37 °C. The spectrum was recorded, using a 350 ms mixing time, and a digital resolution of 1.9 Hz with 450 t1 increments. The spectrum was stored in an 1K x 1K real matrix. Zero-filling was performed once in the t1 dimension, followed by application of shifted squared sine bell functions to both dimensions before double Fourier transform. Positive peaks were plotted.
Figure 4.8 Fingerprint region of 400 MHz NOESY spectrum of 5 mM C-PEPTIDE solution in H₂O/D₂O (9/1, v/v) at 37 °C. The identical conditions with Figure 4.7 were used except the mixing time of 500 ms and 350 t₁ increments.
4.2 C-PEPTIDE in complexes

C-PEPTIDE and DPC-d38 mixture was subjected to 1-D and 2-D NMR studies. The 1-D spectrum of C-PEPTIDE in complexes is presented in Figure 4.9. Compared to the 1-D spectrum of C-PEPTIDE in solution, some changes were observed. The strongest methyl peaks shifted downfield and a large splitting was visible. This was assigned to DPC-d38 or some impurities. Peaks from aliphatic protons including αH and side chain protons showed intensity changes and shifts due to the binding of the peptide to DPC-d38. The second indication of the binding was the chemical shift and the intensity changes of the peaks from amide and aromatic protons.

Figure 4.9 1-D spectrum of C-PEPTIDE in complexes with DPC-d38 in H₂O/D₂O (9/1, v/v) at 37 °C.
4.2.1 Spin System

From a phase-sensitive DQF-COSY, four lysines were assigned based on the splitting patterns (Figure 4.10). One arginine and one valine were assigned by their chemical shift values (Figure 4.10). Five medium chain residues were recognized by comparing chemical shift values with each other and with those of C-PEPTIDE in the solution. Some unexpected cross peaks appeared in the DQF-COSY spectrum including those from residual protons of DPC-d38 molecules and other impurities. Three empirical assignments for phenylalanines and a tryptophan are shown in Figure 4.11. The fingerprint region of the DQF-COSY was very crowded (Figure 4.12). Almost all cross peaks consisted of double peaks. Thirteen out of nineteen residues were recognized in this region by connectivities to the proper spin systems (Figure 4.12). The details of the cluster area of Figure 4.12 is enlarged in Figure 4.13. Three aromatic residues experienced the most dramatic chemical shift changes when binding of C-PEPTIDE to DPC-d38. Twisted peak patterns for a tryptophan were recognized in Figure 4.14. A similar pattern was observed when glucagon bound to perdeuterated DPC-d38 micelles (Figure 1, in Wider et al., 1982). There might be two conformations for the tryptophan.

TOCSY experiments were performed to provide more information to complete the assignments. The upfield region of the TOCSY spectrum is in Figure 4.15. A valine and a leucine were displayed along with some cross peaks from impurities.
Figure 4.10  Aliphatic region of DQF-COSY of 5 mM C-PEPTIDE in complexes with DPC-d38 at a ratio of 1 to 5 (mol/mol) in H$_2$O/D$_2$O (9/1, v/v) at 37°C. A total of 208 scans was collected for each experiment and 320 experiments were recorded. The digital resolution of FID was 2.0 Hz. The apodization function was shifted sine bell function for both dimensions. (A, alanine; K, lysine; R, arginine; M, methionine; E, glutamate; Q, glutamine; V, valine)
Figure 4.11 Cross peaks between $\alpha$H and $\beta$H of aromatic residues of C-PEPTIDE in complexes in the DQF-COSY spectrum. (F, phenylalanine; W, tryptophan; S, serine)
Figure 4.12 Fingerprint of the DQF-COSY of C-PEPTIDE in complexes.
(K, lysine; R, arginine; V, valine; Q, glutamine; E, glutamate; A, alanine; M, methionine; W, tryptophan; L, leucine)
Figure 4.13 Extended fingerprint region of the DQF-COSY of C-PEPTIDE in complexes. (E, glutamate; W, tryptophan; M, methionine; L, leucine; K, lysine)
Figure 4.14 Aromatic region of the DQF-COSY of C-PEPTIDE in complexes. (W, tryptophan)
Figure 4.15 Side chain region of TOCSY spectrum of 5 mM C-PEPTIDE in DPC-d38 complexes at a ratio of 1 to 5 (mol/mol) at 37 °C. The spin lock time was 70 ms and field strength was 9.9 kHz. 2K complex data points were collected for each FID and 512 experiments were recorded. The spectrum width was 4032 Hz in $t_2$. The spectrum was zero-filled to 1K in $t_1$ dimension and apodized with a phase-shifted squared sine bell function. (V, valine; L, leucine)
Couplings between αH and methyls and methylenes are presented in Figure 4.16. Relay cross peaks for amino acids in the long and the medium chain groups are clearly displayed. In Figure 4.17 four relay cross peaks of lysines were found along with αH-βH cross peaks for phenylalanines and a tryptophan. Serines were located relatively close to the diagonal (Figure 4.18). There should be two more cross peaks for threonine in Figure 4.18. The relay connectivities from amides in a TOCSY spectrum are displayed in Figure 4.19. Some spin systems were visible. From relay connectivities to the upfield peaks, amides of a leucine, a valine, a threonine, and alanines were found. Two lysines were distinguished by the relay cross peaks from both amides and the εNH. The glutamate, the glutamine and the methionine were identified using the NOE information. Relay cross peaks of arginine were missing from this region. The assignment of the cross peaks in the fingerprint region of the TOCSY spectrum was achieved by relay connectivities to side chains from αHs to amides (Figure 4.20). There were three relay cross peaks for NH of the arginine. At least two cross peaks had to be assigned to the amide of alanine. This kind of cross peak pattern for alanine appeared in TOCSY experiments for peptide KIT as well (see Chapter 5). There were eighteen αH-amide cross peaks and five relay cross peaks assigned in the fingerprint region (Figure 4.20).
Figure 4.16 Cross peaks between $\alpha$Hs and side chain protons in the TOCSY spectrum of C-PEPTIDE in complexes. (A, alanine; T, threonine; K, lysine; M, methionine; E, glutamate; Q, glutamine; K, lysine; R, arginine; V, valine)
Figure 4.17 Cross peaks between $\alpha$Hs and $\beta$Hs for amino acids in the medium chain group of C-PEPTIDE in complexes in the TOCSY spectrum. (K, lysine; F, phenylalanine; W, tryptophan)
Figure 4.18 Cross peaks of serines of C-PEPTIDE in complexes in the TOCSY spectrum. (S, serine)
Figure 4.19 Relay cross peaks of C-PEPTIDE in complexes in the TOCSY spectrum. (V, valine; L, leucine; T, threonine; A, alanine; K, lysine; R, arginine; E, glutamate; Q, glutamine; M, threonine)
Figure 4.20 Fingerprint region of TOCSY spectrum of C-PEPTIDE in complexes. (K, lysine; V, valine; Q, glutamine; F, phenylalanine; E, glutamate; A, alanine; R, arginine; S, serine; T, threonine; L, leucine; M, methionine)
A large number of cross peaks were found in the aromatic region of the TOCSY spectrum (Figure 4.21) although there are only three aromatic amino acids. The final assignments of the strong cross peaks to one tryptophan in two different conformations was the only possibility. This confirmed the earlier prediction from the DQF-COSY spectrum (Figure 4.14). The positions of 4H and 7H of the tryptophan shifted so much that the whole spin system was changed. Two phenylalanines showed relatively weak cross peaks in the downfield part (Figure 4.22). One of the two had a spin system with the cross peaks between 4H and 2,6H too weak to be seen or missing.

4.2.2 Sequential assignments

NOESY experiments were performed to study connectivities between residues through NOE cross peaks. Three mixing times (75, 150, and 300 ms) were used. It seemed that the mixing time of 300 ms was long enough to show most NOE cross peaks. The NOE connections are partially displayed in Figure 4.23. The intensities of NOE cross peaks were weaker than those of J-coupling peaks. Sequential assignments procedure was used to connect residues with their neighbors by NOE cross peaks. Lysines, glutamates and phenylalanines were distinguished at this stage. Some medium range NOE cross peaks were used later to determine the secondary structure of the peptide. NOE cross peaks from side chains are shown in Figure 4.24. The side chain connectivities between valine and lysines helped to confirm the above assignments of lysines (Figure 4.24). NOE cross peaks between NHs and side chain protons are presented in Figure 4.25. The complete proton assignments of C-PEPTIDE in complexes are summarized in Table 4.2.
Figure 4.21 Cross peaks from aromatic ring of tryptophan of C-PEPTIDE in complexes in the TOCSY spectrum. (W, tryptophan)
Figure 4.22 Cross peaks from two phenol groups of C-PEPTIDE in complexes in the TOCSY spectrum. (F, phenylalanine)
Figure 4.23 Fingerprint region of NOESY spectrum of 5 mM C-PEPTIDE in complexes in H2O/D2O (9/1, v/v) at 37 °C. The mixing time was 300 ms. 48 FID of 2K complex data points were co-added in each of 512 experiments. The spectrum was stored in a 1K x 1K real matrix after zero-filling once in the $t_1$ domain and apodization with shifted squared sine bell functions in both $t_1$ and $t_2$ dimensions were performed. Positive peaks were plotted.
Figure 4.24 Side-chain NOE cross peaks of C-PEPTIDE in complexes.
Figure 4.25 Relay NOE cross peaks of C-PEPTIDE in complexes.
Table 4.2  Complete assignments of C-PEPTIDE in complexes

<table>
<thead>
<tr>
<th></th>
<th>NH</th>
<th>αH</th>
<th>βH</th>
<th>γH</th>
<th>δH</th>
<th>εH</th>
<th>NH</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>5</td>
<td>8.40</td>
<td>4.37</td>
<td>1.30</td>
<td>1.59</td>
<td>3.25</td>
<td>7.43</td>
</tr>
<tr>
<td>K</td>
<td>18</td>
<td>8.16</td>
<td>3.74</td>
<td>1.74</td>
<td>1.27</td>
<td>1.60</td>
<td>2.86</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>7.72</td>
<td>4.07</td>
<td>1.89</td>
<td>1.51</td>
<td>1.65</td>
<td>2.97</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>7.76</td>
<td>4.20</td>
<td>1.96</td>
<td>1.58</td>
<td>1.73</td>
<td>2.99</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>8.23</td>
<td>4.26</td>
<td>1.82</td>
<td>1.45</td>
<td>1.70</td>
<td>3.02</td>
</tr>
<tr>
<td>L</td>
<td>19</td>
<td>7.82</td>
<td>4.28</td>
<td>1.60,1.75</td>
<td>1.82</td>
<td>0.88 (2)</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>15</td>
<td>8.04</td>
<td>3.55</td>
<td>2.15</td>
<td>0.92,1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>4</td>
<td>8.06</td>
<td>4.44</td>
<td>2.29</td>
<td>2.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>17</td>
<td>7.88</td>
<td>4.07</td>
<td>2.20</td>
<td>2.57,2.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>8.35</td>
<td>4.16</td>
<td>2.19</td>
<td>2.55,2.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8.35</td>
<td>4.25</td>
<td>2.20</td>
<td>2.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>13</td>
<td>7.96</td>
<td>3.87</td>
<td>2.20</td>
<td>2.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>11</td>
<td>7.96</td>
<td>4.18</td>
<td>3.95</td>
<td>1.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>8.69</td>
<td>4.41</td>
<td>1.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>7</td>
<td>8.29</td>
<td>4.31</td>
<td>3.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>12</td>
<td>8.48</td>
<td>4.08</td>
<td>3.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>3.98</td>
<td>3.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>9</td>
<td>8.02</td>
<td>4.23</td>
<td>3.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>4.60</td>
<td>3.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Aromatic

<table>
<thead>
<tr>
<th></th>
<th>2H</th>
<th>4H</th>
<th>5H</th>
<th>6H</th>
<th>7H</th>
<th>NH</th>
</tr>
</thead>
<tbody>
<tr>
<td>W  7</td>
<td>7.28</td>
<td>7.11</td>
<td>6.74</td>
<td>7.00</td>
<td>7.37</td>
<td>9.38</td>
</tr>
<tr>
<td>W' 7</td>
<td>7.36</td>
<td>7.18</td>
<td>6.90</td>
<td>7.07</td>
<td>7.45</td>
<td>9.20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2,6 H</th>
<th>3,5 H</th>
<th>4H</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>12</td>
<td>7.62</td>
<td>8.03</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>7.71</td>
<td>8.15</td>
</tr>
</tbody>
</table>

-88-
4.2.3 Conformation

NH(i)-NH(i+1) NOE cross peaks of C-PEPTIDE in complexes (Figure 4.26) and major NOE cross peaks between αHs and amides (Figure 4.23) were used to provide NOE connectivities. All NOE constraints collected from NOESY spectra are summarized in Figure 4.27. NOE cross peaks between α protons and amide protons of residues at the N-end of C-PEPTIDE were also found, but they are not included in the table because of the questionable assignments due to peak overlappings. NOE cross peaks between amide protons of N-terminal residues of C-PEPTIDE were not found.

A partial helical conformation was supported by NOE peak patterns although the intensities of αH-NH cross peaks were a little stronger than those of NH-NH cross peaks. The helical region included about ten residues at the C-end of the peptide. The N-end of C-PEPTIDE might stay in an unstructured conformation.

The secondary shift plot of C-PEPTIDE αHs in complexes is shown in Figure 4.28. The constant upfield shifts for peptides were observed in other α-helices (Szilágyi and Jardetzky, 1989; Dalgarno et al., 1983). There appears to be two helices. The first helix, in the center of the sequence, did not gain much support from NOE data while the second one, at the C-terminal end, did. Secondary shifts for the residues at the C-end displayed a type of pattern with the large shift appearing every third residue. A twisted helix with three residues in each turn matches such a pattern.

Secondary shift of amide NH of C-PEPTIDE in complexes is shown in Figure 4.29. Upfield shifts were formed for most of the residues in the C-end as seen for αHs. Since the secondary shifts of amides could be affected by temperature and pH etc., this figure was used only to confirm the αH secondary shift results.
Figure 4.26 NOE cross peaks between amides and aromatic ring protons of C-PEPTIDE in complexes.
Figure 4.27 Sequence of C-PEPTIDE and the NOE data used to characterize the secondary structure of C-PEPTIDE in a complex.


\[ \begin{align*}
\alpha H \cdot NH \\
\alpha H \cdot \beta H (i,i+3) \\
NH \cdot NH \\
NH \cdot NH (i,i+2)
\end{align*} \]

Figure 4.28 Secondary shifts of $\alpha H$ of C-PEPTIDE in complexes. The canonical values (Wüthrich, 1986) were used as the standards. (positive values are upfield).
4.3 Proton Exchange

A series of 1-D spectra were recorded in a three-hour time period. For C-PEPTIDE in solution, the amide protons exchanged completely in three hours. This was an indication of a loose structure of the peptide chain. On the other hand, progressive exchanges were observed for amide NH of the C-end of the molecule (Figure 4.30). It seems that an ordered structure was adopted. More work, possibly a 2-D NMR experiment, would be needed to achieve detailed assignments of the exchangeable peaks.

Some slowly exchanging labile protons were observed in the series of 1-D spectra of C-PEPTIDE in complexes (Figure 4.31). The signals from those protons changed very slowly in the three-hour time course studies. Comparison of the 1-D spectra in Figure 4.31 with those in Figure 4.23, yielded preliminary assignments of about ten residues at the C-
end of C-PEPTIDE. These results provided evidence for the conclusion that the lipid associated domain is at the C-end helix of C-PEPTIDE. Further work is needed to investigate the binding mode.

Figure 4.30 1-D spectra of C-PEPTIDE in solution. 16 scans were co-added for each experiment at room temperature. The digital resolution of FID was 0.12 Hz.
Figure 4.31 1-D spectra of C-PEPTIDE in complexes. Experiments were performed at room temperature. 32 scans were recorded in each experiment. Same digital resolution with that for C-PEPTIDE in solution was used.
4.4 Discussion

The micelle system was first introduced as a replacement of natural membranes, based on the findings that membrane-bound enzymes remained active when solubilized from membranes by detergents. The binding of peptides to detergent micelles could be monitored in fluorescence experiments. It was observed, for melittin, that a strong increase of tryptophan fluorescence intensity and an approximately 10 nm blue shift occurred when raising the concentration of detergents above the critical micelle concentration, which were similar to changes observed for the binding of melittin to native and artificial phospholipid membranes (Braun, et al., 1983; Wider, et al., 1983). In NMR studies, melittin bound to micelles gave a spectrum which was different from the spectra of monomeric and tetrameric melittin in aqueous solution. These observations were used to prove the binding of melittin to micelles. The size and the stoichiometry of melittin/DPC micelles were determined by ultracentrifugation and light scattering. A micelle was approximate 50Å in diameter. There were 56±5 DPC molecules per micelle in the absence of melittin, and 1.1±0.2 melittin molecules and 32±10 DPC molecules per micelle in complexes (Lauterwein, et al., 1979).

The structures of two proteins have been determined successfully using distance constraints obtained from NOESY experiments in micelle systems (Braun, et al., 1983; Wider, et al., 1983; Lauterwein, et al., 1979).

The molar ratio of C-PEPTIDE to DPC was 1-2/26 in this study, which was calculated using the weights of C-PEPTIDE and DPC under the assumption that all C-PEPTIDE bound to micelles and all DPC formed micelles. There was no evidence from NMR experiments that unbound C-PEPTIDE was present. Fluorescence studies (Treleaven, unpublished data) also were consistent with C-PEPTIDE binding to DPC. The complexes of C-PEPTIDE/DPC were smaller than other peptide/DPC complexes (Lauterwein, et al., 1979; Wider, et al., 1983) on EM pictures and the narrow linewidths of
2-D NMR spectra reflected a small size of complexes, too. From the calculation and the observation, C-PEPTIDE was assumed to be a monomer in complexes.

The binding of C-PEPTIDE to the detergent micelles was demonstrated by chemical shift changes (Figure 4.33, Figure 4.34). These chemical shift changes were due to two reasons. One of them would be the interactions between detergent molecules and peptide. The other was the formation of secondary structures.

Based on NOE constraints, a secondary structure for C-PEPTIDE in complexes is proposed consisting of a partial helical and partial extended conformation. The helix spanned approximately ten residues at the C-terminal end of the C-PEPTIDE while the N-end of C-PEPTIDE likely remained unstructured. The secondary shifts (Figure 4.33, Figure 4.34) of C-PEPTIDE in complexes suggested that the contact of the peptide to micelles occurred at the C-end of the peptide, which was also seen in the deuterium-proton exchange experiments. Two possibilities exist for the binding mode (Figure 4.32). The helix might be attached to the surface of a micelle. The residues having large secondary shifts for their $\alpha$Hs would be buried in the micelle hydrophobic interior. The other side of the helix would be exposed to aqueous solution. Alternatively, the whole helix might be buried in the micelle with an extended tail (the N-terminal end) staying outside of the micelle. From the secondary shift profile of C-PEPTIDE (Figure 4.33), it is hard to distinguish these two modes. Assignments of exchangeable protons are needed to obtain the exchange rates. If all protons in the helical region have a similar exchange rate, the second mode would be preferred. Otherwise, the first mode might be the case.
Figure 4.32 Two binding modes of C-PEPTIDE to a micelle.

Figure 4.33 Secondary shifts of $\alpha$H of C-PEPTIDE in complexes. The chemical shift values of C-PEPTIDE in solution were taken as the standards.

Figure 4.34 Secondary shifts of amide NH of C-PEPTIDE. The same calculation procedure as in Figure 4.33 were followed.
C-PEPTIDE behaved as an amphipathic peptide generally in this study. It was soluble in H₂O and could form complexes with DPC-d38. In response to the change from a hydrophilic environment to a hydrophobic environment, C-PEPTIDE experienced conformational changes. The study showed that, in solution, C-PEPTIDE stayed in a conformation composed of unstructured conformation and β-sheet, and when C-PEPTIDE was mixed with DPC-d38, the conformation of C-PEPTIDE changed to a helix with an extended chain region.

The conformational changes of C-PEPTIDE essentially followed the amphipathic model of formation of a partial helix upon binding to lipids. On the other hand, the finding of a substantial random chain of C-PEPTIDE in complexes was unexpected and beyond the speculation of the amphipathic theory (see Chapter 6 for details). The amphipathic helix required at least 18 residues to form several turns of a helix according to the theory. Our study found that the helix shorter than 18 residues was stable in complexes with DPC-d38. A question about the hydrophobic face for C-PEPTIDE was posed. Comparing the amphipathic helix (Figure 4.35) and the secondary shift plot (Figure 4.28 and Figure 4.29), residues in the hydrophobic face of the helix, Thr11, Phe12, Val15 and Leu19 (Figure 4.35), were not those residues with big secondary shifts (Figure 4.28 and Figure 4.29). Because the environmental changes partially caused the chemical shift changes, those residues buried in the membrane should show great chemical shift changes. Our observation suggested that Thr11, Phe12, Val15 and Leu19 were not all in a hydrophobic environment.
Figure 4.35 Amphipathic helix of C-PEPTIDE in complexes. The numerals are sequential number of the residue. The bold characters are residues contained in the helix as shown by 2-D NMR studies.

The aggregation problem is a concern in most NMR studies of proteins. For large proteins, there are some methods to verify the existence of monomers. For example, use of chemical cross linking agents would identify oligomeric species. Some methods work less well for small peptides, for example, gel chromatography and sedimentation. Other methods used for small molecules are boiling point elevation but this has obvious drawbacks when dealing with biomolecules. C-PEPTIDE was assumed to be a monomer in this study but this assumption might be inconsistent with the 2-D NMR results for C-PEPTIDE in complexes. In future, this can be verified by other techniques. For example, peptide-lipid complexes can be separated using ultracentrifugation and the peptide/lipid ratio in complexes can be determined by chemical analyses. From the peptide/lipid ratio, we can know the numbers of C-PEPTIDE molecules in a micelle. In the case of one molecule per micelle, C-PEPTIDE must be a monomer. Otherwise, there is a possibility of oligomerization.

Two sets of chemical shifts were found for the aromatic ring protons of the tryptophan and multi-peaks appeared in the TOCSY and the NOESY spectra for the $\epsilon$NH of
the arginine and the methyl group of the alanine. This indicated a complicated secondary structure of C-PEPTIDE in complexes with DPC. One possibility accounting for the extra peaks would be that C-PEPTIDE exists as a dimer in the complex. However, this is not likely. All other peptide protons, including the backbone amide protons and α-protons, had single assignments, hence they experienced a similar magnetic environment. The multiple peaks only occurred in the side-chain region of three residues at the N-terminal end. If dimers were to exist the C-terminal end, which we have shown to be a helix, would have to be identical in both monomers. As we showed above with KIT, when dimers are present usually many more peaks from each monomer are resolved. If dimers were the case, the proton assignments of C-PEPTIDE, the NOE patterns and the structure model in this study discussed above, would apply to the parts of C-PEPTIDE which gave identical NOE cross peaks. In a future study, $^{13}$C spectra might be used to detect the different oligomers.

The current model did not include two positions for the ring of the tryptophan and side chains of alanine and arginine. The ring might exist in two positions relative to the main chain. Since tryptophan is about at the boundary of helical and disordered regions, it is possible that the N terminal end exists in two or more orientations. For this to occur, the part of the helix near the tryptophan is suggested to have a slightly different orientation in the micelles. It is necessary to invoke such a model since there are only two other residues, alanine-2 and arginine-5, which had multiple resonances. More experiments are needed to explain these multi peaks.
Chapter 5. KIT

KIT contains 27 amino acid residues (Figure 3.1). It has one cysteine as the C-terminal residue which allows the peptide molecule the potential to form multimers. A 1-D proton NMR spectrum of KIT is shown in Figure 5.1. The spectrum contained sharp lines for a peptide indicating that the solution structure of KIT was solvable.

![Figure 5.1 A 1-D spectrum of KIT solution in H2O/D2O (9/1, v/v) at 37 °C and pH 2.43. Inset: expansion of region between 7.8 ppm and 9.0 ppm.](image-url)
5.1 Spin System

The fingerprint region of the DQF-COSY spectrum of KIT is given in Figure 5.2. On the upper right corner, two clusters of peaks arose between εNH\textsubscript{3} and εCH\textsubscript{2} of lysines and δCH\textsubscript{2} and NH of arginines. Two arginines and two lysines were identified from those cross peaks. There were fifteen αH-NH cross peaks in the αH-NH region of the DQF-COSY spectrum. Figure 5.3 is the enlarged fingerprint region. Only strong peaks are shown along with assignments. The fact that about half of the αH-NH cross peaks were missing implied that the peptide was not in a regular secondary structure.

In the aromatic region of the DQF-COSY spectrum, a typical tryptophan cross peak pattern was recognized. Cross peaks from the phenylalanine appeared to be so close to the diagonal that only two cross peaks could be seen (Figure 5.4).

Figure 5.5 gives the aliphatic region of the DQF-COSY spectrum. Peaks in the region were from one valine, two isoleucines, and two leucines. The two CH\textsubscript{3} groups of the valine were resolved. The cross peaks arose between CH and CH\textsubscript{3} of the valine in Figure 5.5. Four CH\textsubscript{3} from two isoleucines were difficult to assign. Two broad cross peaks were seen between δCH\textsubscript{3} and γCH\textsubscript{2} from the two isoleucines. The γCH\textsubscript{3} of the isoleucines were related to βH by two individual cross peaks. Leucines showed two different patterns. One of the leucines had separate cross peaks between CH\textsubscript{3} and γCH and between γCH and βCH\textsubscript{2}. The second leucine had only one big cross peak connecting all protons.
Figure 5.2 A plot of a 400 MHz DQF-COSY spectrum of 5 mM KIT in H₂O/D₂O (9/1, v/v) at 37 °C and pH 2.43. The spectrum was recorded in a digital resolution of 2.2 Hz with 512 t₁ increments. A 1K x 1K real matrix was created by zero-filling once in the t₁ dimension. Prior to the double Fourier transform, shifted sine bell window functions were applied to both dimensions. The positive peaks were plotted using 9 contours (dark ones) while the negative peaks using 2 contours. The meanings of symbols are: G, glycine; R, arginine; K, lysine.
Figure 5.3 Fingerprint region of the DQF-COSY spectrum of KIT. Thirteen cross peaks were assigned. Symbols and their meanings are: R, arginine; E, glutamate; Q, glutamine; K, lysine; I, isoleucine; L, leucine; T, threonine.
Figure 5.4 Aromatic-NH region of the DQF-COSY spectrum of KIT. (F, phenylalanines; W, tryptophan)
Figure 5.5 Fingerprint region of long chain amino acids of the DQF-COSY spectrum of KIT. The meanings of symbols are: I, isoleucine; L, leucine; V, valine.
Two lysines and two arginines were identified from the cross peaks between $\varepsilon CH_2$ and $\delta CH_2$ of lysines and between $\delta CH_2$ and $\gamma CH_2$ of arginines (Figure 5.6.).

Cross peaks between $\alpha CH$ and $\beta CH$ of amino acids in the long and medium chain groups are shown in Figure 5.7. The assignments were made by combining the information in Figure 5.7 and in Figure 5.5. Cross peaks between $\alpha CH$ and $\beta CH$ of amino acids in the short chain group are present in Figure 5.8. Five more cross peaks between $\alpha H$ and $\beta H$ of a phenylalanine, a tryptophan, a histidine, a serine and an aspartate were expected. Glycine was missing in the side chain region, which implied a degeneracy between the two $\alpha H$ of the glycine. There were not enough clues to make assignments for short chain group in the DQF-COSY spectrum.

TOCSY spectra were recorded for KIT in solution as a supplement to DQF-COSY experiments. Figure 5.9 was the TOCSY spectrum counterpart of the region of Figure 5.5. Some new cross peaks appeared due to relay couplings, which were used to confirm the assignments for the above mentioned five amino acids. The aromatic protons of the histidine were assigned in the aromatic region of the TOCSY spectrum in Figure 5.10.

One proline was recognized by an unique cross peak pattern (Figure 5.11). The $\beta CH_2$ of the proline were coupled to the $\delta CH_2$ through the $\gamma CH_2$. A cis conformation was assumed for the proline. Two other amino acids, glutamate and glutamine were assigned in Figure 5.11 based on the assumption that an amide group caused a further downfield shift than a carbonyl group.
Figure 5.6 Cross peaks between $\gamma$CH$_2$ and $\delta$CH$_2$ of arginine and between $\delta$CH$_2$ and $\epsilon$CH$_2$ of lysine in the DQF-COSY spectrum of KIT. (R=arginine, K=lysine)
Figure 5.7 Fingerprint region for long and medium chain amino acids in the DQF-COSY spectrum of KIT. (A, alanine; I, isoleucine; V, valine; P, proline; E, glutamate; Q, glutamine; K, lysine; T, threonine; L, leucine)
Figure 5.8 Fingerprint region for short chain amino acids in the DQF-COSY spectrum of KIT. Arrows indicated the cross peaks from amino acids in the short chain group.
Figure 5.9 Long chain amino acid fingerprint region of a 400 MHz TOCSY spectrum of 5 mM KIT solution in H$_2$O/D$_2$O (9/1, v/v) at 37 °C and pH 2.43. The spectrum was recorded using a 70 ms MLEV-17 mixing time and a field strength of 8.8 kHz. The digital resolution of FIDs was 2.2 Hz and 512 $t_1$ increments were collected. Prior to double Fourier transform, zero-filling to 1K in the $t_1$ dimension was performed and a shifted sine bell window function was applied to both dimensions. Only the positive peaks were plotted. Arrows indicated the multi-relay cross peaks.
Figure 5.10 Aromatic region of the TOCSY spectrum of KIT. (H, histidine)
Figure 5.11 Aliphatic region of the TOCSY spectrum of KIT. (P, proline; E, glutamate; Q, glutamine)
The assignments for amino acids in the long and the medium chain groups were confirmed in the TOCSY spectrum. The long relay cross peaks in the TOCSY spectrum are given in Figure 5.1. The βCH₂ of phenylalanine and tryptophan were assigned using cross peaks to aromatic rings. The fingerprint region of the TOCSY spectrum (Figure 5.12) contained too many broad cross peaks to allow any assignments.

Overall assignments for 23 residues are listed in Table 4.7. Lysines, alanines, the histidine (part) and an aspartate were not assigned definitively for various reasons. There exists the possibility of degeneracy of lysines. In the DQF-COSY and the TOCSY spectra, seven to eight relay connections appeared to be from alanines (Figure 5.12) but there are only four alanines. Other amino acids, glycine, serine, tryptophan, phenylalanine and cystine were assigned empirically. Additional information from NOESY spectra were needed to finish the complete proton assignments for KIT.

### 5.2 Sequence-specific assignment

A series of NOESY spectra were recorded with different mixing times. But the overlapping of the peaks were so severe that assignments of the NOE cross peaks were impossible (Figure 5.14).
Figure 5.12 Relay cross peak region of the TOCSY spectrum of KIT. (A, alanine; R, arginine; E, glutamate; Q, glutamine; K, lysine; I, isoleucine; L, leucine; T, threonine)
Figure 5.13 Fingerprint region of the TOCSY spectrum of KIT.
Figure 5.14  Fingerprint regions of 400 MHz NOESY spectra of 5 mM KIT solution in H₂O/D₂O (9/1, v/v) at 37 °C and pH 2.43. Three different mixing times were used. (a). 300 ms; (b). 450 ms; and (c). 550 ms.
Table 5.1 Proton assignments of KIT in solution

<table>
<thead>
<tr>
<th></th>
<th>NH</th>
<th>αH</th>
<th>βH</th>
<th>γH</th>
<th>δH</th>
<th>εH</th>
<th>εNH</th>
<th>NH</th>
</tr>
</thead>
<tbody>
<tr>
<td>R 1.</td>
<td>8.42</td>
<td>4.44</td>
<td>1.93,1.88</td>
<td>1.74</td>
<td>3.31</td>
<td>-</td>
<td>7.27</td>
<td>6.77</td>
</tr>
<tr>
<td>R 2.</td>
<td>8.49</td>
<td>4.50</td>
<td>1.93,1.88</td>
<td>1.70</td>
<td>3.30</td>
<td>-</td>
<td>7.24</td>
<td>6.71</td>
</tr>
<tr>
<td>K 1.</td>
<td>8.33</td>
<td>4.46</td>
<td>1.96</td>
<td>1.54</td>
<td>1.80</td>
<td>3.11</td>
<td>7.62</td>
<td>-</td>
</tr>
<tr>
<td>K 2.</td>
<td>8.23</td>
<td>4.34</td>
<td>1.77</td>
<td>1.38</td>
<td>1.70</td>
<td>3.04</td>
<td>7.58</td>
<td>-</td>
</tr>
<tr>
<td>L 1.</td>
<td>8.04</td>
<td>4.30</td>
<td>1.50</td>
<td>1.40</td>
<td>0.88,0.93</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L 2.</td>
<td>7.91</td>
<td>4.20</td>
<td>2.18</td>
<td>2.18</td>
<td>1.04</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I 1.</td>
<td>8.11</td>
<td>4.26</td>
<td>1.98</td>
<td>1.60,1.30</td>
<td>0.98</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I 2.</td>
<td>8.16</td>
<td>4.60</td>
<td>2.00</td>
<td>1.64,1.30</td>
<td>1.08</td>
<td>1.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>-</td>
<td>4.49</td>
<td>2.44</td>
<td>2.18,2.07</td>
<td>3.82,4.00</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: "?" indicates peak not assigned.
5.3 Discussion

Since distance geometry information was not available for KIT, the precise conformation could not be solved. Some comments could be made based on the results obtained these far. The absence of cross peaks between NH and NH suggested a non-helix conformation. However, secondary chemical shift of KIT displayed a upfield shift trend which was correlated to helical structures (Dalgarno et al., 1983; Szilagyi and Jardetzky, 1989, Figure 5.15).

![Figure 5.15](image_url)

Figure 5.15 Secondary shift of alpha protons of KIT as a function of the primary sequence. Blank spaces are residues without assignments.
Chapter 6. Amphipathic Peptides

6.1 Structure prediction

6.1.1 Amphipathic helix

C-PEPTIDE and LAP-20 were amphipathic peptides as both of them clearly showed two "faces" on the presentations of the "Schiffer-Edmundson" helical wheel (Figure 6.1). LAP-20 had a slightly larger nonpolar "face" than C-PEPTIDE had. But C-PEPTIDE had three positive charges to bind to negative charges of phosphotidylcholine while LAP-20 only had two. Both C-PEPTIDE and LAP-20 had high hydrophobicity indexes (Table 6.1) for the hydrophobic faces indicating that they were good amphipathic peptides.

The three small peptides, KIT, B19, and SH3, were selected from three different proteins. Those peptides had two "faces" as well (Figure 6.1). B19 had the largest polar "face" with four ion pairs. KIT showed a typical amphipathic helix with two ion pairs. SH3 was a reversed amphipathic helix with positive charges in the center of the polar face and negative charges between the polar and the non-polar faces.

Table 6.1 Hydrophobic Indexes of the peptides

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Hydrophobicity Index of nonpolar &quot;face&quot;</th>
<th>Ion pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-PEPTIDE</td>
<td>4.1</td>
<td>3</td>
</tr>
<tr>
<td>LAP-20</td>
<td>4.7</td>
<td>2</td>
</tr>
<tr>
<td>KIT</td>
<td>3.3</td>
<td>2</td>
</tr>
<tr>
<td>B19</td>
<td>3.0</td>
<td>4</td>
</tr>
<tr>
<td>SH-3</td>
<td>1.8</td>
<td>2</td>
</tr>
</tbody>
</table>
6.1.2 Hydrophobic Moment

Figure 6.2 shows the hydrophobic moment profiles of the five peptides we studied. C-PEPTIDE and B19 had strong peaks at 100° indicating that α-helix structures are predicted. The others had maxima either 10° below or above 100° while KIT showed a double maxima at 90° and 110°. The peak at 160°-180°, which corresponded to β-sheets (Eisenberg et al., 1984), existed for all five peptides. In some cases, LAP-20, SH3 and KIT, the 160°-180° peak was as or more intense than the 100° band. From the hydrophobic moment profiles, α-helices could be predicted as the most likely conformations for C-PEPTIDE and B19 in complexes with lipid membranes while both α-helices and β-sheets might exist for LAP-20, KIT, and SH3 to be amphipathic peptides. The third feature of the hydrophobic moment profiles was the peak at 0° which has not been assigned to any type of structure to date.
Figure 6.1. "Schiffer-Edmundson" helical wheel expressions of helices for KIT, B19, and SH3 and the purported amphipathic helices for C-PEPTIDE and LAP-20.
Figure 6.2. The hydrophobic moment profiles for peptides.
6.2 Formation of Complexes

The formation of complexes of peptide and DMPC was demonstrated using electron microscopy (EM) and in turbidity experiments. Discoidal particles of complexes were seen on negative stain electron micrographs, which were different from vesicles of DMPC (Figure 6.3; Figure 6.4; Figure 6.5). EM pictures of discs are characterized by long stacks or Rouleaux. The Rouleaux of KIT/DMPC were short compared to B19/DMPC and LAP-20/DMPC. Three complexes had a similar diameter (Table 6.2). Turbidity changes were observed after addition of peptides to DMPC aqueous solution. In the best case, LAP-20 turbidity cleared in 5 minutes at room temperature. B19/DMPC took about one hour before turning to a clear solution. KIT was incubated with DMPC at 50° for 3 days to form a clear solution. SH3/DMPC mixture turned to a clear solution after incubation at 50° for 3 hours. Decreased turbidity is due to complex formation of amphipathic peptide (J.P. Segrest, private communication). At room temperature, the complexes of SH3/DMPC were in a gel state. Shaking could break the gel into a type of solution, but once the shaking stopped, it went back to the gel state again in a few minutes.

Table 6.2 The diameters of the different complexes of peptide and DMPC

<table>
<thead>
<tr>
<th></th>
<th>LAP-20/DMPC</th>
<th>KIT/DMPC</th>
<th>B19/DMPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>diameter (nm)</td>
<td>21</td>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>deviation (nm)</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 6.3. The negative stain electron micrograph of complexes of LAP-20 and DMPC (magnification 90,000x).

Figure 6.4. The negative stain electron micrograph of complexes of KIT and DMPC (magnification 90,000x).
6.3 FTIR studies

6.3.1 LAP-20

6.3.1.1 LAP-20 in solution

Figure 6.6.a is the infrared spectrum, between 1570 and 1720 cm\(^{-1}\), of LAP-20 in D\(_2\)O. The conformation-sensitive amide I region consisted mainly of two broad contours with maxima at 1612 cm\(^{-1}\) and 1673 cm\(^{-1}\). As with many other proteins, the amide I contours were composites of overlapping component bands which represented different elements of protein secondary structure. Most of these components could be resolved by the computational procedure of band narrowing by Fourier self-deconvolution. A width
of 15 cm\(^{-1}\) and an enhancement factor of 1.7 were used to narrow the bands in the self-deconvolution process. The deconvoluted spectrum of LAP-20, bottom trace in Figure 6.6.a, showed six component bands in the region between 1600 cm\(^{-1}\) and 1700 cm\(^{-1}\) (Table 6.5). Four strong peaks were at 1612, 1640, 1655, and 1673 cm\(^{-1}\). Two additional peaks were poorly resolved under the biggest and the second biggest peaks at 1612 and 1673 cm\(^{-1}\), respectively. The component at 1655 cm\(^{-1}\) was in the region characteristic of \(\alpha\)-helices (Byler and Susi, 1986; Surewicz and Mantsch, 1988). The band at 1627 cm\(^{-1}\) and the component at 1673 cm\(^{-1}\) were highly characteristic of the peptide backbone in a \(\beta\)-sheet conformation (Byler and Susi, 1986; Surewicz and Mantsch, 1988). The band at 1612 cm\(^{-1}\) was assigned to an intermolecular \(\beta\)-sheet which may include some contribution from side chain absorption. A similar band was observed by other groups for peptide M (Muga, \textit{et al.}, 1990), apocytochrome c (Muga, \textit{et al.}, 1991), poly(L-lysine) (Carrier, \textit{et al.}, 1990) and atriopeptin (Surewicz, \textit{et al.}, 1987). The band at 1640 cm\(^{-1}\) might be due to random structures or \(\beta\)-sheets. We have assigned it to a random structure (Table 6.3). The band at 1686 cm\(^{-1}\) was indicative of the peptide chain in turn structures. Two bands, out of the region between 1600 and 1700 cm\(^{-1}\), were not assigned to any structures but were included in the curve fitting. Figure 6.6.b shows the results of curve fitting along with the eight component bands.

The infrared spectra of LAP-20 in solution clearly indicated an ordered structure and that the predominant conformation was intra- and intermolecular \(\beta\)-sheets (Table 6.7) which were characterized by the component bands at 1673 cm\(^{-1}\) and 1612 cm\(^{-1}\), respectively. The \(\alpha\)-helix content of the LAP-20 from infrared studies agreed with CD results (Pownall, \textit{et al.}, 1980).
Figure 6.6. (a) Amide I range of the IR spectrum of LAP-20 in free form in D$_2$O (thin line) plus simulated spectrum (thick line; see (c)); (b) the deconvoluted spectrum using a line width of 15 cm$^{-1}$ and an enhancement factor of 1.7; (c) Curve fitting of LAP-20 spectrum. The simulated spectrum (summation of component bands) is the top trace. Fitting components are shown as bands at the bottom.
6.3.1.2 LAP-20 in complexes with DMPC

The infrared spectrum, between 1580 and 1710 cm\(^{-1}\), of LAP-20 in complexes with DMPC is shown in Figure 6.7.a (top trace). The overall features of the amide I region of LAP-20 in complexes with DMPC was dramatically different from LAP-20 in solution. The broad amide I band contour exhibited a maximum at 1648 cm\(^{-1}\) and two well defined peaks around 1612 and 1673 cm\(^{-1}\), and the 1612 cm\(^{-1}\) band diminished significantly. Fourier self-deconvolution, using a line width of 15 cm\(^{-1}\) and an enhancement factor of 2.1, revealed nine component bands in the region between 1600-1700 cm\(^{-1}\) (Figure 6.7.a, bottom trace). Eight of the nine bands were well resolved. The band at 1683 cm\(^{-1}\) appeared as a small shoulder of the strong peak at 1673 cm\(^{-1}\). The band at 1651 cm\(^{-1}\) was associated with an α-helix structure. The residual β-sheets were represented by two components at 1625 cm\(^{-1}\) and 1635 cm\(^{-1}\) and one component at 1673 cm\(^{-1}\). The band at 1612 cm\(^{-1}\) was assigned to β-sheet structures. The band at 1644 cm\(^{-1}\) indicated random coils. Side chains of amino acids were reported to contribute to peaks below 1620 cm\(^{-1}\) in some cases (Rial et al., 1990, Yang et al., 1991). Thus the assignment did not exclude the absorption from side chains. The bands at 1663, 1683 and 1695 cm\(^{-1}\) represented turns. Table 6.3 gives the assignments and relative areas of the various bands. The infrared spectrum of LAP-20 in complexes of LAP-20/DMPC revealed a conformation featured by a large portion of normal and intermolecular β-sheets and a relatively high turn content.
Figure 6.7. Amide I range of the IR spectra of LAP-20 in complexes with DMPC. The same notations as in Figure 6.6 were used. The deconvolution line width was 15 cm$^{-1}$ and the enhancement factor was 2.1.
Table 6.3  Position and fractional areas of the amide I bands of LAP-20 in solution and in complexes with DMPC

<table>
<thead>
<tr>
<th>Band assignment</th>
<th>LAP-20 in solution</th>
<th>LAP-20 in complexes with DMPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-helix</td>
<td>1655 (10%)</td>
<td>1651 (15%)</td>
</tr>
<tr>
<td>β-pleat</td>
<td>1673 (26%)</td>
<td>1673 (15%)</td>
</tr>
<tr>
<td></td>
<td>1627 (8%)</td>
<td>1635 (14%)</td>
</tr>
<tr>
<td></td>
<td>1625 (6%)</td>
<td>1625 (6%)</td>
</tr>
<tr>
<td></td>
<td>1612 (39%)</td>
<td>1612 (14%)</td>
</tr>
<tr>
<td>random</td>
<td>1640 (10%)</td>
<td>1644 (13%)</td>
</tr>
<tr>
<td>turns</td>
<td>1686 (7%)</td>
<td>1695 (1%)</td>
</tr>
<tr>
<td></td>
<td>1683 (7%)</td>
<td>1663 (14%)</td>
</tr>
</tbody>
</table>

6.3.1.3 Discussion

LAP-20 in solution adapted mainly intermolecular and intramolecular β-sheets which were indicated by the bands at 1627 cm⁻¹, 1673 cm⁻¹ and an intense band at 1612 cm⁻¹ (Table 6.3), respectively, and small amounts of helices, which was represented by a small band at 1655 cm⁻¹. Band shifts and new bands appeared for LAP-20 in complexes. Beta structure, which included about 10 a.a. residues, was indicated by four bands (Table 6.3) and the helical structure calculated from the band at 1651 cm⁻¹. Our IR results showed that the conformational changes induced by formation of complexes of LAP-20/DMPC mainly included an increase in turn and helical contents and a decrease in beta content. It was concluded that not only α-helices, but also β-sheets were important for the binding of LAP-20 to DMPC membranes under our experimental conditions.
The band at 1612 cm\(^{-1}\) was assigned to intermolecular \(\beta\)-sheets in this study and in other studies (Surewicz and Mantsch, 1988). Intermolecular \(\beta\)-sheets might be a consequence of a self-association which would explain the high amount of \(\beta\)-structure found for LAP-20 in solution. A similar result was reported for atriopeptin III (Surewicz, \textit{et al.}, 1987). In the presence of lipids, LAP-20 self-association was replaced by peptide-lipid interactions. In response of the new interaction, the conformation of peptide changed from the intermolecular \(\beta\)-sheet to a mixture of \(\beta\)-pleat, helix and turn structures. However, \(\beta\)-sheets remained the main conformation.

The IR results suggested a conformation which was different from the amphipathic helix. Changes from unstructured conformation to a significant amount of \(\alpha\)-helices, reported for many amphipathic model peptides based on CD studies, was not found in FTIR studies (Yang \textit{et al.}, 1991). FTIR spectra suggested an ordered structure for LAP-20 in solution with a high \(\beta\)-sheet content, some turns and some \(\alpha\)-helices under our experimental conditions. A significant decrease in \(\beta\) content and an increase in turns and helix, upon the binding of lipids, were demonstrated by FTIR spectra. However, \(\beta\)-sheets remained the major conformation for LAP-20 in both solution and discoidal complexes with phospholipids. The concentrations used for IR studies are about 50 fold higher than that used for CD, hence we are probably seeing oligomers forming at the higher concentrations. With our present IR equipment we were unable to obtain spectra at the low concentrations necessary for CD, in spite of purchasing a more sensitive detector for the spectrometer.

A model could be proposed that the peptide consisted of frozen domains and flexible domains. The frozen domains included \(\beta\)-sheets which were found in both self-association complexes and LAP-20/DMPC complexes. Bands at 1673, 1635, 1625, and 1612 cm\(^{-1}\) of complexes were associated with \(\beta\)-structures in the frozen domain. The
band at 1612 cm\(^{-1}\) suggested that LAP-20 might exist in multimers. In the complex of LAP-20/DMPC, a flexible domain was found consisting of random structures and turns. The slight increase of random structure in the complexes might result in a more flexible conformation for LAP-20, thus allow it to carry out such biological functions as activating LCAT.

When our IR studies were close to completion, it became apparent that there were several problems associated with using IR to determine protein secondary structure. We therefore used this opportunity to compare IR with other techniques for structure determination.

The errors in a typical FTIR study come from (1) the assumption of a constant extinction coefficient for all secondary structures; (2) the empirical assignments of component bands to secondary structures; (3) the fitting error between original spectrum and the fitting spectrum; (4) error in obtaining deconvoluted spectra and (5) the use of RMS in FTIR to evaluate closeness between original spectrum and fitting curve. Unfortunately, RMS is proportional to the intensity of the spectrum. Thus, we have used IR only in a semiquantitative manner to estimate the secondary structure of our peptides. From above error analyses, it is concluded that it is impossible to even approximate the total errors in IR measurements theoretically.

Another problem in the error estimation is, in most cases, lacking of the real protein structure to compare with. From very beginning, FTIR, as a method to determine protein structures, was based on the comparison of IR results for proteins to results from x-ray studies of the same proteins, but not on theoretical evaluation, such as using of the extinction coefficients. In the earlier FTIR study, attention was focussed on deconvolution and band assignments. The percentages of structures were calculated by adding the areas of all bands assigned to each structure and expressing the sum as a fraction of the total amide I
band area. Such an approach obviously neglected the differences among the integrated intensities of amide I bands arising from segments with different secondary structures. But it did lead to acceptable empirical results for the proteins studied (Byler and Susi, 1986). FTIR and CD results were usually expressed as percentages and x-ray results could be converted to percentage (Levitt & Greer, 1977). A direct comparison of the percentages, an important part to establish FTIR methods, was carried out. It was found that the discrepancy among FTIR, CD and x-ray results was varied upon proteins and secondary structures (Byler and Susi, 1986). However, FTIR results and x-ray results for 11 proteins were agreed within 4%, 2.3% in average (Byler and Susi, 1986). A matrix method was applied to recent FTIR study. Because the matrix method is out of the scope of my thesis, I will not introduce the procedure here, but give the references (Sarver and Krueger, 1991a,b). This approach avoided deconvolution and prior assignments of component bands, which were the most controversial parts of a traditional FTIR method. The average percent difference between the IR (10 proteins base) and x-ray calculated secondary structures was 10.3% for helix, 6.5% for β-sheet, 6.7% for β-turn and 6.3% for other. It was found that CD gave a better helical estimation while FTIR worked better on β-sheet.

Recent FTIR study on Streptokinase brought a new question to FTIR results (Fabian, et al., 1992). It was found that the side chain of amino acids can contribute to absorption in the amide I band. These absorptions are usually confused with those bands associated with β-structures. Since the absorption of side chain groups is varied upon environment, it will not be right to use the absorptions of single amino acids to approximate the absorptions of side chains in peptides. The solution to the problem has not been found yet. However, it was suggested from the same study that the side chain contributed a maximum of less than 10% to the entire β and random range (Fabian, et al., 1992). Considering the side chain effect, FTIR may overestimate β-structures although the work of Saver & Krueger
(1991a,b) indicated secondary structures analyze to within 15% of the x-ray structure and analyze to within 8% of the x-ray 55% of the time.

6.3.2 Other peptides

KIT, B19 and SH3 were peptides which were selected from membrane-binding domains of different proteins. All three peptides had some biological functions.

6.3.2.1 KIT

KIT is the fragment from residues 187 to 214 of the transmembrane receptor kinase encoded by the mouse c-kit gene. The top trace in Figure 6.8.a shows the IR spectrum, between 1595 and 1710 cm\(^{-1}\), of KIT in D\(_2\)O solution. The amide I mode between 1600 and 1700 cm\(^{-1}\) consisted of two broad peaks with maxima at 1645 and 1673 cm\(^{-1}\), respectively. Overall features of the amide I contour were characteristic of peptide chain in random coils and \(\beta\)-sheets. Fourier self-deconvolution using a band width of 16 cm\(^{-1}\) and an enhancement factor of 2.4 gave eight component bands (Figure 6.8, Table 6.4). A band at 1659 cm\(^{-1}\) represented \(\alpha\)-helix. The biggest band at 1673 cm\(^{-1}\) along with two small bands at 1632 cm\(^{-1}\) and 1622 cm\(^{-1}\) indicated the peptide chain in \(\beta\)-sheet conformations. The broad band at 1648 cm\(^{-1}\), accounting for 21% of the total area, and a band at 1640 cm\(^{-1}\) were characteristic of the peptide chain in random coils. A small band, at 1683 cm\(^{-1}\), was associated with turns. A band at 1610 cm\(^{-1}\) was assigned to side-chains. Thus it was not counted in area calculation but it was used in curve-fitting process. Figure 6.8.b demonstrates the curve fitting results (top trace) along with component bands. Figure 6.9 depicts the IR spectrum, from 1595 cm\(^{-1}\) to 1710 cm\(^{-1}\), of KIT/DMPC complexes in D\(_2\)O and the deconvoluted spectrum using a line width of 19 cm\(^{-1}\) and an enhancement factor of 2.0. The amide I band contour of KIT in complexes had the same overall shape, of two broad peaks, as KIT in solution. Seven peaks were visualized by Fourier self-
deconvolution. The band at 1659 cm\(^{-1}\) was associated with \(\alpha\)-helix. The predominant structure, \(\beta\)-sheet, was represented by three bands at 1673, 1638, and 1628. The band at 1612 cm\(^{-1}\) was from contributions of side chain groups. The band shifts from 1622 and 1632 cm\(^{-1}\) to 1628 and 1638 cm\(^{-1}\), respectively, indicated a new type of \(\beta\)-sheet structures existing for KIT in complexes. The band at 1648 cm\(^{-1}\) was assigned to random coils. The band at 1686 cm\(^{-1}\) indicated turn structures. The assignments are summarized in Table 6.4.

Table 6.4 Position and fractional areas of the amide I bands of KIT in solution and in complexes with DMPC

<table>
<thead>
<tr>
<th>Band assignment</th>
<th>KIT in solution</th>
<th>KIT in complexes with DMPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)-helix</td>
<td>1659 (19%)</td>
<td>1659 (12%)</td>
</tr>
<tr>
<td>(\beta)-pleat</td>
<td>1673 (32%)</td>
<td>1673 (45%)</td>
</tr>
<tr>
<td></td>
<td>1632 (6%)</td>
<td>1638 (6%)</td>
</tr>
<tr>
<td></td>
<td>1622 (2%)</td>
<td>1628 (8%)</td>
</tr>
<tr>
<td>random</td>
<td>1648 (21%)</td>
<td>1648 (28%)</td>
</tr>
<tr>
<td></td>
<td>1640 (14%)</td>
<td></td>
</tr>
<tr>
<td>turns</td>
<td>1683 (6%)</td>
<td>1686 (&lt; 1%)</td>
</tr>
</tbody>
</table>
Figure 6.8. Amide I range of the IR spectra of KIT in free form in D$_2$O. The notations in Figure 6.6 were used except that the deconvolution line width was 16 cm$^{-1}$ and the enhancement factor was 2.4.
Figure 6.9. Amide I range of the IR spectra of KIT in complexes with DMPC. The notations in Figure 6.6 were used except that the deconvolution line width was 19 cm\(^{-1}\) and the enhancement factor was 2.0.
The IR results demonstrated a decrease of random coil, α-helix and turns, and an increase of β-sheet when KIT binds to DMPC membranes. β-sheets remained the predominant conformation for KIT both in solution and in discoidal complexes with DMPC.

6.3.2.2 B19

The IR spectrum, from 1595 cm\(^{-1}\) to 1705 cm\(^{-1}\), of B19 in D\(_2\)O is presented in Figure 6.10.a. The amide I mode exhibited a maximum at 1648 cm\(^{-1}\) and two pronounced shoulders at 1613 cm\(^{-1}\) and 1672 cm\(^{-1}\), respectively. The spectrum after Fourier self-deconvolution, using a line width of 20 cm\(^{-1}\) and an enhancement factor of 2.3, revealed, in addition to the major band at 1648 cm\(^{-1}\), three other relatively strong bands at 1671, 1737, and 1613 cm\(^{-1}\) and weak bands at 1639, 1693, 1681, and 1661 cm\(^{-1}\). Two bands at 1630 cm\(^{-1}\) and 1681 cm\(^{-1}\) were poorly resolved. The broad band contour consisted of a major band at 1648 cm\(^{-1}\) along with bands around 1637 and 1661 cm\(^{-1}\).

Among the four basic structures, traces of α-helices were missing from the amide I region of B19 (Table 6.5). The β-structure of B19 was represented by the bands at 1613 (partial), 1630, 1637, and 1671 cm\(^{-1}\). The band at 1613 cm\(^{-1}\) we have assigned to intermolecular β-sheets and side chains (see above). The band at 1648 cm\(^{-1}\) was assigned to random coil structures. Three other bands at 1695, 1681, and 1661 cm\(^{-1}\) were associated with turns. The band fitting curve is shown in Figure 6.10.b along with the component bands.

Binding of B19 to DMPC membranes changed the relative intensities of the shoulder at 1613 cm\(^{-1}\), and the broad band around 1649 cm\(^{-1}\) (Figure 6.11.a, top trace). Five band components were under the broad band around 1649 cm\(^{-1}\) as revealed by Fourier self-deconvolution using a line width of 18 cm\(^{-1}\) and enhancement factor of 2.5. The major
changes of the spectrum from the one of the B19 solution were the appearance of bands at 1649 cm\(^{-1}\) and 1658 cm\(^{-1}\). These two bands were assigned to \(\alpha\)-helices. An \(\alpha\)-helical band at 1649 cm\(^{-1}\) observed for apocytochrome c in complexes with DMPC was believed to be due to the hydrogen-deuterium exchange of amide groups in a helix (Muga, et al., 1991).

The same assignment strategy as in B19 solution was applied to B19 in complexes. The detailed information about the bands and the assignments are listed in Table 6.5.

The IR results of B19 in solution indicated that the dominant structure was \(\beta\)-sheet (Table 6.5). An equal amount of \(\alpha\)-helix, \(\beta\)-sheet and random coils were observed for B19 in complexes (Table 6.5). It was concluded that, upon the binding to DMPC vesicles, B19 changed conformation from \(\beta\)-sheets to \(\alpha\)-helices and random coils.

Table 6.5 Position and fractional areas of the amide I bands of B19 in solution and in complexes with DMPC

<table>
<thead>
<tr>
<th>Band assignment</th>
<th>B19 in solution</th>
<th>B19 in complexes with DMPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)-helix</td>
<td>1658 (15%)</td>
<td>1649 (19%)</td>
</tr>
<tr>
<td>(\beta)-pleat</td>
<td>1671 (28%)</td>
<td>1673 (19%)</td>
</tr>
<tr>
<td></td>
<td>1637 (32%)</td>
<td>1632 (1%)</td>
</tr>
<tr>
<td></td>
<td>1630 (1%)</td>
<td>1624 (2%)</td>
</tr>
<tr>
<td></td>
<td>1613 (11%)</td>
<td>1613 (8%)</td>
</tr>
<tr>
<td>random</td>
<td>1648 (17%)</td>
<td>1641 (33%)</td>
</tr>
<tr>
<td>turns</td>
<td>1695 (3%)</td>
<td>1684 (3%)</td>
</tr>
<tr>
<td></td>
<td>1681 (1%)</td>
<td>1661 (7%)</td>
</tr>
</tbody>
</table>
Figure 6.10. Amide I range of the IR spectra of B19 in free form in D2O. The notations in Figure 6.6 were used except that the deconvolution line width was 20 cm\(^{-1}\) and the enhancement factor was 2.3. See Table 6.5 for the band frequencies and areas.
Figure 6.1. Amide I range of the IR spectra of B19 in complexes with DMPC. The notations in Figure 6.6 were used except that the deconvolution line width was 18 cm\(^{-1}\) and the enhancement factor was 2.5.
6.3.2.3 SH3

The top trace in Figure 6.12.a shows the amide I region of the IR spectrum, between 1575 cm\(^{-1}\) and 1705 cm\(^{-1}\), of SH3 solution in D\(_2\)O. The amide I mode was dominated by a strong peak at 1615 cm\(^{-1}\). The overall feature was intermolecular \(\beta\)-sheet structures. Six component bands were found in the deconvoluted spectrum in the region between 1600 to 1700 cm\(^{-1}\). A line width of 20 cm\(^{-1}\) and enhancement factor of 2.1 were used to narrow the component bands. One band appeared as a shoulder at 1629 cm\(^{-1}\). The same assignment strategy as above was used. The curve fitting and the component bands are given in Figure 6.12.b. The results are listed in Table 6.6.

The amide I mode, between 1585 and 1705 cm\(^{-1}\), of SH3 in complexes with DMPC is the top trace in Figure 6.13.a. It had the same features as the one for SH3 in solution. Fourier self-deconvolution using a line width of 18 cm\(^{-1}\) and an enhancement factor of 2.3 revealed eight component bands. The component bands and the fitting curve are in Figure 6.13.b. The band assignments are given in Table 6.6.

SH3 showed little conformational change in the mixture of SH3/DMPC except that more intermolecular \(\beta\)-sheets were observed. From the IR spectra of SH3 and SH3/DMPC, it appears that SH3 was in self-associated complexes even in the presence of DMPC. This might be explained by the fact that SH3 needed to form a multimer to shield its charged residues. The binding of the SH3 to DMPC membranes was not verified by EM due to the gel state structure of SH3/DMPC mixture. Other peptides have been shown to form gels due to multimer formation (J.C. Vederas, private communication).
Figure 6.12. Amide I range of the IR spectra of SH3 in free form in D₂O. The notations in Figure 6.6 were used except that the deconvolution line width was 20 cm⁻¹ and the enhancement factor was 2.1. See Table 6.6 for the band frequencies.
Figure 6.13. Amide I range of the IR spectra of SH3 in complexes with DMPC. The notations in Figure 6.6 were used except the deconvolution line width was 18 cm\(^{-1}\) and the enhancement factor was 2.3. See Table 6.6 for other parameters.
Table 6.6 Position and fractional areas of the amide I bands of SH3 in solution and in complexes with DMPC

<table>
<thead>
<tr>
<th>Band assignment</th>
<th>SH3 in solution</th>
<th>SH3 in complexes with DMPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-helix</td>
<td>1657 (11%)</td>
<td>1651 (8%)</td>
</tr>
<tr>
<td>β-pleat</td>
<td>1674 (16%)</td>
<td>1672 (8%)</td>
</tr>
<tr>
<td></td>
<td>1629 (12%)</td>
<td>1632 (10%)</td>
</tr>
<tr>
<td></td>
<td>1615 (49%)</td>
<td>1616 (60%)</td>
</tr>
<tr>
<td>random</td>
<td>1643 (8%)</td>
<td>1645 (3%)</td>
</tr>
<tr>
<td>turns</td>
<td>1690 (4%)</td>
<td>1694 (3%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1683 (6%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1662 (1%)</td>
</tr>
</tbody>
</table>

6.3.2.4 Discussion-Hydrophobic Peptides

KIT and B19 showed two "faces" in the "Schiffer-Edmundson" helical wheel expressions and had relatively high hydrophobicities for the nonpolar faces. The hydrophobic moment profiles of these two peptides predicted a possibility of a β-sheet structure for both. Their membrane binding abilities were proven by EM. FTIR studies of KIT suggested conformational changes related mainly to β-sheet structures. However, some increases of α-helix content were observed in complexes of KIT/DMPC. On the amphipathic helical wheel, B19 had smaller nonpolar face and less ion pairs than KIT. So B19 had less amphiphilicity than KIT. A large increase in α-helix content was observed for B19 in complexes with DMPC but β-structure remained the dominant secondary structure.
SH3 was a reversed amphipathic helix (Figure 6.1). It was expected to adopt an α-helix or β-sheet (Figure 6.2). FTIR spectra were typical of self-association complexes whether lipids were present or not.

Table 6.7 The percent secondary structure of the peptides determined by FTIR.

<table>
<thead>
<tr>
<th></th>
<th>α-helix</th>
<th>β-pleat</th>
<th>random</th>
<th>turns</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAP-20</td>
<td>11</td>
<td>73</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>LAP-20/DMPC</td>
<td>15</td>
<td>50</td>
<td>13</td>
<td>22</td>
</tr>
<tr>
<td>KIT</td>
<td>19</td>
<td>40</td>
<td>35</td>
<td>6</td>
</tr>
<tr>
<td>KIT/DMPC</td>
<td>12</td>
<td>59</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>B19</td>
<td>0</td>
<td>74</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>B19/DMPC</td>
<td>33</td>
<td>30</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td>SH3</td>
<td>11</td>
<td>77</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>SH3/DMPC</td>
<td>8</td>
<td>81</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

(Note: Bands at 1612-1616 cm⁻¹ were calculated as intermolecular β-sheet in preparation of this Table)

6.4 The amphipathic helical model

In this study, the amphipathic helical model was studied on four peptides using FTIR spectra. The designed amphipathic model peptide, LAP-20, did not form an amphipathic helix under our experimental conditions. Three other peptides, with different amphiphilicities, displayed three different pictures. The best amphipathic peptide among the three, KIT, showed an increase in β-sheet upon binding to lipids. The second most amphipathic peptide, B19, was the one with a large increase of α-helical content in complexes with DMPC. The last one, SH3, predicted to be a reversed amphipathic
peptide, was possibly in a self-associated state. Our data fit the predictions made based on the hydrophobic moment profiles, which indicated that B19 was the best α-helix and LAP-20, KIT and SH3 could be in β-sheets.

The results from our FTIR study implied a different structure of LAP-20 in solution from the one predicted earlier from circular dichroism (CD) studies which were used to establish the amphipathic helix model. The discrepancy between the two methods (IR and CD) was possibly due to the different scale of concentration of the sample. FTIR spectra require a relatively high concentration while the CD spectra require a low concentration of sample. As a consequence of this, it is possible that the FTIR spectra provide the structural information of peptides in self-association while CD spectra give a picture of a single molecule of peptide. However, both methods should agree when dealing with a peptide with no self-association or in lipid complexes. As mentioned above, CD is unreliable in the presence of lipid complexes due to the particle scattering the incident light. Although there are some drawbacks to using IR to assign secondary structures to protein, it appears to be more reliable than CD.
CONCLUSION

Interactions between proteins and other biological macromolecules have been a major subject of interest in biochemistry for a long time and form the basis of protein structure-function relationships. The amphipathic helical model was proposed from the earlier CD studies on lipoproteins to describe an amphipathic motif in lipoproteins (Segrest, et al., 1976). The motif consists of a helix with special arrangements of hydrophilic and hydrophobic amino acid residues such that the hydrophilic residues occupy one side of the helix (polar face) and the hydrophobic residues distribute at the other side of the helix (apolar face). There is a charge gradient which contains negative charges at the center of the polar face which has positive changes on the periphery.

The amphipathic helix model has been widely used to explain the structural feature of various other proteins (Segrest, et al., 1990). Very recently, the three-dimensional structure of the LDL receptor-binding domain of apo E has been determined at 2.5 Å resolution by x-ray crystallography. The protein formed a four-helix bundle. As predicted by amphipathic theory from the primary sequence, four helices of the protein were arranged with the hydrophobic residues of helices adjacent and the hydrophilic residues exposed to solvent. It was proposed that the amphipathic helical repeats of apolipoprotein A-I were responsible for the binding of HDL to cells (Leblond & Marcel, 1991). However, there is no experimental evidence, apart from CD, that the lipid-binding domains of lipid-associating proteins, including apolipoproteins, are amphipathic helices. The so-called amphipathic helix for lipid binding proteins has not been observed in either x-ray or NMR experiments, which are the best tools for structural determination. Considering the finding of amphipathic β-sheets, which had a hydrophobic residue every 1.1 residue and were the major conformation of B factor of G protein (Orban, et al., 1992), it appeared that the
amphipathic feature exists commonly, but the helix might not be the only ordered structure which could fit the amphipathic model.

The first part of this study was aimed at elucidating the detailed 3D structure of an amphipathic peptide in solution and in complexes with DPC and to test the features of an amphipathic helix in complexes with DPC. An amphipathic peptide, C-PEPTIDE, residues 38 to 53 of apolipoprotein C-I, which occurs in the lipid binding domain is purported to form a strong amphipathic helix in the presence of lipids. The COSY type experiments, TOCSY and DQF-COSY, were used to achieve proton assignments for C-PEPTIDE in solution and in complexes. The unique residues were determined first from either DQF-COSY or TOCSY spectroscopy based on the identification of the spin-system of side chain protons. Other residues were assigned using a procedure called sequential assignment strategy in which NOE information was used to make connections between adjacent residues. For C-PEPTIDE in both solution and complexes, all protons, except a few aromatic protons, were assigned.

The complete proton assignments of C-PEPTIDE enabled me to make a detailed model of C-PEPTIDE/lipid complex. Instead of the predicted amphipathic helix in DPC complexes C-PEPTIDE appeared to have a partial ordered secondary structure and this was also observed with C-PEPTIDE free in solution. This two-domain structure consisted of the N-terminal end of about 10 residues (N domain) and the C-terminal end of 9 residues (C domain). The two domain structure followed from the NOE patterns of C-PEPTIDE in solution and in complexes, from the secondary shift plots of $\alpha$H of C-PEPTIDE and from the results of the proton-deuterium exchange experiments. In solution, the N domain of C-PEPTIDE was in random coil while the C domain formed a partial $\beta$-structure with about 4-5 residues. Experimentally, no NOE connectivities were found for the residues at the N domain and part of the C domain. This observation indicated an unstructured conformation.
for most of C-PEPTIDE in solution. Since weak NOE cross peaks appeared only between αH and NH in the middle of the C domain region and no NH-NH NOE cross peaks existed for the region, a helical conformation was excluded and β-sheet was proposed for this small region of the C-PEPTIDE. The secondary shift plot showed negative peaks which supported the proposed β-sheet structure. A dramatic spectral change was observed for C-PEPTIDE in complexes compared with C-PEPTIDE in solution. In complexes with DPC-d38, C-PEPTIDE experienced a conformational change in the C domain, which adopted a helical conformation with about 2 turns starting from residue 10 or 11, ending at residue 18. The N domain remained unstructured. Cross peaks were dispersed and resolved. There was little change for the N domain in terms of chemical shifts and NOE patterns, which indicated that the N domain stayed in an unstructured conformation. The C-domain experienced the major change which was observed on the NOE connectivities and the secondary chemical shift plot. Continuous NOE connectivities between αH(i) to NH(i+1) spanned nine residues in the C domain and some NH-NH connectivities were also observed. The Medium αH(i)−βH(i+3) NOEs, were found between side chains for the residues from eleven to eighteen in the C domain. These NOE patterns demonstrated a helical conformation for most of the C domain. Positive peaks were displayed on the secondary shift plot which were indications of a helical structure for the C domain.

The apolar face of C-PEPTIDE appeared to differ from the one predicted by the amphipathic helix model because the hydrophobic residues at the apolar face on the Schiffer-Edmundson wheel of C-PEPTIDE did not show a big chemical shift change upon the formation of C-PEPTIDE/DPC complexes. It was certain in the deuterium-proton exchange experiments and from the secondary shift plots that the binding of C-PEPTIDE to micelles happened at the C domain. There were two possible orientations of C-PEPTIDE on a micelle, parallel to the micelle surface and perpendicular to the surface. In the first mode, C-PEPTIDE used one side, possibly the apolar face, to interact with the surface. If
this was the case, the residues on the apolar face should have big chemical shift changes, because of the hydrophobic environment in the lipid layer, while the residues on the polar face should show slight chemical shift changes. In the second mode, C-PEPTIDE used entire C domain to interact with the surface so that all residues in the C domain should have big chemical shift changes. My results did not match either of these two modes. The secondary shift plot of $\alpha$H showed changes for all residues at the C-end of C-PEPTIDE with some residues having greater changes than others, and one residue, out of three, on the apolar face did not show big chemical shift change as expected. In the deuterium-proton exchange experiments, peaks were roughly assigned to the residues at the C terminal of C-PEPTIDE but not to the individual residues. So information of proton exchange rates as indications of accessibility to protons of C-PEPTIDE were not available. The binding mode of C-PEPTIDE to DPC micelles, attaching or inserting, and the detailed helical structure remained unclear.

Three residues of C-PEPTIDE, all located in the N-domain, showed more than one set of signals for some of their protons. The residues were Ala-2, Arg-5 and Trp-7; all of which occur in the flexible N-terminal end of the molecule. Two possible explanations for these multiple peaks are proffered here. One possibility is that C-PEPTIDE exists as a dimer in the lipid complex. A second possibility is that Trp-7, which occurs close to the hinge between the helix and the unstructured domain, assumes two different orientations. An interaction between the Trp and the lipid is given as the reason for this explanation. Further NOE and other studies might detect such an interaction. As for the possibility of dimers existing they must be consistent with the NMR data. The C-domain of the dimers must be magnetically equivalent since all residues in the C-end have single signals per proton. As well, only selected protons of the other two residues were multipeaks. When dimers occur almost all peaks are nonequivalent (see Chapter 5). It is difficult to distinguish
between the two explanations merely on the basis of NMR. For instance, if dimers exist two peptide molecules per particle must be found by accurate measurements.

CD and FTIR are two common and convenient ways for studying protein secondary structures though they only give gross structural features without details. It was found that the accuracy of FTIR and CD results, compared to x-ray results, varied from case to case (Byler and Susi, 1986; Sarver and Krueger, 1991). It seems that CD gives more reliable estimation of α-helical content while FTIR provides information of β-structures. Another controversy between CD and FTIR was in the application to protein-lipid complexes. CD was shown to be not suitable to deal with protein-lipid complexes as the particles are known to scatter light, and CD also overestimates α-helices. FTIR was superior to CD when studying a protein-lipid system as the amounts of helix and β-sheet are resolved in D2O solution. But some drawbacks exist in performing FTIR experiments (see Section 6.3.1.3) and extracting information from those experiments.

In the second part of this study, Fourier transform infrared (FTIR) results were compared with CD results, when applied to a peptide-lipid complex system. An amphipathic model peptide, LAP-20, and three other peptides in both solution and complexes were subjected to FTIR study. Results showed significant β-sheet structures for LAP-20, but not for B19 in complexes or for KIT in solution. Along with the findings that FTIR gave better estimation of β-sheet and CD worked better on helices, this study suggested that the amphipathic theory might need to be examined more carefully, especially the existence of β-sheet conformation.

Finally, it is noteworthy that a significant amount of irregular structure (random and turn structures) was found in complexes of LAP-20, KIT and B19 with lipids similar to that which I observed from the 2D NMR studies of C-PEPTIDE. Considering the fact that turns, or loops, were often found in active sites of various proteins or enzymes, the
irregular structures (random or turn structures) found in this study may imply that "non-structured" elements of the bound peptide may be just as important as the helical moiety for biological activity.

All the findings mentioned above were missed in early circular dichroism (CD) studies on amphipathic peptides, so this NMR and FTIR study brings us closer, at least qualitatively, to the understanding of the details of amphipathic peptide-lipid interactions.

In addition, a direct comparison of NMR results and FTIR results was available from this study where IR and NMR were performed for KIT in solution. IR indicated a significant amount of β-structures. Detailed NMR analysis was precluded because preliminary spectra indicated at least dimers with nonequivalent subunits. The structure of KIT in solution was not resolved due to the fact that many more NH cross peaks were obtained in the TOCSY spectrum than the number of amide groups in a KIT molecule. A rapid qualitative method of determining structure has been suggested by the external examiner, i.e. αH-NH cross peaks occurring for β-structures have been found to low field of the water signal. For KIT, all αH-NH cross peaks were above the water resonance. In another words, chemical shifts for all α protons of KIT were shifted upfield from the water signal. The results argued against a β-sheet structure. The explanation of this discrepancy between FTIR and NMR results is not very clear and more NMR work is necessary.

Several factors need to be considered. Because of the drawbacks of FTIR (see Section 6.3.1.3), the amounts of secondary structures determined by IR must be viewed as qualitative only. On the other hand, NMR experiments can provide a more detailed structure only when NOE information is available. An example is given in this study. From NOE patterns, a helical conformation was located at the C-end of C-PEPTIDE which counts for 40-50% total structure. This meant that the possibility, although remote, of β structures for KIT existed and a direct comparison of β structures of KIT with results of
NMR experiments would not be conclusive. However, FTIR will not be recommended in future structural studies from my experiences because, in a IR experiment, the procedure is not well established and the fundamental theory needs to be strengthened. Much more effort is needed to study the IR method theoretically before application of such method to peptide secondary structure determinations.

In general, apolipoproteins help to stabilize lipoproteins structurally and functionally in the catabolism of lipoproteins. It has been a goal for a long time to understand how the structures of apolipoproteins correspond to their functions. The amphipathic model was used to fit that need. This work extended the amphipathic theory by providing the detailed conformational changes of an amphipathic peptide. This work also suggested that the lipid-binding domain of apo C-I might be shorter than originally proposed. While my work has helped to elucidate the details of the lipid-binding region of apo C-I, it would be difficult to extrapolate these preliminary results to explain apo C-I mechanism of action.

Based on the findings in this study, multidimensional (2-D, 3-D, 4-D) NMR should be the major approach to determine the structures of lipoproteins. FTIR and CD results should be used to provide complementary information only when NMR is not applicable. In order to approach structures of lipoproteins, molecular biology methods are needed to produce isotope labelled proteins which can be used in heteronuclear NMR experiments. Fortunately, all major apolipoproteins have been cloned and the genes of those proteins are available. A bacteria strain, that grows in a minimal culture medium with inorganic sources of $^{13}\text{C}$ and $^{15}\text{N}$, can incorporate isotopes into proteins. The gene, which codes for the protein of interest, can be transferred into bacteria by many vectors. Then, a large quantity of isotope labelled proteins can be produced by growing bacteria in the minimal culture medium. Using resultant proteins and heteronuclear multidimensional NMR techniques,
detailed conformational changes of lipoproteins upon binding to micelles can be studied. It is hoped that a more precise protein-lipid structural model could be extracted from those studies.

A structure of a protein can be obtained from experimental data using several approaches. One approach is the distance geometry algorithm wherein a rigid protein conformation can be obtained by a process called distance geometry in which NOE data are converted into distance constraints between protons in a protein molecule. Local detailed conformation is calculated by an energy minimizing process. A second approach is to use a molecular dynamics algorithm. In a molecular dynamics simulation, an artificial energy penalty function is created and used to reflect the distance violations in a simulation process. Torda et al. (1990) proposed a time averaged method in which time-averaged distances were used to fit to NOE distances. This approach resulted in a relatively flexible conformation.
REFERENCES


-158-


-159-


55. Lauterwein, J.C., C. Bösch, L.R. Brown, and K. Wüthrich (1979) *Biochem. Biophys. Acta* 556, 244-264


60. Mammi, S and E. Peggion (1990) *Biochemistry* 29, 5265-5269


83. Saudek, V., and J.T. Pelton (1990) *Biochemistry* 29, 4509-4515


Appendix

1. Fitting program

```pascal
PROGRAM SIMREST;
VAR
  WORD:STRING[3];
  F:FILE;
  M1:ARRAY[0..255] OF INTEGER;
  BANDS,TOP5:INTEGER;
  RMS,FACTOR:REAL;
  SCALE,BLOCKNO,SECS:INTEGER;
  POINT,TOTALN:INTEGER;
  INTENSIT:ARRAY[230..500] OF REAL;
  WAVENM:ARRAY[230..500] OF REAL;
  SUM: ARRAY[250..500] OF REAL;
  BADC:ARRAY[1..20] OF REAL;
  MAX: ARRAY[1..20] OF REAL;
  HWIDTH:ARRAY[1..20] OF REAL;
  GAUP:ARRAY[1..20] OF REAL;
  CONSA: ARRAY[1..20] OF REAL;
  CONSB: ARRAY[1..20] OF REAL;
  CONSC: ARRAY[1..20] OF REAL;
  CONSD: ARRAY[1..20] OF REAL;
  BANDINT: ARRAY[1..20,230..500] OF REAL;
  AREA: ARRAY[1..20] OF REAL;
  Y: ARRAY[0..1023] OF INTEGER;
  BANDST: ARRAY[1..126] OF INTEGER;
PROCEDURE SPECTRUH;
VAR
  X: ARRAY[0..1023] OF INTEGER;
  I,J,K:INTEGER;
  N,T:REAL;
  XEP,XSP:REAL;
  POINTNM:REAL;
  BLOCKNO:INTEGER;
  LER,RER:REAL;
BEGIN
  RESET(F,‘FNA.DATa=02’);
  J:=BLOCKREAD(F,M1,1,-1);
  IF M1[46]<>0
     THEN T:=1+M1[46]*0.1/1677721.5
     ELSE T:=M1[46]*0.1/1677721.5;
  WRITELN;
  XEP:=M1[45]+T;
  IF M1[48]<>0
     THEN T:=1+M1[48]*0.1/1677721.5
     ELSE T:=M1[48]*0.1/1677721.5;
  XSP:=M1[47]+T;
```
POIUTHM:=M1[41];
FACTOR:=POIUTHM/(XSP-XEP);
SCALE:=1;
FOR I:=M1[44] TO 12 DO SCALE:=SCALE*2;
I:=0;
FOR BLOCKNO:=0 TO 3 DO
BEGIN
    J:=BLOCKREAD(F,X[I],1,BLOCKNO);
    I:=I+256;
END;
WRITE('LEFT END OF THE RANGE =');
READLN(LER);
WRITE('RIGHT END OF THE RANGE =');
READLN(RER);
IF LER>RER THEN
BEGIN
    N:=LER;
    LER:=RER;
    RER:=N;
END;
IF XEP<XSP THEN
BEGIN
    N:=LER;
    LER:=RER;
    RER:=N;
END;
N:=FACTOR*(RER-XEP);
TOTALN:=TRUNC(N)+1;
N:=FACTOR*(LER-XEP);
POINT:=ROUND(N);
TOYO:=TOTALN-POINT;
WRITE('DO YOU WANT TO PRINT OUT THE SPECTRUM? Y OR N. ');
READLN(WORD);
FOR K:=POINT TO TOTALN DO
BEGIN
    WAVE[N]:=XEP+(K)/FACTOR;
    INTENSIT[N]=(X[K]/1024/SCALE);
    IF WORD=YES THEN
BEGIN
  WRITELN('WAVENUMBER= ',WAVENM[K],
            INTENSITY= ',INTENSIT[K]);
  IF ((K DIV 19)*19=K) THEN READLN;
END;
END;
PROCEDURE INPUTARS;
VAR
  N,K:INTEGER;
  TEMPER:REAL;
  CHART:STRING[6];
BEGIN
  WRITELN('READIN = READIN,  INPUT = INPUT');
  READLN(CHART);
  CASE CHART OF
    'READIN': BEGIN
      RESET (F,'FNA.F'ARM=D2');
      SECS:=BLOCKREAD(F,PAE4DSTv1,0);
      BANDS:=BANDST[1];
      FOR N:=1 TO BANDS DO BEGIN
        BADC[N]:=(BANDST(6*N+1))/1677721.5*0.2;
        MAX[N]:=(BANDST(6*N+2))/1677721.5*0.2;
        HWIDTH[N]:=(BANDST(6*N+3))/100000;
        GAU[N]:=(BANDST(6*N+4))/100000;
        CONSA[N]:=MAX[N];
        CONSD[N]:=SQRT(HWIDTH[N]);
        CONSB[N]:=CONSD[N]/LN(2);
        CONSC[N]:= CONSA[N]*CONSD[N];
        WRITELN('BAND NO. = ',N);
        WRITELN('POSITION = ',BADC[N]);
        WRITELN('INTENSITY = ',MAX[N]);
        WRITELN('LINWIDTH = ',HWIDTH[N]);
        WRITELN('GAUSIER % = ',GAU[N]);
      END;
    END;
    'INPUT': BEGIN
      WRITE('HOW MANY BANDS DO YOU HAVE? (>=1)');
      READLN(BANDS);
      FOR N:=1 TO BANDS DO BEGIN
        WRITELN(DASH);
        WRITELN('BAND No. = ',N);
        WRITE('THE BAND CENTRE = ',BADC[N],
                  THE NEW ONE = ');
        READLN(BADC[N]);
        WRITE('THE MAXIMUM INTENSITY = ', MAX[N],
                  THE NEW VALUE = ');
        READLN(MAX[N]);
        WRITE('THE HALF WIDTH OF THE BAND = ', HWIDTH[N],
                  THE NEW ONE = ')
        READLN(HWIDTH[N]);
        WRITE('THE GAUSSIER CONTENT = ', GAU[N],
                  THE NEW PERCENTAGE = ');
  END;
END;
READLN(GAUP[N]);
IF GAUP[N]>1 THEN GAUP[N]:=GAUP[N]/100;
CONSA[N]:=MAX(N);
CONSD[N]:=SQR(HWIDTH[N]);
CONSB[N]:=CONSD[N]/LN(2);
CONSC[N]:=CONSA[N]*CONSD[N];
END;
END;
WRITELN(DASH);
END;
PROCEDURE CALCUL);
VAR
CBAND, SECS,BLOCKNO:INTEGER;
XVAR,REAL;
DVAT,TEMPER:REAL;
M,N: INTEGER;
I,J,K:INTEGER;
CHART: STRING[3];
BEGIN
CBAND:=1;
WRITELN;
WRITELN('YOU HAVE A CHANCE TO MAKE CORRECTION?');
REPEAT
FOR J:=1 TO CBAND DO
BEGIN
WRITELN(DASH);
WRITE('BAND No. = ');READLN(TEMPER);
IF TEMPER)=1.0
THEN N:=ROUND(TEMPER)
ELSE N:=1;
WRITE('THE BAND CENTRE = ',BADC[N],'
THE NEW ONE = ');READLN(TEMPER);
IF TEMPER)>0 THEN BADC[N]:=TEMPER;
WRITE('THE MAXIMUM INTENSITY = ', MAX[N],'
THE CHANGES = ');READLN (TEMPER);
MAX[N]:=TEMPER+MAX[N];
WRITE('THE HALF WIDTH OF THE BAND = ', HWIDTH[N],'
THE NEW ONE = ');
READLN(TEMPER);
IF TEMPER)>0 THEN HWIDTH[N]:=TEMPER;
WRITE('THE GAUSSIER CONTENT = ', GAUP[N],'
THE NEW PERCENTAGE = ');READLN(TEMPER);
IF TEMPER)>1 THEN GAUP[N]:=TEMPER/100 ELSE GAUP[N]:=TEMPER;
CONSA[N]:=MAX(N);
CONSD[N]:=SQR(HWIDTH[N]);
CONSB[N]:=CONSD[N]/LN(2);
CONSC[N]:=CONSA[N]*CONSD[N];
END;
WRITELN(DASH);
WRITELN;
WRITE(’SHOW THE COMPARISON? ’);
READLN(CHART);
IF CHART=’YES’ THEN
BEGIN
WRITELN(DASH);
WRITELN(’WAVENUMBER: INTENSITY—SIMULATION= DEVIATION’);
WRITELN(DASH);
END;
FOR N:=1 TO BANDS DO AREA[N]:=0;
RMS:=0;
FOR K:=POINT TO TOTALN DO
BEGIN
M:=0;
SUM[K]:=0;
FOR N:=1 TO BANDS DO
BEGIN
XVAR:=SQR(WAVENM[K]-BADC[N]);
BANINT[N,K]:=GAUP[N]*CONSA[N]*EXP(-XVAR/CONSB[N]);
BANINT[N,K]:=BANINT[N,K]+(1-GAUP[N])*CONSC[N]/(CONSD[N]+XVAR);
IF K=POINT THEN AREA[N]:=AREA[N]+(BANINT[N,K]+BANINT[N,K-1])/2/ABS(FCTOR);
SUM[K]:=SUM[K]+BANINT[N,K];
IF (CHART=’YES’) AND (XVAR<0.0001) THEN M:=N;
END;
DVAT:=INTENSIT[K]-SUM[K];
RMS:=SQR(DVAT)+RMS;
IF CHART=’YES’ THEN
BEGIN
WRITE(WAVENM[K]:10:5,INTENSIT[K]:11:8,SUM[K]:11:8,DVAT:11:8);
IF (M=1) OR (M=BANDS) THEN WRITE(‘* ’,SPACE,BANINT[M,K]:11:8);
IF (M=1) AND (M=BANDS) THEN
WRITELN(’* ’,BANINT[M-1,K]:11:8,BANINT[M,K]:11:8,BANINT[M+1,K]:11:8)
ELSE WRITELN;
IF ((K DIV 16)*16=K) THEN READLN;
END;
END;
RMS:=SORT(RMS/TOPO);
WRITELN(DASH);
WRITELN;
WRITELN(’***** RMS= ’,RMS,’ *****’);
WRITELN;
WRITE(’HOW MANY BANDS DO YOU WANT TO CORRECT? ’);
READLN(CBAND);
UNTIL CBAND=0;
END;
PROCEDURE STOR File;
VAR
BLOCKNO,I,J,N:INTEGER;
BEGIN
WRITE(‘HAVE YOU DONE THE FITTING? ’);
READLN(WORD);
IF WORD=’YES’ THEN BEGIN
  WRITELN(DASH);
  WRITELN(’BAND No. * WAVENUMBER * INTENSITY * BAND WIDTH * GAUSIER % * AREA ’);
  WRITELN(DASH);
  FOR N:=1 TO BANDS DO BEGIN
    WRITELN(N:7,’ ,BADC[N],’ ,MAX[N],’ ,HWIDT[N],’ ,GAU[N],’ ,AREA[N]);
    CASE N OF
      1: REWRITE(F,’BAND1.DATA=D2’);
      2: REWRITE(F,’BAND2.DATA=D2’);
      3: REWRITE(F,’BAND3.DATA=D2’);
      4: REWRITE(F,’BAND4.DATA=D2’);
      5: REWRITE(F,’BAND5.DATA=D2’);
      6: REWRITE(F,’BAND6.DATA=D2’);
      7: REWRITE(F,’BAND7.DATA=D2’);
      8: REWRITE(F,’BAND8.DATA=D2’);
      9: REWRITE(F,’BAND9.DATA=D2’);
     10: REWRITE(F,’BAND10.DATA=D2’);
     11: REWRITE(F,’BAND11.DATA=D2’);
     12: REWRITE(F,’BAND12.DATA=D2’);
     13: REWRITE(F,’BAND13.DATA=D2’);
     14: REWRITE(F,’BAND14.DATA=D2’);
     15: REWRITE(F,’BAND15.DATA=D2’);
     16: REWRITE(F,’BAND16.DATA=D2’);
     17: REWRITE(F,’BAND17.DATA=D2’);
     18: REWRITE(F,’BAND18.DATA=D2’);
     19: REWRITE(F,’BAND19.DATA=D2’);
     20: REWRITE(F,’BAND20.DATA=D2’);
    END;
    SECS:=BLOCKWRITE(F,M1,1,-1);
    FOR I:=0 TO 1023 DO Y[I]:=0;
    FOR I:=POINT TO TOTALN DO
      Y[I]:=ROUND(BANINT[N,I]*1024*SCALE);
    J:=0;
    FOR BLOCKNO:=0 TO 3 DO BEGIN
      SECS:=BLOCKWRITE(F,Y[J],1,BLOCKNO);
      J:=J+256;
    END;
  END;
END;
END;
BANDST[1]:=BANDS;
BANDST[2]:=ROUND(RMS*1000000);
FOR N:=1 TO BANDS DO
BEGIN
  BANDST[6*N]:=TRUNC(BADC[N]);
  BANDST[6*N+1]:=ROUND((BADC[N]-BANDST[6*N])*1677721.5/0.2);
  BANDST[6*N+2]:=ROUND(MAX[N]*1677721.5/0.2);
  BANDST[6*N+3]:=ROUND(HWIDTH[N]*100000);
  BANDST[6*N+4]:=ROUND(GAUF[N]*100000);
  BANDST[6*N+5]:=ROUND(AREA[N]*100000);
END;
REWRITE (F,'FNA.PARM=D2');
SECS:=BLOCKWRITE(F,BANDST,1,0);
CLOSE(F,LOCK);
REWRITE(F,'BAND_DATA=D2');
SECS:=BLOCKWRITE(F,M1,1,-1);
FOR I:=0 TO 1023 DO Y[I]:=0;
FOR I:=POINT TO TOTALN DO Y[I]:=ROUND(SUM[I]*1024*SCALE);
J:=0;
FOR BLOKNO:=0 TO 3 DO
BEGIN
  SECS:=BLOCKWRITE(F,Y[J],1,BLOKNO);
  J:=J+256;
END;
CLOSE(F,LOCK);
END;

BEGIN
  SPECTRUM;
  INPUTPARS;
  CALCUL;
  STOREFIL;
  WRITELN;
  WRITELN;
  WRITELN('SPACE,'SPACE,'MERRY CHRISTMAS, HAPPY NEW YEAR !');
  WRITELN;
END.
2. Secondary derivative

```
PROGRAM SEDERIVATIVE;
VAR
  F:FILE;
  I,J,SEC,BLOCKNO,K:INTEGER;
  DIVI,FIRSTDATA,LASTDATA:REAL;
  ELE: ARRAY[0..255] OF INTEGER;
  RANDATA: ARRAY[0..1280] OF INTEGER;
  DATA:ARRAY[0..1280] OF INTEGER;

BEGIN
  WRITELN;
  RESET(F,'INFILE.DATM');
  SEC:=BLOCKREAD(F,ELE,1,-1);
  IF (ELE[46])=0)
    THEN FIRSTDATA:=ELE[45]+ELE[46]*0.1/1677721.5
    ELSE FIRSTDATA:=ELE[45]+1+ELE[46]*0.1/1677721.5;
  IF (ELE[48])=0)
    THEN LASTDATA:=ELE[47]+ELE[48]*0.1/1677721.5
    ELSE LASTDATA:=ELE[47]+1+ELE[48]*0.1/1677721.5;
  BLOCKNO:=ROUND(ELE[41]/256);
  DIVI:=ABS(LASTDATA-FIRSTDATA)/ELE[41];
  I:=0;
  FOR J:=0 TO BLOCKNO DO
    BEGIN
      SEC:=BLOCKREAD(F,RAWDATA[I],1,J);
      I:=I+256;
    END;
  DATA[I]:=O;
  DATA[ELE[41]-1]:=O;
  K:=ELE[41]-2;
  FOR I:=1 TO K DO
    DATA[I]:=ROUND(4*(RAWDATA[I]+1)+RAWDATA[I-1]-2*RAWDATA[I])/DIVI/DIVI;
  REWRITE(F,'OUTFILE.DATM=DATA=D2');
  SEC:=BLOCKWRITE(F,ELE,1,-1);
  I:=0;
  FOR J:=0 TO BLOCKNO DO
    BEGIN
      SEC:=BLOCKWRITE(F,DATA[I],1,J);
      I:=I+256;
    END;
  CLOSE(F,LOCK);
END.
```