NMR STUDIES OF MEMBRANE-BINDING PEPTIDES IN MONOOLEIN CUBIC PHASES

by

Laila Maria Rani Singh

Honours Bachelor of Science, University of Waterloo, 1997

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the

Department of Molecular Biology and Biochemistry

SIMON FRASER UNIVERSITY

Summer 2005

Approval Page

Name: Laila Maria Rani Singh
Degree: Doctor of Philosophy

Title of thesis: NMR Studies of Membrane-Binding Peptides in Monoolein Cubic Phases

Examining Committee:

Chair: Dr. Peter J. Unrau, Assistant Professor

Dr. Jenifer L. Thewalt, Associate Professor
Department of Molecular Biology and Biochemistry, Simon Fraser University
Senior Supervisor

Dr. Rosemary B. Cornell, Professor
Department of Molecular Biology and Biochemistry, Simon Fraser University
Committee Member

Dr. Lawrence P. McIntosh, Professor
Department of Biochemistry and Molecular Biology, University of British Columbia
Committee Member

Dr. Mark Paetzel, Assistant Professor
Department of Molecular Biology and Biochemistry, Simon Fraser University
Public Examiner

Dr. Richard M. Epand, Professor
Biochemistry Department, McMaster University
External Examiner

Date Approved: Wednesday August 3rd, 2005

Abstract

Membrane proteins are of particular interest to structural biologists since it is believed that these proteins comprise approximately thirty percent of the proteins encoded for by genomes. Despite the biological significance of these proteins, the atomic resolution structures of membrane proteins are being solved at a much slower rate than their soluble counterparts. This is primarily due to difficulties that arise in maintaining the lipid interactions required to retain the structural and functional integrity of the proteins during structural studies. One approach to resolving this problem would be to develop alternative membrane-mimetic environments for structural studies of membrane proteins. The cubic phases formed by mixtures of lipids and water have been identified as such an environment. In recent years cubic phases have been used to crystallize membrane proteins for structural studies using X-ray crystallographic techniques. This application of cubic phases, along with information provided from previous studies, indicated that cubic phases possessed the necessary properties to make them ideal membrane-mimetic environments for solution NMR studies of membrane proteins. I investigated the suitability of the cubic phases formed by mixtures of monoolein and water as membrane-mimetic environments for the solution NMR study of incorporated membrane proteins. For my studies I used two transmembrane peptides (WNALAAVAAALAAVAAALAAGKSKSKS and alamethicin), and one membrane surface-associating peptide (methionine-enkephalin), as models of membrane proteins. In the NMR spectra collected on the transmembrane peptides, only the residues found in the interfacial regions of the cubic phase were observed, whereas all of the residues of the membrane surface-associating peptide were observed. In order to gain insights into the behaviour of the incorporated peptides, the diffusion rates of lipid, peptide and water molecules were measured in peptide-containing cubic phases using solution NMR techniques. The data that I have collected provide the first examples of solution NMR spectra collected on peptides incorporated into a lipid cubic phase. The NMR spectra that were obtained suggest that the monoolein:water cubic phase may be a suitable membrane-mimetic environment for the study of membrane surface-associating peptides and proteins, and the inter-helical loops of integral membrane proteins.

Reader's Summary

Membrane proteins are the proteins associated with the biological membranes of all cells. These proteins perform many important functions which are essential for life, including the transportation of molecules across cell membranes, the conversion of energy and the communication of signals into, and out of cells. It has been estimated that about 30 % of the genes in a genome encode for membranes proteins. Due to their abundance, and involvement in numerous biological processes, there is a considerable amount of interest in understanding both the structure and function of membrane proteins as they represent many important therapeutic drug targets. It is believed that a complete understanding of the mechanisms of action of a protein can only be achieved following the elucidation of a protein's structure to atomic resolution. There are many examples of membrane proteins that perform remarkable functions. One of these proteins is the human aquaporin AQP1. In order for our kidneys to filter 180 liters of blood per day, water must permeate through human AQP1 at a rate of 3 *billion* water molecules per second (Werten *et. al.*, 2002). This example illustrates why there has been, and still is, a considerable amount of interest in understanding how membrane proteins function. However, despite their biological importance, the high resolution structural information available about membrane proteins pales in comparison to the information available for soluble proteins (which make up the majority of the proteins in a cell). By the end of 2003, only 75 unique structures of membrane proteins had been solved to atomic resolution, compared to the greater than 3000 unique structures that were available for soluble proteins (White, 2004). In order to study the structure of membrane proteins it is necessary to remove them from their native environment, the biological membrane, without disrupting their structure. My PhD. research has focused on developing a technique for studying membrane protein structure using nuclear magnetic resonance (NMR) spectroscopy. I have been investigating the use of the cubic phase formed by mixtures of lipids and water as a mimic of biological membranes (which are primarily composed of lipids in the lamellar phase). Cubic phases are an arrangement of lipids that have many characteristics that are similar to those of biological membranes, yet the cubic phase is amenable to studies using solution NMR spectroscopy (whereas lamellar phases are not). For my research I studied peptide models of membrane proteins. The data that I have collected provide the first examples of solution NMR spectra collected on peptides incorporated into a cubic phase, and suggest that cubic phases may be suitable membrane-mimetic environments for structural studies of the peptides and proteins that are associated with biological membranes.

To my parents Agneta and Kishore for their enthusiasm and commitment to lifelong learning.

Professional Acknowledgements

I would like to thank my supervisor, Dr. Jenifer Thewalt for giving me the opportunity to pursue this very interesting and challenging research project for my PhD. I would also like to thank my supervisory committee members, Dr. Rosemary Cornell and Dr. Lawrence McIntosh, for their continued support and enthusiasm during my tenure as a graduate student at SFU.

I would like to thank the following people whose generous contributions made the successful completion of this thesis possible:

First and foremost I would like to thank Dr. Brian Sykes in the Department of Biochemistry at the University of Alberta for taking an active interest in my project and for collecting, and helping me to interpret, my NMR data. I wish to extend my thanks to members of the Sykes lab (past and present) for answering my many questions and for making me feel welcome in the lab during my visits, in particular: Peggy, Tharin, Grant, Steffen, Jan, Monica, Xu, Angie and Robert. I also wish to thank Dr. Leo Spyracopoulos for the collection of NMR data and Dr. Ron McElhaney for helpful discussions about transmembrane peptides and cubic phases.

I would like to thank Dr. John Vederas in the Department of Chemistry at the University of Alberta for his continued interest and support of my project over the years through the provision of labelled and unlabelled samples of methionine-enkephalin, the use lab resources and expertise and valuable discussions. I wish to thank Dr. Nathaniel Martin for the synthesis of the methionine-enkephalin samples while he was a PhD. student in the Vederas lab. I also wish to thank current and former members of the Vederas lab, in particular: Tara, Kamaljit, Belen, Hanna and Andrew for technical help and excursions to the Powerplant and Elk Island.

I would like to thank Dr. Joe O'Neil in the Department of Chemistry at the University of Manitoba for his generosity in providing a significant quantity of labelled alamethicin for my NMR studies, and for a number of helpful discussions regarding the NMR of membrane proteins.

I would like to thank Dr. Peter Tieleman in the Department of Biological Sciences at the University of Calgary and Dr. Siewert-Jan Marrink at the University of Gröningen, The Netherlands, for building a model of the monoolein:water G-type cubic phase and attempting to run molecular dynamics simulations of the system. I also wish to thank Peter for inviting me to visit his lab to begin to learn how to set-up and run molecular dynamics simulations of peptides in monoolein bilayers, and for many stimulating discussions regarding cubic phases.

I would like to thank Dr. Jim Davis in the Department of Physics at the University of Guelph for his interest in this project, for the collection of solid state NMR data and for informative

discussions. I also wish to thank two former members of the Davis lab, Dr. Christophe Farés and Dr. Scott Houliston for assistance with data collection and analysis, interesting discussions and moral support.

I would like to thank Dr. Valerie Booth in the Department of Biochemistry at Memorial University of Newfoundland for her patience and persistence in coaching me through the process of NMR structure calculation and analysis. I also wish to thank Valerie for her mentorship in many matters related to research and graduate school, and for many enjoyable lunches and dinners in Montréal.

At the University of British Columbia, I would like to thank Dr. Mark Okon in the Laboratory for Molecular Biophysics for running many NMR experiments on my samples over the years, in particular the experiments to measure molecular diffusion. I would like to thank Dr. Suzana Straus in the Department of Chemistry for many helpful discussions of my NMR data and my project in general. I would like to thank past and present members of the McIntosh lab, in particular Dr. Cameron Mackereth, for technical help and advice in many areas related to graduate studies.

At Simon Fraser University, I would like to thank Dr. Alan Tracey and Marcy Tracey for assisting me in getting this project started through their help with NMR and invaluable advice regarding working with liquid crystals. I wish to thank Linda Pinto for tirelessly answering my microbiology and biochemistry questions. I am grateful to Dr. Mark Paetzel for lending me an HPLC column, to Dr. Jamie Scott for HPLC use, and to Oscar Pan for assistance in running the HPLC. I wish to thank Sara Leo and Chris Robertson, two outstanding undergraduate students who helped me optimize peptide overexpression and purification. I wish to thank past and present members of Jenifer Thewalt's lab, in particular: Elana Brief, Ya-Wei Hsueh, Amy Rowat, Sherry Leung, John Cheng, Carolyn Young, Gillian Royle, Myrna Monck, Christian Code and Ralph Giles. I wish to thank many of the faculty at SFU for their support over the years, in particular: Drs. Andy Beckenbach, Mike Smith, Dave Baillie, Esther Verheyen, Nancy Hawkins, Willie Davidson, Erika Plettner and Dipankar Sen. I wish to thank past and present SFU post docs, research associates and staff members, in particular: Ethel Hammerly, Duncan Napier, Kathleen Fitzpatrick, Christine Crosby, Kate Scheel, Theresa Kitos, Kathryn Coukell, Nancy Suda and Karen Beckenbach for 'going the extra mile' for all members of the Department. I wish to thank the MBB, Biology and Chemistry graduate students for their support of fellow graduate students over the years.

Personal Acknowledgements

I wish to thank my Mom, my Dad and my brother Zubin for their encouragement, support and understanding during my PhD. A special thanks goes to my husband Jim whose unwavering support helped me make it through some very tough times during my graduate studies.

I wish to thank my family in South Africa, in particular: uncle Rajen, great auntie Devi, uncle Ami and auntie Dewa, auntie Roshini, auntie Kasthuri and cousins Sanjay and Shikha, and my family in Sweden: Mormor, aunt Birgitta and uncle Roland and cousins Lotta and Annika, for their encouragement and understanding over the past 8 years.

I wish to thank my friends for their support and understanding while I was a graduate student, especially: Peggy and Chris, Cameron, Tara, Leah, Bomina, Michelle, Cathy, Rasmus, David, Annick and Kevin, Anat, Connie, Jillian, Melissa, Andrea, Valerie, Jennifer and Steve, Markus and Lisa, Susanne, Jason, Diana and Cris, Amanda, Ian, Adrienne D., Adrienne B., Morven, Alison and Terry, Tricia, Kath, John, Caroline, Nicole, Jilly, Nat, Marika, Paulina, Pete and Tina, Jacqui and Susan. I feel very lucky to have had the opportunity to meet and become friends with so many wonderful people over the years.

I wish to thank Dr. Mark Abramovitz for being a fabulous supervisor and mentor and for his contagious enthusiasm for research and unsolicited positive feedback. I wish to thank Dr. Liz Meiering for her excellent teaching and mentorship at a crucial point in my career. I wish to extend a sincere thanks to both Mark and Liz for teaching me many of the skills that were essential for success in graduate school.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

A

AFM	atomic force microscopy
APS	ammonium persulfate

B

bp	base pair (DNA or RNA)

C

CD	circular dichroism
CNBr	cyanogen bromide
COSY	correlated spectroscopy
CSI	chemical shift index

D

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
DAG	diacylglycerol
DCM	dichloromethane
DCPC	dicaproylphosphatidylcholine
DMPC	dimyristoylphosphatidylcholine
DNA	deoxyribonucleic acid
DOPC	dioleoylphosphatidylcholine
DOPE	dioleoylphosphatidylethanolamine
DOPS	dioleoylphosphatidylserine
DPC	dodecylphosphocholine
DPPC	dipalmitoylphosphatidylcholine
dsDNA	double stranded DNA
ssDNA	single stranded DNA
DSS	2,2-dimethyl-2-silapentane-5-sulfonic acid
DTT	dithiothreitol

E

E.coli	*Escherichia coli*
EM	electron microscopy

F

Fmoc	*O*-fluorenylmethoxycarbonyl
FT	Fourier transformation
FW	formula weight

G

g	grams

× g	times gravity

H

h	hour
HPLC	High Performance Liquid Chromatography
HSQC	heteronuclear single quantum coherence
HMQC	heteronuclear multiple quantum coherence

I

IPTG	isopropyl-β-D-thiogalactopyranoside

K

kb	kilo bases or kilo basepairs of DNA or RNA
kDa	kilo Dalton
KSI	ketosteroid isomerase

L

lysoPC	lysophosphatidylcholine

M

MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization Time of Flight
MAS	magic angle spinning
mL	millilitre
MO	monoolein, 1-monooleoyl-rac-glycerol, C18:1c9
MW	molecular weight

N

ng	nanogram
NMR	Nuclear Magnetic Resonance spectroscopy
NOE	nuclear Overhauser enhancemen
NOESY	nuclear Overhauser enhancement spectroscopy

P

PC	phosphatidylcholine
PCR	polymerase chain reaction
PE	phosphatidylethanolamine
pH	p of hydrogen in solution
pI	isoelectric pH
pmol	picomole
PO	palmitoylolein, 1-monopalmitoleyl-rac-glycerol, C16:1c9
POPC	palmitoyloleoylphosphatidylcholine
POPG	palmitoyloleoylphosphatidylglycerol
PS	phosphatidylserine
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate

R

RMSD	root mean square deviation
RNA	ribonucleic acid
rpm	revolutions per minute

S

SDS	sodiumdodecylsulfate

T

TE	Tris-EDTA
TEM	transmission electron microscopy
TEMED	N,N,N',N'-tetramethylethylenediamine
TFE	trifluoroethanol
TOCSY	total correlation spectroscopy
TCP	total cellular protein

U

UV	ultraviolet light

V

vis	visible light
vol/vol	volume/volume

W

wt/wt	weight/weight

Greek Symbols

μ

μL	microlitre
μg	microgram

Chapter 1: Introduction

1.1 Membrane Proteins

Membrane proteins represent an important class of proteins that are encoded for by about 30 % of the genes in a typical genome (Werten *et al.*, 2002). These proteins are associated with biological membranes where they perform many vital functions in cells, ranging from solute transport and the conversion of energy, to signal transduction (Landau and Rosenbusch, 1996). Although the term 'membrane protein' is often associated with *integral* membrane proteins, those proteins with regions which span both leaflets of a biological membrane, it also includes *peripheral* membrane proteins, the proteins that are associated with the surfaces of biological membranes. Peripheral membrane proteins can be *reversibly* or *irreversibly* associated with membranes depending on the nature of the protein-membrane interaction. Proteins can interact with membranes via an amphiphilic membrane binding domain, they can be 'tethered' to the membrane via a lipid anchor or they can associate with the membrane through electrostatic interactions. The original fluid mosaic model of biological membranes with embedded membrane proteins is shown in Figure 1.1. The different modes of membrane protein binding to biological membranes are shown in Figure 1.2. Proteins which reversibly associate with membranes (i.e. they can exist in both water soluble and membrane bound forms) are called 'amphitropic proteins' (Epand, 1998). These proteins often interact with membranes using a combination of the possible membrane association mechanisms shown in Figure 1.2. The activity of these types of proteins is often modulated by membrane binding.

In the case of integral membrane proteins, there are only two membrane-spanning structural motifs that have been observed, either α-helices or β-barrels. This is because these types of structures allow for sufficient internal hydrogen bonding to compensate for the high energetic

Figure 3 from Singer & Nicolson (1972) *Science* **175:**720-731.

Figure 1.1. Fluid Mosaic Model of Biological Membranes

Figure 10 from Jensen & Mouritsen (2004) *Biochimica et Biophysica Acta* **1666:**205-226.

Figure 1.2. Schematic Illustration of the Modes of Membrane Protein Association with Biological Membranes

cost of dehydrating the peptide bonds upon insertion into a membrane (White, 2003). The vast majority of integral membrane proteins that have been identified to date are helical in nature. The principal structural features of these helical polytopic membrane proteins are transmembrane helices connected by interhelical loop(s), a motif that is identified in 20 - 30 % of the proteins encoded in a typical genome (Wallin and von Heijne, 1998). There is a considerable amount of interest in understanding both the structure and function of membrane proteins as they represent several crucial therapeutic drug targets (Rouhani *et al.*, 2002; Fernández and Wüthrich, 2003).

1.1.1 Biological Membranes

Membrane proteins reside in, or are associated with, biological membranes. Biological membranes play many important roles in cells. In addition to acting as the interface between a cell and the outside world, they are involved in dividing the cell into many intracellular compartments with diverse biological functions. The physical properties of membranes, such as fluidity, bilayer thickness, interfacial polarity, net charge and curvature strain are involved in modulating the functions of membrane proteins within cells.

1.1.1.1 Lipid Modulation of Membrane Protein Activity

The activity of membrane proteins is often dependent on the lipid environment in which they reside, or on the presence of specific lipids (Lee, 2004). There are many mechanisms by which lipids can modulate the activity of membrane proteins. For example, the propensity of a membrane to form non-lamellar phases can modulate the activity of membrane-bound proteins (Epand, 1998). Changes in the hydrophobic thickness of the lipid bilayer can also have an effect on membrane protein organization, and direct or indirect effects on membrane protein activity. The discovery of the roles that lipids and membranes play in modulating the activity of membrane proteins has helped to change the perception of the role of membranes from an inert structural one, to an active one involved in many important cellular functions.

Membrane Thickness

In studies of membrane protein activity upon reconstitution in lipid bilayers, it has been observed that the functional activity of the proteins is affected by the thickness of the bilayer (Dumas *et al.*, 1999). It is believed that this effect is modulated through mismatch between the hydrophobic length of the membrane spanning segments of the proteins and the hydrophobic thickness of the bilayer (Killian, 2003). In cases of 'positive' mismatch where the transmembrane segments of a protein are *longer* than the hydrophobic thickness of the bilayer, hydrophobic side chains may protrude out of the bilayer into a polar environment which would be energetically unfavourable. Consequently a protein may: oligomerize (thus shielding the exposed residues from exposure to a

polar environment), alter its backbone conformation, tilt away from the bilayer normal (thereby reducing its effective transbilayer length), or change the orientation of side chains near the lipid/water interface. In cases of 'negative' mismatch where the transmembrane segments of a protein are *shorter* than the hydrophobic thickness of the bilayer, a protein may respond in a similar manner as for 'positive' mismatch, except for the tilting of transmembrane segments. It is also possible that changes in lipid organization within a membrane may occur in response to mismatch. Such a response could include stretching or disordering of lipid acyl chains, or reorganization of the lipids such that the ones which best matched the hydrophobic length of a protein would become preferentially located in the immediate vicinity of the protein.

1.1.2 Factors Influencing Membrane Protein Structure

There are a number of physical influences that play important roles in determining the structure of a membrane protein. These include the interactions of membrane proteins with water, each other, the hydrophobic interior of the bilayer, the interfacial region of the bilayer and cofactors (White, 2003). Membrane proteins can interact with the lipids in cell membranes in a variety of ways. These can include hydrophobic interactions with the hydrocarbon interior of the bilayer and electrostatic, hydrogen bonding and dipolar interactions with the lipid/water interfacial region of the bilayer (Killian, 2003). Another factor influencing a proteins' structure in a lipid bilayer is the lateral pressure of the membrane (van den Brink-van der Laan *et al.*, 2004). It is believed that the release of this pressure plays an important role in the observed lability of membrane proteins upon detergent solublization.

1.1.3 Challenges of Studying Membrane Proteins

In order to study membrane proteins using most standard biochemical techniques, it is first necessary to remove the proteins from their native environment, the biological membrane. This is often done through detergent solubilization of the membrane in which the proteins reside. Once the proteins are removed from the membranes, they are exposed to environments with properties that are significantly different from those in their native environment. Consequently, membrane proteins are often prone to denaturation or aggregation once they have been extracted out of their native bilayer environment (Landau and Luigi Luisi, 1993; Hasler *et al.*, 1998). It has however been found that many factors can influence the stability of purified membrane proteins. These include the purification procedure, pH, ionic strength, detergents, protein concentration and the presence of specific protein-bound lipids (Werten *et al.*, 2002). It is therefore necessary to optimize the conditions used to reconstitute and study each membrane protein of interest because as a general rule, a procedure that works for one membrane protein will rarely work for another.

When working with membrane proteins that contain cysteine residues, large amounts of reducing agents need to be added to samples to maintain the native protein conformation and to prevent aggregation. One approach that has been used to circumvent the need for large amounts of reducing agents has been to replace the cysteine residues with serine residues. This strategy was successfully employed in structural studies of the MerF protein where cysteine to serine substitutions increased both the length of time, and the protein concentrations, over which samples were stable (Howell *et al.*, 2005).

1.1.3.1 Membrane Protein Expression, Solubilization and Purification

More often than not, membrane proteins cannot be purified from their natural sources in the amounts required for structural, and other types of biochemical studies. It is therefore necessary to use an efficient heterologous system for the expression of large quantities of the proteins of interest. There are currently four major types of expression systems available for these purposes. These involve the use of bacteria or yeast cultures, or insect or mammalian cell lines for membrane protein expression. It is desirable to use an expression host that most closely resembles the organism from which the protein originated. Mammalian proteins often require specific lipid environments, the presence of certain chaperones, and specific post-translational modifications otherwise misfolding or loss of functionality can occur when the protein is expressed. So for mammalian proteins, a mammalian expression system would likely give the best results in terms of the structure and function of the expressed protein, however these types of expression systems are quite costly, require special handling procedures and often yield less protein per liter of culture than other methods. It is therefore advisable to investigate a number of different expression systems to determine which one will give the highest yields of functional protein at the lowest cost, in particular when proteins are being isotopically labelled for NMR studies (Fernández and Wüthrich, 2003). A well-designed and flexible cloning procedure that allows for the efficient cloning of the cDNA of interest into various expression systems for optimization of protein production is recommended.

Once a suitable host has been identified for overexpression of the protein of interest, the protein must be purified. This involves the screening of detergents to identify one that solubilizes the protein efficiently, allows for the purification of the protein, does not disturb the structure and function of the protein, and allows for reconstitution of the protein into the desired system for study. This can be a tedious and time consuming task. In some cases it may also be necessary to strip the membranes of other membrane-associated proteins prior to solubilization of the cell membranes and purification of the protein of interest. The solubilized protein can be purified using classical chromatographic methods, or using affinity purification if the appropriate tags

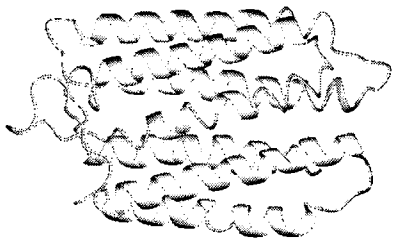
(i.e. histidine-tags) have been engineered into the protein. Since the tags can change the properties of the protein, they should ideally be cleaved off following purification.

1.1.4 Why is structure determination so important?

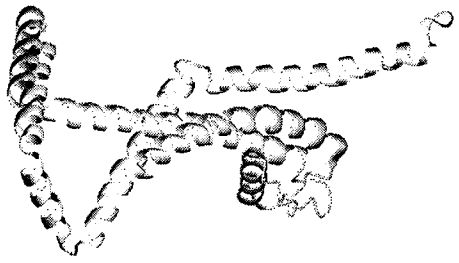
A complete understanding of the mechanisms of action of a protein can only be achieved following the elucidation of a proteins' structure to atomic resolution. The current level of understanding of structure-function relationships in membrane proteins still has a long way to go before it will be comparable to the level of understanding that has been achieved for soluble proteins. One of the aspects of membrane protein structure determination which is particularly challenging for structural biologists is the structure determination of the transmembrane regions of proteins. Often the important biological functions of a protein, such as signal transduction or ion transport, are mediated through the transmembrane domain. However, very little is currently known about the structure and function of the transmembrane domains of proteins at the atomic level.

One of the most extensively studied membrane proteins has been bacteriorhodopsin. It is comprised of seven transmembrane helices that are neatly packed into a bundle. The apparent simplicity of this 'archetypal' membrane protein has encouraged the belief that membrane protein structure prediction should be relatively straightforward: first identify the transmembrane segments using hydropathy plots and then apply helix-packing constraints (White, 2003). However, the recent determination of the three-dimensional (3D) structure of the ClC chloride channel demonstrated that these methods could not be relied on to accurately predict the location or topology of the 17 helices of this protein (Dutzler *et al.*, 2002). Hydropathy plot analyses also failed to correctly identify the helical hairpins that are part of the 3D structure of the KvAP voltage-gated potassium channel (Jiang *et al.*, 2003). In the case of the bacterial mercury transport protein MerF, its atypical hydropathy plot had led researchers to propose structural and mechanistic models that were found to be inconsistent with the 3D structure of the protein that was recently solved using solution NMR techniques (Howell *et al.*, 2005). The structures of bacteriorhodopsin, MerF, the ClC chloride channel and the KvAP voltage-gated potassium channel are shown in Figure 1.3.

There are a number of membrane proteins that perform some remarkable functions, so it is not surprising that there is a considerable amount of interest in understanding how these proteins perform these functions. Since structure often leads to insight into protein function, there has been a drive to solve the structures of some of these membrane proteins in hopes that the structures will provide insights into the mechanisms by which these proteins perform their functions. An example of one of these remarkable proteins is human aquaporin AQP1. In order



Bacteriorhodopsin
PDB Accession Code 1WAZ



KvAP Potassium Channel
PDB Accession Code 1ORQ

CIC Chloride Channel
PDB Accession Code 1KPK

Mercury Transport Protein
PDB Accession Code 1WAZ

Figure 1.3. Structures of the Membrane Proteins Bacteriorhodopsin, KvAP Voltage-Gated Potassium Channel, CIC Chloride Channel and MerF

for the kidneys to filter 180 liters of blood per day, water must permeate through human AQP1 at a rate of 3 billion water molecules per second (Werten *et al.*, 2002). So there has been considerable interest in understanding how these channels function. How can they prevent the flow of protons, yet specifically allow water to pass through at such a high rate? The atomic level structural information provided by the crystal structure of AQP1 has been used in conjunction with molecular dynamics simulations to try to elucidate how this protein functions at the atomic level (de Groot and Grubmüller, 2001; Sui *et al.*, 2001).

1.2 Membrane Protein Structure Determination

Despite their biological importance, the high resolution structural information available on membrane proteins can be considered sparse compared to that of their soluble counterparts. By the end of 2003, only 75 unique structures of membrane proteins had been solved to atomic resolution, compared to the greater than 3000 unique atomic resolution structures that were available for soluble proteins (White, 2004). Most of the membrane protein structures that had been determined were of proteins of bacterial origin, so membrane protein conformational space was not (and still is not) very well represented in the protein structure data base. In 2002, of the 67 membrane protein structures that had been deposited into the protein data base, 52 were of bacterial origin (Werten *et al.*, 2002). In general, it appears that bacterial membrane proteins are more easily produced, purified and crystallized than eukaryotic proteins. Since many membrane proteins are potential therapeutic targets, there is considerable interest in solving the structures of both prokaryotic and eukaryotic membrane proteins so that the function, dynamics and ligand interactions of these proteins can be more well understood.

Three principal techniques have been successfully applied to the high resolution structure determination of membrane proteins. These are: solution state nuclear magnetic resonance (NMR) spectroscopy, solid state NMR spectroscopy and X-ray crystallography. The approach which has met with the greatest amount of success thus far has been X-ray crystallography, however this technique is limited by the restricted availability of well diffracting membrane protein crystals. Solution NMR has been widely applied to the study of membrane proteins with a single membrane-spanning domain. The major advantage of solution NMR has been that it utilizes readily available equipment and does not require the generation of protein crystals. The technique which facilitates the study of membrane proteins in the most native-like environment is solid state NMR. Using this approach it is possible to study proteins in bilayer environments.

1.2.1 X-ray Crystallography

Most of the 3D structures of membrane proteins that are available have been determined using

X-ray crystallography. This technique is well suited to membrane protein structure determination as it is not subject to limitations in the size of molecule that can be studied, unlike solution NMR techniques. However, X-ray crystallography has been limited in its application to membrane protein structure determination by the 'crystallization bottleneck', the limited availability of high quality 3D crystals of membrane proteins. In order to study a protein's structure by X-ray crystallography it is first necessary to generate a well diffracting 3D crystal of the protein. For membrane proteins this is not a trivial task since it requires the removal of the protein from its native bilayer environment, followed by reconstitution into an environment amenable to crystallization without perturbation of the proteins' structure which can often lead to denaturation, aggregation or even degradation of the protein. It is the nature of membrane proteins that makes them difficult to crystallize. The hydrophobic surfaces and anisotropic orientation of membrane proteins are serious obstacles to producing well-ordered 3D crystals suitable for X-ray analysis (Landau and Rosenbusch, 1996).

Although crystallographic methods provide atomic resolution structures of protein, they only reveal the static parts of a structure, and do not provide any information about the dynamics of a protein. Depending on the nature of the protein being studied, this information can be important for gaining insights into a protein's function. Concern has also been expressed that the structures solved for membrane proteins in their crystalline forms may not accurately represent the structures of these proteins in biological membranes because the physical environment in a protein crystal is quite different from the environment within a biological membrane.

1.2.2 Solid State NMR

Solid state NMR studies have been used to determine the 3D structures of membrane peptides and proteins, as well as their orientation when associated with lipid bilayers. One advantage of employing solid state NMR techniques to the study of membrane peptide and protein structure is that studies can be conducted in lipid bilayers whose properties are most similar to those of the native biological membranes in which these peptides and proteins reside. Another advantage of this technique is that large proteins and protein complexes that are not amenable to study using solution NMR approaches due to their size, can be studied using solid state NMR. There are currently two major approaches to solid state NMR studies of membrane proteins: studies using oriented samples and studies using magic angle spinning (MAS) (Griffin, 1998).

1.2.2.1 Oriented Samples

Mechanically Oriented Samples

Mechanically oriented samples are prepared by sandwiching preparations of membrane peptides and proteins between thin glass plates. Studies of these types of samples allows for the collection

of orientation dependent information about a peptide or protein because the orientation of the samples is known (Sanders II and Landis, 1994; Chekmenev *et al.*, 2005). For example, the orientation of the membrane active peptide magainin was determined using ^{15}N solid state NMR in oriented bilayer samples. It was found that the helix axis of the peptide was oriented parallel to the bilayer surface at peptide concentrations ranging from 0.8 - 3 mole % (Bechinger, 1997). The use of this technique can be advantageous in that samples can be prepared with higher peptide:lipid ratios than can be achieved using bicellar systems.

Magnetically Oriented Samples

Bicelles are magnetically orientable mixtures of phospholipids that can be used as membrane-mimetic environments for oriented sample NMR studies of membrane-associating peptides and proteins. They are formed by mixtures of certain phospholipids and detergents and can be described as "bilayered discoidal mixed micelles" that are readily oriented by a strong magnetic field (Sanders II and Landis, 1995). Although the exact structure of bicelles are not known, they are believed to have a disk-like shape with a planar surface that has properties similar to those of biological membranes, making them attractive model membranes for structural studies of membrane peptides and proteins (Gaemers and Bax, 2001). They can be described as a class of model membrane systems that lie at the interface between vesicles and classical micellar systems. These membrane-mimetic environments are well suited to NMR studies of membrane protein structure as the detergents used to prepare bicelles do not generally denature water-soluble or transmembrane proteins (Sanders II and Landis, 1994). A variety of integral and peripheral membrane proteins have been reconstituted into bicelles and studied using solid state NMR techniques (Vold, 1997; Losonczi *et al.*, 2000; Whiles *et al.*, 2001; Andersson and Mäler, 2002; Glover, 2002; De Angelis *et al.*, 2004; Ellena *et al.*, 2004; Marcotte *et al.*, 2004). The suitability of bicelles as mimics of biological membranes was demonstrated with the successful incorporation of the integral membrane enzyme, diacylglycerol kinase, into bicelles in a fully functional and stable state (Sanders II and Landis, 1995). For some peptides and proteins it may even be possible to collect both solution and solid state NMR data on the same samples by varying the lipid:detergent ratios in the samples to convert an oriented bicellar solution to an isotropic one.

However, not all membrane peptides and proteins are amenable to study using this technique. Some proteins can alter or disrupt the magnetic orientation of the bicelles, whereas other proteins can interact with the lipids and induce phase changes or phase separation, promoting bicelle fusion or alterations in bicelle morphology (Sanders II and Landis, 1995). All of these interactions would have the potential to interfere with the collection of good quality

NMR spectra of these proteins in bicellar systems. Another challenge of working with bicellar systems is that the optimal sample preparation conditions vary from protein to protein and must be worked out independently for each peptide or protein to be studied.

1.2.2.2 Magic Angle Spinning

This solid state NMR technique which involves fast spinning of the sample at 54.7° (hence the name, 'magic angle spinning'), is used to collect highly resolved NMR spectra (Bechinger *et al.*, 2004). It is generally used on unoriented samples of membrane peptide and proteins reconstituted into multilamellar lipid dispersions (Bloom, 1995; Bouchard *et al.*, 1995; Davis *et al.*, 1995; Farés *et al.*, 2002; Lemaitre *et al.*, 2004).

1.2.3 Solution NMR

With respect to membrane protein structure determination, solution NMR techniques have been most successfully applied to the 3D structure determination of membrane peptides and membrane proteins with a *single* membrane-spanning domain. The reason that this technique has been primarily limited to the study of proteins with a single transmembrane domain is the requirement for isotropic reorientation of the protein molecules in solution on the NMR timescale (microseconds). If a molecule is too large it will not tumble rapidly enough to facilitate study using solution NMR techniques. In the case of membrane proteins, the effective size of the protein in solution is often much larger than the actual size of the protein due to the presence of the detergent molecules required for protein solubilization. This results in an inherent limitation to the size of membrane proteins whose structure can be studied using solution NMR techniques.

One of the major advantages of using solution NMR to study membrane protein structure is that samples can be studied in native-like environments where a wide variety of conditions (temperature, pH, ionic strength, buffers and detergents) can be used. This allows for the investigation of many different aspects of protein structure, such as studies under native and denaturing conditions, or studies of intermolecular interactions and their dependence on external factors such as ionic strength and pH. One of the major challenges faced in solution NMR studies of membrane proteins is how to optimize sample stability while maximizing sample concentration, since the tendency of a protein to aggregate generally increases with increasing protein concentration. Solution NMR can also be used to study protein dynamics in the context of protein structure. These types of studies can provide useful information about the processes underlying molecular recognition, and can help to provide insights into the relationships between a protein's structure and its function. NMR can be used to characterize the internal dynamics of proteins in a direct, quantitative or semi-quantitative manner depending on the nature and frequency of the process being studied (Wüthrich, 1986).

1.2.3.1 Membrane Mimetic Environments

Membrane peptides and proteins cannot be studied in their native membrane environment using solution NMR techniques due to their anisotropic orientation within the bilayers which would result in spectra with very broad signals that would be impossible to analyze (Milon *et al.*, 1990). So one of the major challenges in solution NMR studies of membrane proteins is to reconstitute the protein in a suitable membrane-mimetic environment for study.

Organic Solvents

Solution NMR studies of membrane peptides and proteins have been conducted in different types of membrane-mimetic environments. The most simple of these are mixtures of organic solvents such as methanol, trifluoroethanol (TFE) and hexafluoro-2-propanol (Bazzo *et al.*, 1988; Jarvis *et al.*, 1995; Fregeau Gallagher *et al.*, 1997; Miskolzie and Kotovych, 2003; Booth *et al.*, 2004; Zubkov *et al.*, 2004). An advantage of using solvent mixtures for NMR studies is that the line broadening observed in studies using detergent micelles can be avoided (Bazzo *et al.*, 1988). These systems have been successfully applied to the study of membrane peptides, however their application to the study of membrane proteins raises concerns about the potential for structural artifacts given the significant differences in the properties of an organic solvent mixture versus a biological membrane (Howell *et al.*, 2005).

Detergent Micelles

Detergent micelles have been routinely used as mimics of biological membranes in solution NMR studies of membrane peptides and proteins (Brown and Wüthrich, 1981; Graham *et al.*, 1992; Franklin *et al.*, 1994; Papavoine *et al.*, 1994; Tessmer and Kallick, 1997; Yan *et al.*, 1999; Schibli *et al.*, 2001; Raussens *et al.*, 2003; Howell *et al.*, 2005; Roosild *et al.*, 2005; Zamoon *et al.*, 2005). Preparation of NMR samples involves the careful selection of a detergent for protein solubilization, followed by optimization of the detergent concentration and detergent-to-peptide ratio, as well as conditions of temperature, pH and ionic strength to maximize sample stability (Fernández and Wüthrich, 2003). Although many different lipids can be used for solution NMR studies of membrane proteins, the most reliable for maintaining native functional proteins and yielding well-resolved solution NMR spectra are the detergents sodiumdodecylsulphate (SDS), and dodecylphosphocholine (DPC) (Howell *et al.*, 2005). The packing of phospholipids and detergents in micelles is remarkably different from the packing of phospholipids in biological membranes. There has been some concern that the conformation of a micelle-bound peptide may differ from the bilayer-bound conformation of the peptide. There have been a few documented examples where micelles have been shown to influence the structure of membrane-binding peptides (Milon *et al.*, 1990; Chou, 2002). Since detergent micelles do not provide a bilayer

environment, it is possible that they may not be suitable environments for the reconstitution of some integral membrane proteins in their native conformations (Sanders II and Landis, 1995).

1.2.3.2 Isotopic Labelling of Proteins for NMR Studies

The isotopic labelling of polypeptides is routinely used in solution NMR studies of proteins because it facilitates the application of multidimensional heteronuclear NMR techniques to resolve the complicated 1D ^{1}H NMR spectra of proteins into two, three and four dimensions (Marion *et al.*, 1989a; Marion *et al.*, 1989b; Cavanagh and Rance, 1992; Sklenár *et al.*, 1993; Kainosho, 1997; Rovnyak *et al.*, 2004). These techniques are particularly useful for the study of membrane proteins, since the chemical shift dispersion of the residues in membrane-spanning regions, in particular those in α-helices, is generally poor (Fernández and Wüthrich, 2003). Structural studies of membrane proteins are also complicated by the broad line widths that are often observed for detergent solublized proteins in solution.

1.2.4 Other Techniques

Advances have been made in the use of electron crystallography and atomic force microscopy to study fully functional membrane proteins in their native membrane environments.

1.2.4.1 Cryo-electron Microscopy

An alternative method of studying membrane protein structure is to reconstitute proteins into two-dimensional (2D) crystals in the presence of lipids where the proteins can then be studied using cryo-electron microscopy (EM). This lipid environment is very much like the native environment of the protein, and both the structure and biological activity of the protein can be restored within this environment (Hasler *et al.*, 1998). The 3D structure of a membrane protein reconstituted into 2D crystals can be determined at close to atomic resolution using cryo-EM (Werten *et al.*, 2002).

1.2.4.2 Atomic Force Microscopy

Atomic force microscopy (AFM) can be used to study biological membranes in aqueous solutions. It is possible to image the surfaces of individual protein molecules at sub-nanometer resolution, and to even monitor the movement of single polypeptide loops (Werten *et al.*, 2002). This technique is useful for sampling the conformational states of a protein.

1.3 Lipid Polymorphism

Biological lipids exhibit polymorphic behaviour, the ability to exist in different structural arrangements known as 'phases', depending on conditions of hydration, temperature and lipid composition. The lamellar phase is perhaps the most familiar phase to the biochemist, as biological membranes are predominantly arranged in lamellae (more commonly referred to as bilayers). However, membrane lipids are also capable of forming a fascinating array of different

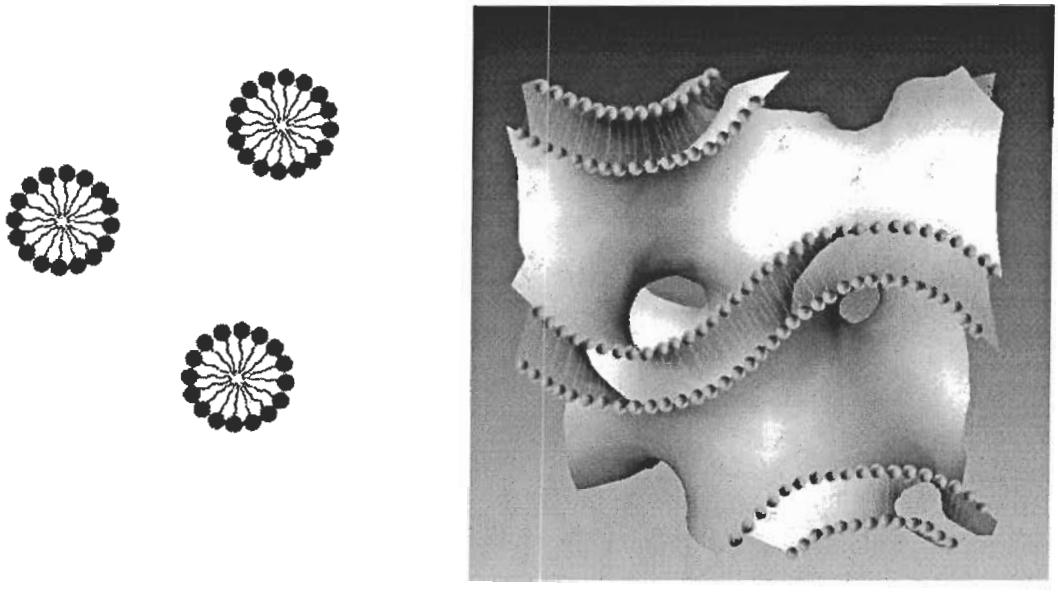
non-lamellar structures to an extent that is not seen in any other class of chemical compounds (Luzzati, 1997). It was observed that in their purified forms, some of the lipid components of biological membranes would spontaneously form non-lamellar phases (Luzzati and Husson, 1962). These non-lamellar phase forming lipids have been shown to play important functional roles in biological membranes where they can modulate the physical properties of membranes or engage in specific interactions with membrane-associated proteins (Dan and Safran, 1998; Epand, 1998). Conversely, it has been shown that peptides and proteins can influence the tendency of lipids to form non-lamellar phases. These types of protein (or peptide) interactions with lipids have effects on membrane stability and consequently on processes such as the formation of pores in membranes, or the fusion of membranes.

The non-lamellar phases formed by membrane lipids have different physical properties than those of lamellar phases. The most striking of these is that the non-lamellar phases have a curved morphology, whereas the lamellar phase has a flat planar morphology. The non-lamellar phases that can be formed by membrane lipids are the micellar, hexagonal and cubic phases which are shown in Figure 1.4 and Figure 1.5. As one might expect, the packing arrangements of the lipids in these phases would be quite different from those in a lamellar phase. So when lipids which normally form highly curved phases are packed in a planar environment there will be some 'strain' in the system as this is not the optimal structural arrangement for the lipids. This strain has been termed 'curvature stain', and it relates to the tendency of the lipids to "want" to adopt curved structures as opposed to the planar structure of the bilayer. It has been found that this property of biological membranes can modulate the activity of membrane-bound proteins (Epand, 1998).

The propensity of lipids to form different phases can be described pictorially using the idea of the 'molecular shape' of the lipids. Lipids can be described as having one of three general shapes: a cone, an inverted cone or a cylinder. The shapes of biological lipids and the phases which they preferentially form are illlustrated in Figure 1.4. Cone-shaped lipids have relatively small headgroups (in comparison to the cross-sectional area of their acyl chains), and they preferentially organize into inverted phases such as reversed hexagonal or cubic phases. Inverted cone-shaped lipids have relatively large headgroups (in comparison to the cross-sectional area of their acyl chains), and they prefer to organize into micelles or normal hexagonal phases. Cylindrically-shaped lipids have headgroups and acyl chains with approximately equal cross-sectional areas, and these lipids preferentially organize into bilayers. In reality, the phase behaviour of lipid systems is quite complicated and cannot be accurately represented solely on the

Figure 1 from Lindblom & Rilfors (1989) *Biochimica et Biophysica Acta* **988**:221-256.

Figure 1.4. Anisotropic Phases That Can Be Formed By Membrane Lipids



Micellar Phase (M)

Gyroid Type Cubic Phase (G)

Reprinted from *Origins of Life and Evolution of the Biosphere*, Vol. 34, 2004, Page 127, " Lipid aggregates inducing symmetry breaking in prebiotic polymerisations", Stefano Piotto, Figure 1, Copyright 2004 Kluwer Academic Publishers, with kind permission of Springer Science and Business Media.

Figure 1.5. Isotropic Phases That Can Be Formed By Membrane Lipids

basis of the molecular shape of the lipids, however molecular shape can be used to illustrate the principles behind lipid organization in the different phases.

1.3.1 Lipid Cubic Phases

Cubic phases are highly viscous, transparent, glass-like materials formed by mixtures of lipids and water under a variety of conditions ranging in temperature from 0 - 150°C, and in water content from 0 % (wt/wt) to excess water (Lindblom and Rilfors, 1989). These phases can be formed by biological lipids such as monoglycerides and phospholipids, as well as by non-biologically derived amphiphiles such as anionic and cationic soaps, and nonionic and zwitterionic surfactants (Lindblom and Rilfors, 1989; Hochkoeppler *et al.*, 1995). The organization of the lipid 'aggregates' which form these phases can be described by a 3D lattice having cubic symmetry which gave rise to the name 'cubic phase' (Lindblom and Rilfors, 1989). One of the unique features of cubic phases is that they are optically isotropic, in contrast to the lamellar and hexagonal phases which are optically birefringent. This property can make cubic phases difficult to discover experimentally, so it is often necessary to employ several methods for the unambiguous identification of these phases (Lindblom and Rilfors, 1989).

1.3.1.1 History

Cubic phases formed by membrane lipids and water were first described in the late 1950s. In 1958 Brokaw described a 'gel state' formed by monomyristin:water and Luzatti had described a 'cubic state' (Lutton, 1965). Initial characterizations of the phase behaviour of lipid:water systems were made using qualitative descriptions of sample viscosity and sample birefringence as determined through observations of the samples through crossed polarizers. Using these techniques, the 'viscous isotropic' phases formed in monoglyceride:water systems were first described and systematically studied (Lutton, 1965). These 'viscous isotropic' phases were believed to have a considerable amount of structure because of their stiff consistency, and the appearance of strong low-angle diffraction lines in X-ray diffraction studies (Lutton, 1965).

1.3.1.2 Characterization

Visual Inspection

Cubic phases can be identified by visual inspection when working with well characterized systems. When a cubic phase is formed the sample becomes completely transparent, very viscous and optically isotropic (Ericsson *et al.*, 1983). The homogeneity of the cubic phase can be confirmed using centrifugation to verify that no phase separation is occurring in the sample (Ericsson *et al.*, 1983).

Electron Microscopy

Cubic phases have been studied by freeze fracture electron microscopy, a technique that has been

applied to the study of many liquid crystalline phases (Landau and Rosenbusch, 1996). Electron micrographs of different cubic phases showed different types of arrangements of lipidic particles. The shape and size of the lipidic particles, as well as the spatial arrangement of these particles, was dependent on the particular lipid:water mixture that was studied (Lindblom and Rilfors, 1989). It was found that the closed aggregate and bicontinuous cubic phases could not be distinguished from one another based on the appearance, and arrangement of the lipidic particles in the electron micrographs (Rilfors *et al.*, 1986).

X-ray Diffraction

X-ray diffraction has been used since the 1960s to determine the structures of a large number of cubic phase-forming amphiphile systems (Lindblom and Rilfors, 1989). One of the major challenges faced in X-ray diffraction studies of cubic phases is the unambiguous determination of the cubic phase structure. Difficulties arise when an insufficient number of reflections are observed, making it impossible to obtain a unique identification for the structure. Also, since this technique detects only those structures with long-range order, it is not possible to rule out sample heterogeneity based solely on the observed X-ray diffraction patterns. For these reasons X-ray diffraction data is often complemented with other methods, such as NMR, in order to facilitate more accurate cubic phase structure determinations.

A relatively recent development in X-ray diffraction methodology has facilitated the study of the structure, dynamics and mechanisms of phase transitions in lipid:water systems. The technique is called time-resolved X-ray diffraction and it involves the use of synchrotron light and a 2D live time X-ray imaging device to study transitions involving lamellar, cubic, reversed hexagonal and fluid isotropic phases (Winter and Köhling, 2004).

NMR Spectroscopy

The solution NMR spectra of bicontinuous cubic phases formed by membrane lipids have narrow well resolved peaks. This is because the molecular motions within these phases are isotropic and consequently average out all nonscalar interactions. For these reasons the NMR spectra of cubic phases can be very similar to those of micellar solutions. The presence of narrow well resolved peaks in the NMR spectra of cubic phases also makes them amenable to study using pulsed field gradient NMR techniques. These types of NMR experiments facilitate the direct measurement of lipid and water translational diffusion coefficients which can be used to differentiate between the two fundamentally different types of cubic phases, the closed aggregate and the bicontinuous. Solid-state ^{2}H and ^{31}P NMR have been used to study the phase behaviour of cubic phase-forming lipid:water systems (Lindblom and Rilfors, 1989).

1.3.1.3 Structure of Cubic Phases

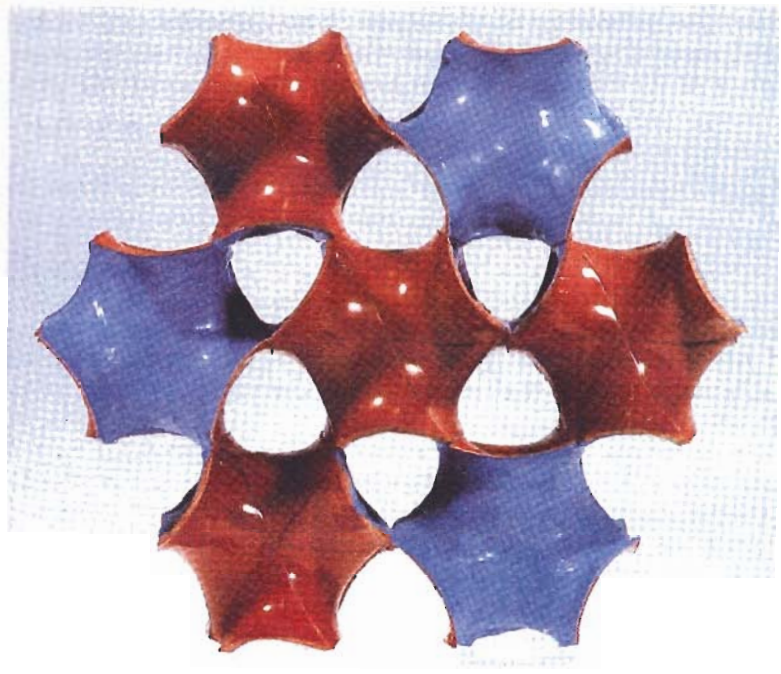
The structures of the cubic phases formed by mixtures of lipids and water were elucidated using data provided from both X-ray and NMR studies. Comparison of lipid and water translational diffusion coefficients made it possible to distinguish between two fundamentally different groups of cubic phases based on the arrangement of the lipid aggregate units which form the phases (Rilfors *et al.*, 1986). The first group are the *closed aggregate* cubic phases which are composed of ordered 3D arrays of lipid micelles in water. In these phases water molecules can diffuse freely in three dimensions, whereas the diffusion of lipid molecules is restricted. The second group are the *bicontinuous* cubic phases whose structure is based on a continuous 3D network of bilayers interleaved with a system of aqueous channels (Landau and Luigi Luisi, 1993). A computer generated model of the gyroid type bicontinuous cubic phase showing the local 'bilayer' structure of these phases is shown in Figure 1.5 and a 3D model of the same cubic phase is shown in Figure 1.6 where opposing surfaces of the 'bilayers' are shown in red and blue. In these phases both water and lipid molecules can diffuse freely in three dimensions. These two groups of cubic phases have many members whose structures have the same general organization (bicontinuous or closed aggregate), but whose lipid aggregate units are arranged in different geometries. The classification of cubic phases within these groups on the basis of their specific geometries is still the subject of debate and will not be discussed further (Lindblom and Rilfors, 1989).

Closed Aggregate

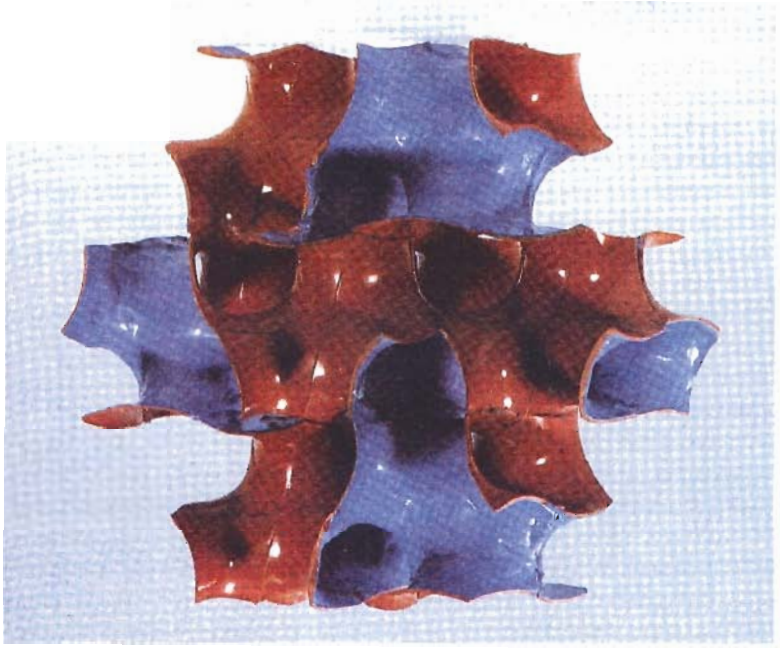
Closed aggregate cubic phases are also called micellar-type cubic phases. They are composed of lipids packed in micelles arranged in a cubic lattice that is interpenetrated by a freely communicating system of aqueous channels (Landau and Rosenbusch, 1996). Lateral diffusion in these types of cubic phases is markedly hindered (Briggs *et al.*, 1996). The cubic phase is built up of short rod-like micelles with an axial ratio of approximately two in a 3D cubic lattice arrangement in water (Lindblom and Rilfors, 1989).

Bicontinuous

In bicontinuous cubic phases lipid and water molecules can diffuse over macroscopic distances without polar groups passing through hydrocarbon regions, and without hydrocarbon chains passing through water regions. The 'bilayer' structure of bicontinuous cubic phases was confirmed by NMR diffusion studies. NMR was used to measure lipid lateral diffusion coefficients in both the lamellar and cubic phases formed by dioleoylphosphatidylethanolamine: dioleoylphosphatidylcholine:water or monoolein:water. In both cases the local lipid diffusion coefficients measured in the cubic phase matched the lipid diffusion coefficients measured in the lamellar phase for the same lipid:water system (Lindblom and Rilfors, 1989).



Top View



Side View

Reprinted from *Zeitschrift für Kristallographie*, Vol. 168, Hyde *et al.*, Pages 213-219, Copyright 1984, with permission from Oldenbourg Wissenschaftsverlag GmbH.

Figure 1.6. Top and Side Views of a Scaled Model of the Gyroid Type Monoolein:Water Cubic Phase

The structure of bicontinuous cubic phases can be described by periodic minimal surfaces. A periodic minimal surface, often referred to as an infinite periodic minimal surface, is a type of mathematical surface defined as having at each point a mean curvature of zero (Lindblom and Rilfors, 1989). Such surfaces can be infinitely extended without self-intersections. The simplest example of such a surface would be the plane formed by lamellar phases which could be viewed mathematically as a volume divided into two physically continuous subvolumes (water and lipid bilayers) without self-intersections. An infinite periodic minimal surface of cubic symmetry continuously extended throughout space can be used to describe the midplane at the center of the 'bilayers' composing the bicontinuous cubic phase (Larsson and Lindblom, 1982). However, in order to accurately represent the 3D structure of the cubic phase, the thickness of the 'bilayer' must be built up by adding the length of the lipid acyl chains to both sides of the minimal surface. For this reason cubic phases built up of 'bilayers' should *not* be referred to as having a zero mean curvature interface (Lindblom and Rilfors, 1989).

1.3.1.4 Cubic Phase-Forming Lipids

There are a variety of lipids that are capable of forming cubic phases under certain conditions of composition and temperature. These include, but are not limited to, monoglycerides, galactolipids, lysophosphatidylcholines (lysoPCs), glycerolipids and total polar lipid extracts from bacterial and plant membranes (Larsson and Lindblom, 1982). A variety of binary lipid mixtures can form cubic phases such as 1-monolein (MO) and water, 1-lauroyl-lysophosphatidylcholine and water, 1-myristoyllysophosphatidylcholine and water, 1-palmitoyl-lysophosphatidylcholine and water, 1-stearoyllysophosphatidylcholine and water, 1-oleoyl-lysophosphatidylcholine and water, 1-linoleoyllysophosphatidylcholine and water, dioleoyl-monoglucosyldiacylglycerol and water, and dioleoylphosphatidylcholine and water (Lindblom and Rilfors, 1989). Ternary mixtures of MO:dioleoylphosphatidylcholine:water, egg phosphatidylcholine:sodium cholate:water, monogalactosyldiacylglycerol:galactosyl-diacylglycerol:water and dioleoylphosphatidylcholine:dioleoylphosphatidylethanolamine:water can also form cubic phases (Lindblom and Rilfors, 1989).

1.3.1.5 Biological Implications

Structures of Biological Membranes

The observation of structures in biological membranes with 3D periodic order reminiscent of cubic phases has sparked much interest in the potential biological roles of cubic phases since these structures efficiently partition 3D space into two unconnected compartments (Landh, 1995). This ability of cubic phases to partition 3D space has been proposed to have had a role in the origins of life, in particular in the origins of biomolecular chirality (i.e. that proteins,

polysaccharides and nucleic acids are made up of monomer units having uniform chirality) (Piotto, 2004). It has been found that total lipid extracts from some biological membranes containing the common membrane constituents phosphatidylcholine (PC) and phosphatidylethanolamine (PE) can form cubic phases (Luzzati, 1997). Based on this information it is believed that some membranes, in particular those forming networks within organelles, may form structures resembling cubic phases.

In plants, electron microscopy revealed that the 3D membrane system that transforms into the thylakoid membranes of chloroplasts had a structure resembling that of bicontinuous cubic phases (Larsson *et al.*, 1980). Similarly, some of the 3D networks of tubules found in the endoplasmic reticulum have structural resemblances to the bicontinuous cubic phases (Luzzati, 1997). Transmission electron microscopy (TEM) was used to prepare 3D reconstructions of the cristae (inner membrane infoldings) of the mitochondria of the amoeba, *Chaos carolinensis* (Deng *et al.*, 1999). These membranes systems are highly curved periodic structures that can be described by mathematical models analogous to the ones used to describe bicontinuous cubic phases. However, the unit cell dimensions of cubic membrane structures are usually larger (> 100 nm) than the spacings observed in lipid systems (which are typically < 20 nm).

A striking feature of the lipid molecules found in archaebacteria is their unusual chemical structure. Archaebacterial lipids have branched chains that are ether-linked to a substituted glycerol, in contrast to the ester-linked unbranched chains of most other organisms. An extreme example of the unusual lipids that can be found in archaebacterial membranes are the bipolar isoprenyl ether lipids. These lipids are essentially dimers of lipid molecules where the ends of two acyl chains are fused together to form one long bilayer-spanning acyl chain (with a length of C_{40}, instead of C_{20}), that has ether linked glycerol head groups at each end (Gulik *et al.*, 1985). Lipids of this nature that have been isolated from the extreme thermoacidophilic archaebacterium *Sulfolobus solfataricus* exist in the bicontinuous cubic phase under conditions that are very similar to the physiological conditions of the organism (in excess water above 80°C) (Luzzati, 1997). Evidence from X-ray scattering studies of these lipid extracts, and of the rigid periodically ordered protein scaffold which supports the plasma membrane *in vivo*, suggest that the plasma membrane of *S. solfataricus* may exist in the cubic phase under physiological conditions.

Membrane Fusion

The transitions between lamellar, cubic and hexagonal phases involve making substantial changes to the topology of these phases and are likely to occur via similar mechanisms to those involved in membrane fusion. For these reasons it has been speculated that the intermediates involved in membrane fusion and lamellar → non-lamellar phase transitions are similar (Siegel, 1999).

Electron micrographs of cubic phases revealed arrays of lipidic particles whose size, shape and distribution were reminiscent of those seen in various biological membranes, and that had been associated with processes such as membrane fusion (Verkleij, 1984). It was also found that the fusion of vesicles, liposomes or membranes could be induced through the addition of monoglycerides or lysoPCs, both of which are capable of forming cubic phases (Rilfors *et al.*, 1986). Studies of liposome fusion induced by phospholipase C (through the generation of fatty acid and diacylglycerol during lipid degradation), revealed the presence of lipid patches whose local structure was similar to those of bicontinuous and closed aggregate cubic phases (Nieva *et al.*, 1995; Basáñez *et al.*, 1997). It has also been shown that diacylglycerol (DAG) and the fusogenic peptide of the influenza virus, can induce the formation of bicontinuous cubic phases in lipid systems (Luzzati, 1997). For these reasons it is believed that cubic phases may be involved in the process of membrane fusion in biological systems, although the exact molecular mechanisms of membrane fusion are not well understood.

Fat Digestion

The process of fat digestion involves a complex series of events that are mediated through the action of pancreatic lipase, colipase and bile salts (Luzzati, 1997). One of the early steps takes place in the stomach and the lumen of the upper intestine where triglycerides are degraded into monoglycerides and fatty acids. A number of different lipid phases are formed during fat digestion, including the lamellar and cubic phases (Patton and Carey, 1979). It was found that cubic phases can be formed by mixtures of monoacylglycerols and protonated fatty acids under conditions corresponding to those in the intestine (Lindblom and Rilfors, 1989). For these reasons it has been speculated that cubic phases formed by monoglycerides and water may be involved in the process of fat digestion.

Drug Delivery

The use of lipid cubic phases as delivery systems for drugs, in particular lipophilic ones, has been investigated. Aqueous dispersions of cubic phases, also referred to as 'nanocubicles' or 'cubosomes', have been developed for the controlled release of drugs (Engström *et al.*, 1999; Siekmann *et al.*, 2002; Yang *et al.*, 2002; Boyd, 2003; Um *et al.*, 2003). One of the main challenges of administering hydrophobic drugs is to solubilize them in a suitable carrier that can be delivered to the gastrointestinal tract, or intravenously into the blood stream, without precipitation of the drug (Boyd, 2003). Since cubic phases could be used to solubilize a variety of amphiphilic, lipophilic and protein molecules, it was believed that perhaps cubic phases could also be used to 'encapsulate' these molecules for delivery to the intestine where they would be taken up by cells (Um *et al.*, 2003). Cubic phases also have a higher bilayer area to particle

volume ratio than liposomes, so they are able to incorporate and deliver larger amounts of lipophilic and amphiphilic drugs per unit volume (Siekmann *et al.*, 2002). Cubic phases are biodegradable *in vivo* where they are easily solubilized by bile salts to produce mixed micelles (Um *et al.*, 2003). In addition to their use to deliver drugs to the intestine, cubic phases can be used to deliver drugs via subcutaneous or intravenous injection (Siekmann *et al.*, 2002; Boyd, 2003).

1.3.1.6 Studies of Peptides and Proteins in Cubic Phases

Cubic phases containing a variety of soluble and membrane-associated peptides and proteins have been successfully prepared and characterized. These studies will be discussed in detail in Chapter 2. It was found that both soluble and membrane proteins could be incorporated into the cubic phase without perturbing the phase behaviour of the system. The structures of the proteins did not appear to be altered by incorporation into the cubic phase. It was also possible to study the function of proteins that were incorporated into the cubic phase.

1.4 Studies of Peptides as Models of Membrane Proteins

Due to the difficulties encountered in the study of membrane proteins, peptide fragments of these proteins are often studied in lieu of studying the whole intact protein (White, 2003). The structure and lipid interactions of peptide fragments of the membrane binding domains of peripheral membrane proteins and the lipid-binding domains of apolipoproteins, as well as many signal peptides have been studied in micelles and organic solvent mixtures using NMR techniques (Chupin *et al.*, 1995; Jarvis *et al.*, 1995; Yin *et al.*, 1995; Chupin *et al.*, 1996; Martinez *et al.*, 1997; Wienk, 2000; Okon *et al.*, 2001; Okon *et al.*, 2002; Raussens *et al.*, 2003; Booth *et al.*, 2004; Ellena *et al.*, 2004).

Peptides derived from the transmembrane regions of integral membrane proteins including: glycophorin A, IsK channel protein, sarcoplasmic reticulum Ca^{2+}-ATPase, *E. coli* phosphatidylglycerophosphate synthase, influenza A M2 protein, multidrug resistance protein EmrE and the human EGF receptor have been studied in membrane-mimetic environments using NMR techniques (Smith *et al.*, 1994; Morein *et al.*, 1996; Aggeli *et al.*, 1998; Rigby *et al.*, 1998; Soulié *et al.*, 1999; Kovacs *et al.*, 2000; Venkatraman *et al.*, 2002).

Membrane-active peptides such as alamethicin and gramicidin have been studied as models of integral membrane proteins, in particular of those with channel activity (Woolley and Wallace, 1992; Bechinger, 1997). These peptides are of interest as therapeutic agents because of their ability to selectively lyse the membranes of target cells. Studies have also been conducted on synthetic peptide models of the transmembrane regions of polytopic membrane proteins (Zhang

et al., 1995b; Zhang *et al.*, 2001; Killian, 2003; Lemaitre *et al.*, 2004). These studies will be discussed in more detail in Chapter 3

1.5 Thesis Objectives

To investigate the suitability of the cubic phases formed by mixtures of monoolein and water as membrane-mimetic environments for the solution NMR study of incorporated membrane proteins, through the study of model transmembrane and membrane surface-binding peptides.

Chapter 2: Preliminary Investigations of Lipid Cubic Phases Suitable for Solution NMR Studies of Membrane Peptides and Proteins

2.1 Objectives

To identify a lipid, or a mixture of lipids, that form cubic phases at room temperature and would be suitable for solution NMR studies of incorporated peptides and proteins.

2.2 Introduction

Although there are many lipids that are capable of forming cubic phases, they often do so only over a very narrow range of temperature and composition. In such cases, the incorporation of additional components would likely have a significant effect on the phase behaviour of the system. There are however some examples of lipid:water mixtures that form cubic phases over a relatively broad range of temperatures and composition. Cubic phases of this type would be more likely to be able to incorporate water soluble, amphiphilic or hydrophobic molecules without significant perturbation of the phase behaviour of the system. It has in fact been shown that some types of cubic phases can incorporate significant amounts of other lipids, detergents and even proteins, without significantly altering the phase behaviour of the system.

Studies have been conducted in which both soluble and membrane proteins have been successfully incorporated into, and in some cases characterized in, lipid cubic phases. Invaluable information necessary for the identification of a cubic phase suitable for use as a membrane-mimetic environment for NMR studies of membrane peptides and proteins was provided by these studies. In order to gather information about: the type and nature of the peptides and proteins studied, sample preparation, sample stability over time and temperature, and peptide and protein conformation and stability following incorporation into the cubic phase, these studies were carefully reviewed (see Section 2.2.1).

2.2.1 Studies of Proteins in Cubic Phases

A number of studies have been conducted in which proteins have been incorporated into the cubic phases formed by mixtures of lipids and water. Since I was interested in using the most simple system possible, only studies involving binary lipid:water mixtures were reviewed.

2.2.1.1 Phase Behaviour of Protein-Containing Cubic Phases

A-gliadin

In one of the first studies of proteins in cubic phases, the amphiphilic protein A-gliadin was incorporated into cubic phases formed by mixtures of MO and water (Larsson and Lindblom, 1982). The authors were able to solubilize a substantial amount of protein, about 10% (wt/wt), in the MO cubic phase. Samples were prepared by mixing solutions of gliadin and MO in a 70:30 ethanol:water mixture (vol/vol) followed by evaporation of the ethanol from the sample. Their results indicated that it was very likely that membrane proteins could be solubilized into the cubic phases formed by MO and water.

Lysozyme and Other Soluble Proteins

In another study a variety of soluble proteins ranging in size from 14 - 150 kDa were incorporated into MO:water cubic phases (Ericsson *et al.*, 1983). The authors were interested in looking at the ability of MO cubic phases to form in the presence of proteins. The proteins: lysozyme (14 kDa), α-lactalbumin (14 kDa), soybean trypsin inhibitor (20 kDa), myoglobin (17 kDa), pepsin (35 kDa), bovine serum albumin (67 kDa), conalbumin (77 kDa) and glucose oxidase (150 kDa), were studied in cubic phases composed of 40 % MO:18 % protein:42 % water (wt/wt). It was assumed that the proteins would be located in the aqueous channels of the cubic phases and not in the lipid 'bilayers'. One observation which supported this assumption was the diffusion of coloured proteins such as myoglobin *into* the cubic phase from an outside protein solution, or *out* of the cubic phase into an outside water solution. Using differential scanning calorimetry it was shown that the enthalpy and temperature of thermal denaturation of lysozyme were the same for the protein in MO cubic phases as for the protein in aqueous solution. This observation provided additional evidence that the proteins were located in the aqueous channels of the cubic phase.

Protein-containing cubic phases were prepared by adding an aqueous protein solution prepared in double distilled water, to MO melted at 40°C. The samples were equilibrated at room temperature for a maximum of 24 hours, or until no change was detected using a polarizing microscope. Centrifugation (8000 × g for 48 h) was used to make sure that no phase separation was occurring in the samples (Ericsson *et al.*, 1983). It was found that MO cubic phases containing lysozyme, α-lactalbumin, bovine serum albumin or pepsin reached equilibrium within

24 hours and formed a viscous transparent optically isotropic phase with the same characteristics as MO cubic phases prepared in the absence of protein. For cubic phases containing myoglobin, an equilibration time of 7 days was needed in order to obtain a homogeneous cubic phase at the protein concentration studied. The authors believed that this difference in the speed of equilibration of myoglobin-containing cubic phases could have been a result of differences in the net charges of the proteins studied since lysozyme and pepsin both carry high net charges (pI = 11 and pI = 1, respectively), whereas myoglobin does not (pI = 7). They believed that the electrostatic repulsion between protein molecules might influence the formation of the cubic lattice. It seemed that the proteins with a high net charge had a greater ability to increase the dimensions of the cubic lattice than did proteins without a high net charge, and that this might have an influence on the time required for samples to reach equilibrium.

To test this hypothesis experiments were conducted in which the protein solutions had *not* been desalted prior to preparation of the cubic phases. It was found that the phase behaviour of these samples differed significantly from that of the original samples which had been prepared with desalted protein solutions. The sample containing myoglobin reached equilibrium as a protein-containing cubic phase in contact with excess protein solution, whereas the samples containing lysozyme or pepsin exhibited similar behaviour to the original samples prepared with desalted solutions, except that they reached equilibrium at an extremely slow rate. These results were as expected, since any salts present in the protein solutions would act to screen any electrostatic repulsion which would increase the relative time required for samples to reach equilibrium.

The X-ray diffraction patterns of the protein-containing cubic phases were found to be similar to those of MO:water cubic phases, strongly suggesting that there were no major structural differences between the phases. The dimensions of the cubic lattice were found to be dependent on both the proportion of protein and the proportion of water present in the cubic phase (Ericsson *et al.*, 1983). When lysozyme was present the cubic phases had a raised water content and the lattice dimensions were increased. The ternary phase diagram determined for MO:lysozyme:water at 40°C using a polarizing microscope and X-ray diffraction is shown in Figure 2.2. Unfortunately the full range of lysozyme concentrations could not be studied due to limitations in the amount of protein that could be incorporated into the samples by direct addition of water solubilized protein because solutions with > 45 % (wt/wt) lysozyme are very viscous (Ericsson *et al.*, 1983). It was found that when the lipid concentration of samples exceeded 80 % (wt/wt) the phase equilibria were very slow, so this region of the phase diagram was omitted.

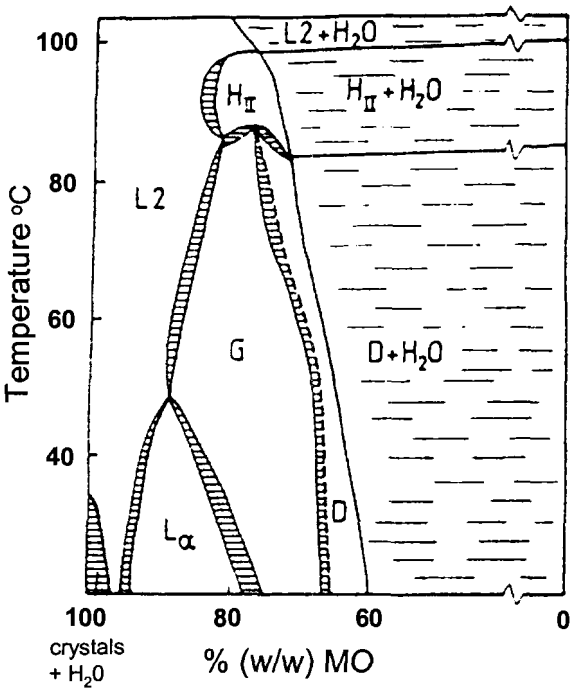


Figure 2.1. Monoolein:Water Phase Diagram

The bicontinuous cubic phases with either a gyroid (G) or diamond (D) geometry are indicated on the phase diagram as 'G' and 'D' respectively. The reversed micellar phase is denoted at 'L_2', the reversed hexagonal phase as 'H_{II}' and the lamellar phase as 'L_α'.

Figure 1 from Ericsson *et al.* (1983) *Biochimica et Biophysica Acta* **729**:23-27.

Figure 2.2. Ternary Phase Diagram of Monoolein:Lysozyme:Water at 40°C

2.2.1.2 Spectroscopic Studies of Proteins in Cubic Phases

A number of different spectroscopic studies had been conducted on soluble proteins and on membrane proteins incorporated into both closed aggregate and bicontinuous cubic phases.

UV/vis and Circular Dichroism Spectroscopy of α-chymotrypsin and Bacteriorhodopsin

The optical transparency of cubic phases has made them amenable for use in spectroscopic studies of proteins. In one investigation the conformation and activity of the hydrophilic enzyme α-chymotrypsin was studied in cubic phases using UV/vis spectroscopy and circular dichroism (CD) spectroscopy (Portmann *et al.*, 1991). Some of the reasons why the authors chose α-chymotrypsin for their studies were that the spectroscopic properties of the enzyme in aqueous solution were well characterized and sensitive to changes in protein conformation. The authors also collected preliminary data on the membrane protein bacteriorhodopsin that had been incorporated into the cubic phase. The cubic phases that were used in these studies were either the closed aggregate cubic phases formed by 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphocholine (lysoPC) and water, or the bicontinuous cubic phases formed by MO and water. The viscoelastic properties of lysoPC:water cubic phases in the presence or absence of protein were studied at 27°C.

LysoPC forms a cubic phase when mixed with water at concentrations between 39 - 45 % (wt/wt). Cubic phase samples of lysoPC were prepared by adding buffer solutions (with or without protein) directly to lipid in 1 mL UV cuvettes. The cuvettes were centrifuged for 1 - 2 days (2900 × g at 22°C) and then they were equilibrated at room temperature on the benchtop for 2 days during which time the UV scattering decreased significantly. The resulting cubic phases were highly stiff transparent gels. Monoolein cubic phases were prepared by slowly adding phosphate buffer (with or without enzyme) to the appropriate amount of melted MO at 40°C. The samples were equilibrated for a few hours at 40°C and then for a few days at room temperature under nitrogen. All of the aqueous solutions used for the preparation of lysoPC cubic phases (with and without protein) contained 18 mM phosphate buffer at pH 8.0. The aqueous solutions used to prepare samples of MO cubic phases contained phosphate buffer at pH 6.0. For the most part, cubic phase samples were prepared directly in spectroscopic cells. This set some limits as to the experiments that could be performed because of limitations in the centrifugal force that commercially available spectroscopic cells could withstand. Circular dichroism measurements were made on cubic phase samples that had been pressed as thin films between quartz plates.

Both α-chymotrypsin and bacteriorhodopsin were easily incorporated into the lysoPC cubic phase. It was possible to measure the UV/vis absorption spectra of both proteins in these cubic phase samples under a variety of different concentrations. The absorption spectra of the proteins in lysoPC cubic phases were found to be essentially the same as for the proteins in aqueous solution over the protein concentration range that was studied. A protein concentration of 30 μM α-chymotrypsin and 14.5 μM bacteriorhodopsin were obtained without compromising the optical transparency of the cubic phases composed of 41 % and 44 % (wt/wt) lysoPC:water respectively. The enzyme activity of α-chymotrypsin in 40 % (wt/wt) lysoPC:water cubic phases was monitored using a water soluble substrate whose absorption spectrum could be measured as a function of time. Measurements were taken at time intervals ranging from 3 to 21 hours following the addition of the substrate to the sample. It was found that the enzyme kinetics in the cubic phase were considerably slower than in aqueous solution.

Circular dichroism studies were used to determine whether or not the conformation of α-chymotrypsin was modified when it was incorporated into the cubic phases formed by lysoPC and water, or MO and water. Since lysoPC is optically active, it was necessary to prepare cubic phases using racemic lysoPC in order to avoid lipid contributions to the CD spectra in the far UV region. The spectra in water and the spectra in the cubic phase were almost identical indicating that the main chain conformation of the protein was not perturbed in this lipid matrix. Spectra could not be collected at less than 200 nm on these samples because of the large absorption band of lysoPC in this region of the spectrum. Instead, CD spectra were collected in this region on α-chymotrypsin in MO cubic phases. These results did not show significant differences between the aqueous and cubic phase samples, indicating that the aromatic side chains of the protein had not undergone significant conformational changes. The samples of proteins in cubic phases were also found to be very stable. In the case of α-chymotrypsin, samples with protein concentrations up to 13 μM remained stable for a least 4 weeks at room temperature. The assessment of sample stability was based on the consistency of spectra collected as a function time. It appeared that the samples would remain stable indefinitely when they were kept in sealed vials. Rheological studies were conducted at 27°C on pure 44 % lysoPC:water (wt/wt) cubic phases and on those containing 9.7 μM α-chymotrypsin. The data that were obtained for the protein-free, and protein-containing, cubic phases were practically identical. This indicated that the incorporation of proteins does not alter the viscoelastic properties of the cubic phase, suggesting that the structure of the cubic phase does not change upon incorporation of protein. The rheological data also indicated that cubic phases behave like solids and not liquids.

Circular Dichroism Studies of Bacteriorhodopsin and Melittin

Studies were conducted on bacteriorhodopsin and melittin in lysoPC:water cubic phases in an attempt to develop a novel membrane-mimetic environment that would facilitate the simultaneous study of a membrane protein's structure and function. The authors had selected an integral membrane protein and a membrane surface-associating peptide for their studies. The integral membrane protein that was selected was bacteriorhodopsin, a 248 amino acid protein with seven transmembrane α-helices. It is found in the halophilic bacterium *Halobacterium halobium* where it functions as a light driven proton pump (Landau and Luigi Luisi, 1993). Melittin was the surface-associating peptide selected for study. It is a 26 amino acid water soluble peptide that is the main component of honey bee venom. This peptide spontaneously binds to membranes where it acts as a stong membrane lytic agent (Landau and Luigi Luisi, 1993).

Cubic phase samples to be used for UV measurements were prepared directly in 1 mL UV cuvettes through the addition of buffer solutions of melittin, or suspensions of bacteriorhodopsin, to lysoPC followed by centrifugation for 1 - 2 days (2900 × g) at 25°C (Landau and Luigi Luisi, 1993). In the case of samples to be used for CD measurements, protein-containing cubic phases were prepared in 1 mL vials and were transferred to quartz plates just prior to data collection. The bacteriorhodopsin suspensions used for cubic phase sample preparation were prepared in 18 mM p-buffer at pH 8.0, and the melittin solutions were prepared in 20 mM Tris-HCl at pH 7.4. In all cases the transformation of the samples to highly stiff and transparent gels was used to indicate that cubic phase formation was complete. Cubic phases were successfully prepared with protein concentrations of up to 1.6×10^{-4} M for bacteriorhodopsin and 1.2×10^{-4} M for melittin. For both bacteriorhodopsin and melittin, cubic phase samples were found to be stable at ambient temperature for many months.

The CD spectra of bacteriorhodopsin in lysoPC cubic phases indicated that the protein had a helical content consistent with the values determined from electron microscopy and electron diffraction studies of the protein in its native purple membranes. These findings suggested that the conformation of the protein was not perturbed by incorporation into the cubic phase. The thermal stability of bacteriorhodopsin was studied over the temperature range in which lysoPC:water cubic phases form (0 - 50°C). Data were collected at 10°C intervals from 5 - 45°C for both heating and cooling of the sample. The spectra were found to be virtually identical at all temperatures, except at 45°C where the cubic phase starts to melt. These results were consistent with those obtained in other studies investigating the thermal stability of bacteriorhodopsin's conformation over a similar temperature range. Circular dichroism spectra were also collected on melittin in lysoPC cubic phases as a function of temperature (at 10°C intervals from 5 - 45°C), for

both heating and cooling of the sample. The conformation of the peptide was found to be temperature independent (except at 45°C where the cubic phase starts to melt) and in good agreement with the helical content determined from the crystal structure melittin.

Visible and CD Spectroscopy of a Photosynthetic Reaction Center

The photosynthetic reaction center from the facultative phototropic bacterium *Chloroflexus aurantiacus* was studied in lysoPC:water cubic phases by visible and CD spectroscopy (Hochkoeppler *et al.*, 1995). The authors were interested in studying the photochemical activity of this integral membrane protein in the cubic phase. Since the kinetics of the photochemical reaction do not depend on diffusionally controlled processes, it was expected that incorporation of the protein into the cubic phase would not affect the kinetics of the reaction. Protein-containing cubic phases were prepared by weighing out the desired amount of lipid directly into spectroscopic cuvettes followed by the addition of an appropriate amount of protein prepared in 50 mM phosphate buffer at pH 8.0. Samples were mixed by centrifugation (2900 × g) back and forth several times for 5 minute intervals, followed by centrifugation overnight and subsequent equilibration under ambient conditions for 48 hours before measurements were made. For CD experiments, the cubic phase samples were pressed between quartz plates just prior to data collection.

Both visible and CD spectra were collected on the *Chloroflexus* photosynthetic reaction center reconstituted into 42% (wt/wt) lysoPC:water cubic phases at a concentration of 1.4 - 4 μM. The results obtained from these studies indicated that the incorporated protein retained its native spectroscopic properties in the lysoPC cubic phase. The photosynthetic reaction center was also incorporated into 35% (wt/wt) MO:water cubic phases. Unfortunately bleaching of one of the absorption bands occurred, perhaps as a result of oleate peroxide formation under the aerobic conditions used for the experiment, so further studies were not conducted using this system. The structural and functional integrity of the protein in lysoPC cubic phases was found to be maintained for at least 5 months for samples stored at room temperature in the dark.

2.2.1.3 Crystallization of Membrane Proteins from Cubic Phases

Bacteriorhodopsin

The applicability of cubic phases for the generation of membrane protein crystals was investigated using bacteriorhodopsin as a model for integral membrane proteins (Landau and Rosenbusch, 1996). It was believed that cubic phases would be well suited to the generation of protein crystals because they had the capacity to provide nucleation sites for the seeding of crystal growth, as well as the ability to feed the growing crystal through the lateral diffusion of protein molecules within the cubic phase 'bilayer'. Studies were conducted using the bicontinuous cubic

phases formed by MO and water or monopalmitolein and water, and the closed aggregate cubic phases formed by lysoPC and water.

Samples 10 - 20 μL in volume were prepared in thin glass capillaries by mixing lipids with aqueous phosphate buffer (1 - 3 M) at pH 5.6 containing protein, detergent (1.2% β-octylglycopyranoside) and precipitant (0.05% methylpentanediol), followed by centrifugation (10,000 × g) for 150 minutes prior to incubation at 20°C in the dark. Within 14 days both the MO:water and the monopalmitolein:water cubic phases had yielded protein crystals. Crystal formation was monitored both spectroscopically and microscopically directly in the glass capillaries. Initially the cubic phases were a homogeneous purple colour, the colour of bacteriorhodopsin, and they had an absorption peak at 550 nm. As crystallization progressed the purple colour of the samples faded and the intensity of the absorption peak at 550 nm decreased, and protein crystals were visible under the microscope. Diffraction data were collected on crystals mounted directly in the glass capillaries with the surrounding cubic phase intact.

The cubic phases which yielded bacteriorhodopsin crystals had a composition of 60 - 70% (wt/wt) MO or monopalmitolein in water. The growth of the crystals did not affect the transparency or the viscoelastic properties of the cubic phase. No fragments of lipid bilayers were incorporated into the crystals suggesting that the cubic phase 'bilayer' recedes during crystal growth. It was found that crystals could only be generated from bicontinuous cubic phases and not from closed aggregate cubic phases, despite extensive crystallization attempts. This finding indicated that the continuity of the diffusion space, in this case the 'bilayer', played a critical role in crystal growth. In unpublished data the authors found that the water soluble enzyme lysozyme could be crystallized from either type of cubic phase. The fact that lysozyme's crystallization was not dependent on the type of cubic phase used was not unexpected since this enzyme resides in the aqueous channels of the cubic phase.

Crystallization was believed to proceed through a partitioning of the membrane protein into the hydrophobic 'bilayer' of the cubic phase which would also act as a sink for detergent monomers, occurence of a nucleation event (perhaps favoured by the large interfacial area between the lipidic and aqueous compartments), followed by subsequent growth of the crystal using protein molecules supplied from locations throughout the cubic phase through lateral diffusion within the 'bilayers'. The crystals of bacteriorhodopsin that were generated diffracted to a 3.7 Å resolution. In subsequent studies by this research group, bacteriorhodopsin crystals diffracting up to a resolution of 1.9 Å were obtained and used for the successful structure determination of bacteriorhodopsin (Pebay-Peyroula *et al.*, 1997; Belrhali *et al.*, 1999).

2.2.2 Candidate Cubic Phases for NMR Studies of Embedded Proteins

A number of criteria were considered in the selection of a lipid cubic phase that would be amenable to solution NMR studies of incorporated peptides and proteins. In addition to the requirement that the cubic phase be one that had been previously well characterized, preferably in studies with incorporated peptides and/or proteins, the lipidic and aqueous components of the cubic phase had to undergo rapid isotropic reorientation. This requirement for rapid isotropic reorientation of components within the cubic phase was necessary to facilitate the study of these phases, and incorporated peptides or proteins, with solution NMR techniques. Other factors that were considered in cubic phase selection were the: commercial availability and cost of the lipid(s), availability of the lipid(s) in deuterated form, complexity of cubic phase preparation (i.e. binary versus ternary lipid mixtures), range of temperature and composition over which the cubic phase exists, and stability of cubic phase samples as a function of time.

There were two lipid mixtures capable of forming cubic phases that had been extensively characterized in the absence and presence of incorporated peptides and proteins. These were the cubic phases formed by lysoPC and water, and MO and water. The cubic phases formed by these lipids, and their potential application as membrane mimetic environments for solution NMR study of incorporated membrane peptides and proteins, are discussed in Section 2.2.2.1 and Section 2.2.2.2.

2.2.2.1 LysoPC:Water Cubic Phases

The ability of the cubic phases formed by lysoPC and water to incorporate peptides and proteins has been well characterized. These cubic phases are able to incorporate significant amounts of peptides and proteins without significant perturbation of the phase behaviour of the system. Solid state ^{2}H, ^{14}N and ^{31}P NMR spectra have been collected on these cubic phases. It was found that the spectra of lysoPC:water cubic phases contained both isotropic (narrow) and not fully averaged anisotropic (broad) signals, even though the cubic phases were optically isotropic when viewed through crossed polarizers (Lindblom and Rilfors, 1989). This information, in conjunction with the data provided from X-ray diffraction studies was used to come up with a model for this type of cubic phase. It is believed that the structure of the cubic phases formed by lysoPC and water are composed of closed rod-like micelles with an approximate axial ratio of 2, surrounded by water (Lindblom and Rilfors, 1989). The restricted diffusion of lipidic components within the closed aggregate structure of the lysoPC:water cubic phases, combined with the lack of rapid isotropic reorientation, would likely make them unsuitable cubic phases for solution NMR studies of incorporated peptides and proteins. Despite this, a considerable amount of useful information about the general properties of cubic phases, methods for the preparation of protein-containing

cubic phases, and peptide and protein stability within cubic phases was provided from the various studies conducted on lysoPC:water cubic phases that could be applied to the preparation and NMR study of protein-containing cubic phases.

2.2.2.2 Monoolein:Water Cubic Phases

The cubic phases formed by MO and water have been extensively studied. Both hydrophobic and water soluble peptides and proteins have been successfully incorporated into these cubic phases without causing significant perturbations of the phase behaviour of the system. The plasticity of MO:water cubic phases has been further demonstrated by the ability of these phases to incorporate significant amounts of other molecules such as dioleoylphosphatidylcholine, glycerol, cholesterol, sodium cholate, tetramethylsilane, oleic acid, monostearin, and 1-oleoyl-lysophosphatidylcholine without perturbation of the cubic phase (Lindblom and Rilfors, 1989). The properties of the cubic phases formed by MO and water are quite remarkable given the simplicity of the system. Monoolein is a monoglyceride whose structure is shown in Figure 2.3. The lipid does not have a head group, and consists of a glycerol backbone and an oleic acid chain at position 1. In addition to its ability to form cubic phases upon mixture with water, MO is used as an additive in foods and is known to be a very effective fusogenic lipid (Lutton, 1965; Hope and Cullis, 1981)

The cubic phases formed by MO and water are of the bicontinuous type. They are composed of a continuous lipid bilayer separating two continuous, but not connected, water channel systems (Larsson and Lindblom, 1982). The structure of MO cubic phases was determined using X-ray diffraction and NMR data. It was found that the lipid lateral diffusion rates measured in the cubic phase compared exactly with those measured in the corresponding lamellar phase providing additional evidence that these cubic phases were composed of bilayer units (Lindblom *et al.*, 1979). The 'bilayer' environment provided by the MO cubic phase has comparable viscoelastic properties to those of biological membranes making them excellent membrane-mimetic environments for the study of membrane peptides and proteins (Landau and Rosenbusch, 1996).

There are two types of bicontinuous cubic phases that can be formed by mixtures of MO and water. These are the gyroid (G), and diamond (D) type cubic phases. In the MO:water phase diagram shown in Figure 2.1, it can be clearly seen that cubic phases exist over a wide range of temperatures and composition (Hyde *et al.*, 1984). It should be noted that only negligible amounts (less than a small fraction of a percent) of monoglycerides dissolve in water, so MO does not form a micellar solution in water but rather exists as a cubic phase in the presence of excess water (Lutton, 1965). The lamellar (Lα) phase forms in mixtures of MO and water at hydration levels that are lower than those required to obtain cubic phases. Under conditions of

extreme dehydration and elevated temperature the system will form a reversed micellar (L_2) phase. The other phase that can be formed by this system is the reversed hexagonal (H_{II}) phase which generally forms under conditions of elevated temperature.

Structure

The structures of the MO cubic phases are formed by infinite lipid bilayers that separate two continuous water networks that never connect. The 3D structure of these cubic phases can be described by an infinite periodic minimal surface located at the center of the 'bilayers' where the methyl groups at the termini of the acyl chains from adjacent lipid monolayers actually meet. The curvature of this periodic minimal surface is such that every lipid head group is at a saddle point so all molecules experience an equivalent packing environment. The tails of the lipids in this type of a structure are less restricted in their motion than the headgroups resulting in a wedge-like molecular shape for MO in the cubic phase (Larsson and Lindblom, 1982). There are two types of bicontinuous cubic phases that can be formed by MO and water depending on the temperature and composition of the samples. One exists at lower water contents, and one at higher water contents, with a narrow range over which both phases coexist in equilibrium with one another. The two types of cubic phases formed by MO and water differ slightly in their geometries and are named the gyroid (G-type), and diamond (D-type) cubic phases after the infinite periodic minimal surfaces that describe them.

The transition between these two cubic phases occurs without any change in the average curvature of the 'bilayers' (Hyde *et al.*, 1984). This would suggest that the enthalpy of transition from the G-type to the D-type cubic phase would be very small. Differential scanning calorimetry was used to measure the enthalpy of this transition and it was found to be less than about 0.01 $kJ \cdot mol^{-1}$ (Hyde *et al.*, 1984). By comparison, the enthalpy of transition for the transition from the cubic to the reversed hexagonal phases formed by MO and water is about 1 $kJ \cdot mol^{-1}$, which is of the same order of magnitude as phase transition enthalpies seen in other lipid systems (Hyde *et al.*, 1984). The transition from the G $\rightarrow$ D type cubic phase (which starts to occur at around 33.5 % (wt/wt) water at 20°C) can be difficult to observe visually because it is only characterized by a reduction in the transparency of the sample within the phase coexistence region (Hyde *et al.*, 1984). However, this transition can also be observed through changes in the water diffusion coefficient measured using NMR. It was found that there is a discontinuity in the water diffusion coefficient, but not in the lipid diffusion coefficient, over the composition range where the transition from the G $\rightarrow$ D type cubic phase occurs (Lindblom and Rilfors, 1989). This observation is consistent with a reorganization of the structure of the cubic phase in which the network of aqueous channels is rearranged without disruption of the lipid bilayer network.

The dimensions of the 'bilayer' in MO cubic phases of varying compositions were determined at 25°C using X-ray diffraction techniques (Chung and Caffrey, 1994). The lipid length in MO:water cubic phases was found to be constant over the composition range studied (66.5 - 76.9 % MO). However, the lattice parameter (the smallest repeating unit of the cubic lattice), was found to vary as a function of sample composition, for example it changed from 125.35 → 147.49 Å when the water content of the sample changed from 23.1 → 33.5 %. As shown in Figure 2.4, the average lipid length in MO:water cubic phases is 17.3 Å and this is assumed to be constant throughout the cubic phase. The head group region varies from 4 - 7 Å in length, leaving an acyl chain region of 10.3 - 13.3 Å in length.

The cross-sectional area of the lipid molecules is larger at the methyl terminus and smaller at the headgroup giving the molecules a dynamically averaged wedge shape as shown in Figure 2.5. The wedge shape becomes more pronounced with decreasing water content so the surfaces become more highly curved and the lattice parameter decreases. With regards to the dimensions of the aqueous pore in MO cubic phases, they are on the order of 50 Å and have been found to vary as a function of electrostatics and water content (Chung and Caffrey, 1994; Giorgione *et al.*, 1998; Aota-Nakano *et al.*, 1999). For example it was found that the size of aqueous channels in MO:water cubic phases containing the anionic lipid 1-palmitoyl-2-oleoyl-3-phosphatidylserine (POPS) were found to vary in size from 39 Å - 62 Å over PS contents of 0 - 10% (Giorgione *et al.*, 1998).

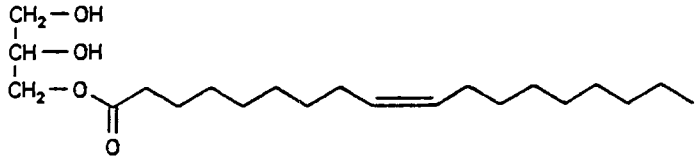
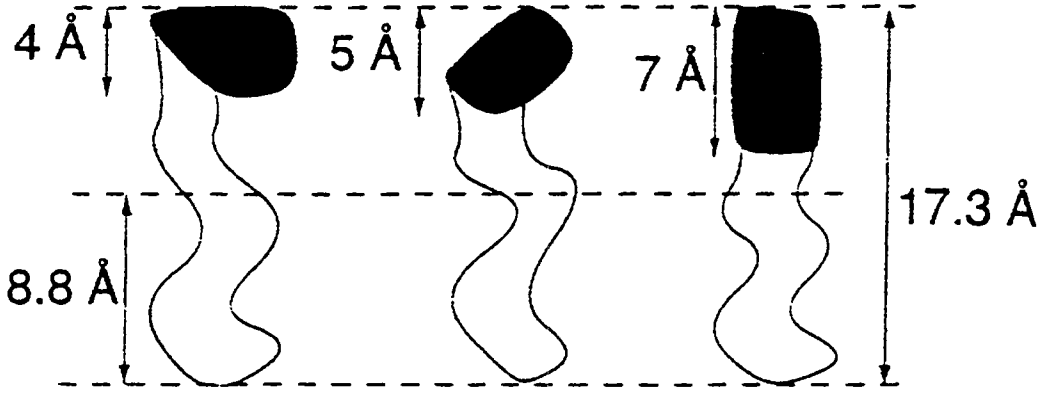
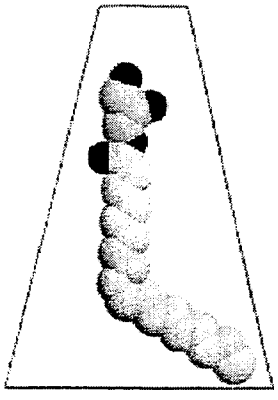


Figure 2.3. Monoolein



Reprinted from *Biophysical Journal*, Vol. 66, Chung & Caffrey, Pages 377-381, Copyright 1994, with permission from the Biophysical Society.

Figure 2.4. Schematic Diagram of the Dimensions of the 'Bilayer' in Monoolein:Water Cubic Phases



Reprinted from *Biophysical Journal*, Vol. 79, Ai & Caffrey, Pages 394-405, Copyright 2000, with permission from the Biophysical Society.

Figure 2.5. Dynamically Averaged Shape of Monoolein Molecules in the Cubic Phase

2.3 Materials and Methods

2.3.1 Selection of a Cubic Phase for NMR Studies

The cubic phases formed by mixtures of MO and water were deemed the most promising for solution NMR studies of incorporated membrane peptides and proteins. The primary reason for this was that the rapid isotropic reorientation experienced by both the aqueous and lipidic components of MO cubic phases make them amenable to study using solution NMR techniques. By analogy, it was believed that an incorporated membrane peptide would also experience rapid isotropic reorientation facilitating its study using solution NMR techniques. Another important factor in the selection of MO cubic phases was the demonstrated ability of these phases to incorporate membrane peptides and proteins, and various hydrophobic or amphipathic molecules to relatively high concentrations without perturbation of the cubic phase. This plasticity of MO cubic phases, along with their extraordinary stability as a function of time, temperature and composition, made them ideal candidate membrane-mimetic environments for the solution NMR study of membrane peptides and proteins.

2.3.1.1 Monoolein

Monoolein was found to be a relatively inexpensive lipid that was readily available in sufficient quantities for the preparation of the sample volumes required for solution NMR studies. Unfortunately MO was not available as a perdeuterated lipid, even by custom synthesis. It would therefore be necessary to use labelled peptides or proteins for any NMR studies conducted on peptides or proteins incorporated into the MO cubic phase.

2.3.2 Preparation of Monoolein:Water Phases

2.3.2.1 Cubic Phase Sample Preparation in 0.6 mL Microcentrifuge Tubes

The first step was to prepare some samples of MO and water of varying compositions and to confirm the identity of the phase(s) that were formed. The information provided in the MO phase diagram shown in Figure 2.1 was used to prepare samples with compositions of 60 %, 65 % and 70 % MO (wt/wt) in water. These samples were expected to form cubic phases of the: D-type with a minor excess of water, D-type and the G-type, respectively at 20°C. The jar containing MO (MW 356.5 g, Sigma Chemical Co., 1-monooleoyl-rac-glycerol (C18:1,[cis]-9) ~99%, Cat. No. M-7765, 1 g) was removed from the -20°C freezer and allowed to equilibrate to room temperature for at least 1 hour before the lipid was weighed out. This was done to prevent the condensation of water on the lipid which would lead to weighing errors. The quantity of MO that was weighed out was the amount needed to prepare a sample with a final volume of approximately 200 μL, assuming that a cubic phase sample with a total mass of 200 mg would have a total volume of 200 μL (Chung and Caffrey, 1994). Monoolein was weighed out on

weighing paper using an analytical balance (Sartorius, Cat. No. BP 210S, readability of 0.1 mg) and carefully transferred to 0.6 mL microcentrifuge tubes. At room temperature MO is composed of relatively large fluffy flakes that are slightly greasy, which made it quite difficult to transfer the required quantity of lipid to the 0.6 mL microcentrifuge tube. An anti-static instrument (Aldrich Chemical Company Inc., Zerostat 3, Cat. No. Z10,881-2) was used to reduce the amount of lipid lost during transfer to the microcentrifuge tubes. After the MO was weighed out the bottle was purged with Ar(g) and the lid was tighten and sealed with laboratory film (American National Can, Parafilm "M") before being returned to the -20°C freezer.

The amount of water to be added to each of the samples was calculated based on the mass of lipid that had been weighed out. The water used for sample preparation was double distilled water that had been filter sterilized using a 0.22 μm syringe filter (Millipore, Millex-GP, Cat. No. SLGPR25LS) into a sterile 50 mL polypropylene conical tube (Becton Dickinson Labware, Falcon Brand, Cat. No. 352070) using sterile technique. The pipettor (Gilson, Pipetman, P200) to be used to measure out the water was calibrated with water equilibrated to ambient temperature just prior to use. This step was undertaken every time samples were prepared to ensure that the correct mass of water would be added to each sample because of significant fluctuations in the ambient temperature of the lab. The desired mass of water was added to each microcentrifuge tube containing solid MO. Attempts were made to vortex mix the samples, unfortunately this was not possible due to the 'solid' nature of the samples resulting from their high lipid/low water content. Sample tubes were wrapped in aluminum foil to protect from light and left to equilibrate at ambient temperature on the benchtop overnight.

The appearance of the samples had changed very little by the following day, so it became apparent that it would be necessary to centrifuge the samples in order to achieve proper mixing of the lipids with water. Samples were centrifuged for 2.5 hours at 10,000 × g (Eppendorf, Centrifuge 5804, FA45-30-11 rotor) at ambient temperature. The sample tubes were removed from the centrifuge and left at ambient temperature overnight on the benchtop wrapped in aluminum foil. Since the appearance of the samples was unchanged the following day, it appeared that it would be necessary to centrifuge the samples for a prolonged period of time to obtain sample homogeneity. An ultracentrifuge had to be used in order for the samples to be centrifuged overnight. Since the samples were too gooey and sticky to transfer to new tubes compatible for use with the ultracentrifuge without incurring substantial sample loss, it was necessary to place the 0.6 mL microcentrifuge tubes (with the lids cut off) inside 1.5 mL thick-walled microcentrifuge tubes (Beckman, Polyallomer Microfuge Tube, Cat. No. 357448) that were compatible for use with the ultracentrifuge. Samples were centrifuged (Beckman, Optima

TLX Ultracentrifuge, TLA-45 rotor) for 16 hours at 10,000 × g and 25°C. Since sample homogeneity had still not been achieved, the samples were heated in a water bath for 4 hours at 54°C, before being equilibrated back to ambient temperature for 2.5 hours. Samples were wrapped in aluminum foil and allowed to come to equilibrium at ambient temperature on the benchtop over the course of 2 weeks.

2.3.2.2 Sample Preparation Directly in 5 mm Solution NMR Tubes

The MO:water cubic phase that was selected for further characterization was the G-type cubic phase because it exists as a homogeneous phase over a broad range of temperature and composition, unlike the D-type cubic phase which co-exists with excess water over most of its existence range. The sample composition that was selected, 70 % MO (wt/wt) in water, was well within the existence range of the G-type cubic phase at the temperatures that would be used for NMR studies. Samples of the lipid phases found on either side of the G-type cubic phase, the lamellar phase (at lower water concentrations), and the D-type cubic phase in excess water (at higher water concentrations), were prepared for NMR characterization and comparison of the results with those obtained for the G-type cubic phase. So NMR samples containing 20 %, 70 % or 85 % MO (wt/wt) in D_2O would be prepared directly in 5 mm solution NMR tubes.

The minimum sample length required for solution NMR studies is ~ 3 cm, which corresponds to a sample volume of ~ 450 μL in 5 mm solution NMR tubes. Samples having an approximate final volume of 450 μL were prepared based on the assumption that a sample with a total mass of 450 mg would have a total volume of 450 μL (Chung and Caffrey, 1994). Quartz NMR tubes (Kontes Glass Company, Cat. No. 897220-0000, 5 mm) were washed three times each with: acetone (Anachemia, Lab Grade, Cat. No. UN-1090, 4L), 95% ethanol (Commercial Alcohols Inc., 4L) and distilled deionized water, and were oven dried overnight at 100°C prior to use. NMR samples were prepared essentially as described in Section 2.3.2.1: the desired amount of MO was weighed out, transferred to a solution NMR tube, and the appropriate volume of deuterated water (Isotec Inc., deuterium oxide, 99.996 atom % D, Cat. No. T82-70003-ML, 0.8 mL) was added using pipettors (Gilson, Pipetman, P200 and P1000) calibrated using H_2O to ensure that no differences in sample composition would result from the slightly different densities of H_2O and D_2O. The deuterated water (D_2O) used in sample preparation had a pH of ~ 7, whereas the distilled deionized water had a pH of ~ 5, as determined using pH paper (EM Science, colorpHast Indicator Strips pH 0-14, Cat. No. 9590). An NMR sample of ~ 2 mg of MO in deuterated chloroform (Aldrich, 99.96 atom % D, Cat. No. 45,328-5 , 1 mL) was prepared for comparison of the ^1H spectrum with the one obtained for MO in the cubic phase.

Adequate mixing of lipid and water in the NMR samples could not be achieved using centrifugation because of a lack of availability of an NMR tube-compatible centrifuge. Instead, the samples were heated at 60°C in a water bath and mechanically mixed using a stainless steel spatula in an attempt to speed up the process of sample equilibration. After mechanical mixing and incubation in the water bath for ~ 5 hours, the samples were removed from the water bath and were equilibrated to room temperature overnight. Additional mixing of the 70 % and 85 % MO samples was conducted the following day using a variety of techniques in an attempt to remove some of the defects (air pockets) from the samples. This approach was unsuccessful in reducing the number of defects present in the samples, but it resulted in significant sample losses onto the implements used for mixing of the samples.

For these reasons the 70 % and 85 % MO NMR samples were prepared a second time using slightly modified procedures. The 70 % MO sample was prepared directly in a 5 mm NMR tube as described previously *except* that the MO was melted at 54°C *prior* to the addition of 54°C D_2O or H_2O. The sample was briefly mixed using a thin wire before being wrapped in aluminum foil and placed on the bench top to equilibrate overnight at ambient temperature. The 85 % MO sample was prepared in a 1.5 mL microcentrifuge using the procedure previously described in this Section (2.3.2.2) with the following changes: the MO was melted at 54°C *prior* to the addition of 54°C D_2O, the sample was vortex mixed and centrifuged at 13,000 × g for a total of 100 minutes at ambient temperature, before being incubated in a 54°C water bath for 2 hours. The sample was equilibrated overnight at ambient temperature on the benchtop and the following day the 85 % MO sample was carefully transferred to a 5 mm NMR tube using a pipettor (Gilson, Pipetman, P1000) with a pipette tip that had been cut off to yield a larger opening to facilitate the pipetting of a highly viscous sample.

At all times the NMR samples (20 %, 70 % and 85 % MO (wt/wt) in D_2O or H_2O) were stored upright on the benchtop in a tall beaker wrapped in aluminum foil to protect from light. The sample of MO in deuterated chloroform was stored at 4°C. ^1H NMR spectra of non-spinning samples were collected at 20°C on a Bruker AMX400 spectrometer. Data were processed using the software program *MestReC* (Cobas and Sardina, 2003). The data ('.fid' files) were imported into *MestReC* and processed using a Fourier transformation (FT). Prior to application of the FT, the spectral data (free induction decay) were zero filled by a factor of two and apodized using an exponential function with a line-broadening value close to the digital resolution of the spectrum. The spectra were phased using the automatic phase correction followed by manual phasing if required. An automatic baseline correction was applied to the spectra before they were referenced either to water or to chloroform. Referencing to water was conducted using the

following relationship that accounts for temperature dependence of the water chemical shift:

$$\delta\ [\text{ppm}] = 5.0057 - 0.0093286 \times T\ (\text{in } ^\circ\text{C}) \qquad [2.1]$$

Where: δ = water chemical shift (ppm)
T = temperature (in °C)

2.3.3 Optimized Solution NMR Sample Preparation of Cubic Phases

A method for the preparation of homogeneous defect-free solution NMR samples of MO G-type cubic phases was developed based on the sample preparation procedures outlined in Section 2.3.2. The method involved the preparation of cubic phase samples using centrifugation, followed by the transfer of the sample into a hollow glass tube that was then placed in a 5 mm solution NMR tube.

2.3.3.1 Apparatus

Custom NMR Tube Inserts

Thin-walled NMR grade quartz tubing of the type used to make Stem Coaxial Inserts for 5 mm NMR tubes was ordered from Wilmad LabGlass (Buena, New Jersey, USA). This item is not available in the company catalogue so a custom quote had to be requested. The outer diameter of the tubing was 4.15 mm which was the maximum size of tubing that would fit into a standard 5 mm solution NMR tube. The thickness of the glass used to make the tubing was ~ 0.36 mm, which was about the same thickness as the glass used to make standard 5 mm solution NMR tubes. The length of the tubing was ~ 203 mm, which was longer than the length of a 5 mm solution NMR tube which is typically around 177 mm. This made it difficult to seal the NMR tube and insert properly. The quartz tubing was very hard and brittle so it could not be safely cut using a glass cutter. It was taken to the glass shop where it was cut into 3 equal pieces ~ 67 mm in length using a saw. The NMR tube inserts (as they will now be referred to), were cleaned prior to use as outlined in Section 2.3.2.2 for 5 mm solution NMR tubes. A full description of the use of these inserts for cubic phase NMR sample preparation can be found below in Section 2.3.3.2.

Cut-off NMR Tubes

It was necessary to find tubes that were of the appropriate size and shape to accommodate the custom NMR tube inserts *and* that could withstand prolonged centrifugation at 13,000 × g. The ideal tubes for sample preparation would be as close in size as possible to the custom NMR tube inserts so as to minimize the amount of sample that was wasted during sample preparation. Sample wastage would be a major concern for the preparation of peptide or protein-containing cubic phases because these samples would be prepared using expensive and difficult to obtain labelled peptides and proteins. The most suitable tubes for cubic phase sample preparation turned

out to be 5 mm solution NMR tubes that had been cut-off to an appropriate length to facilitate centrifugation in a desktop centrifuge. The 5 mm solution NMR tubes were taken to the glass shop where they were cut down to a length of 45 mm using a saw to avoid cracking of the quartz tubes. The cut-off 5 mm solution NMR tubes were cleaned prior to use in the same manner as the regular 5 mm solution NMR tubes were cleaned.

Custom Centrifuge Inserts

In order for the cut-off 5 mm NMR tubes to be centrifuged in the desktop centrifuge appropriate inserts for the centrifuge rotor had to be found. Some of the inserts for the rotor that were commercially available were of the correct size to accommodate the cut-off NMR tubes, however the inserts were hollow tubes that did not have a bottom and the cut-off NMR tubes eventually cracked when they were centrifuged using these inserts. For this reason custom centrifuge rotor inserts had to be made that would support both the sides and the bottoms of the cut-off NMR tubes during centrifugation. The machine shop made plexiglass inserts for the desktop centrifuge rotor (Kendro Laboratory Products, Heraeus Biofuge Fresco, 7500 3325 rotor, fixed angle at 40°) that were designed to accommodate the cut-off NMR tubes.

2.3.3.2 Sample Preparation Procedure

Unless otherwise noted samples were prepared using the same reagents and procedures that were previously described in Section 2.3.2. The desired amount of MO was weighed out on weighing paper using an analytical balance and carefully transferred to a cut-off 5 mm NMR tube using a small plastic funnel to minimize the loss of MO during transfer. The tube was sealed with laboratory film and placed in a floating foam block in a 55°C water bath for ~ 30 minutes. The melted lipid was stirred using a flame-sealed 25 μL glass micropipette (Drummond, Microcaps, Cat. No. 1-000-0250) to remove any bubbles, the tube was re-sealed with laboratory film and returned to the water bath for an additional 10 minutes. Samples were prepared using water containing 10 % D_2O in order to provide the lock signal required for solution NMR studies. Distilled deionized water containing 10 % D_2O (CDN Isotopes, 99.9 atom % D, Cat. No. D-175, 100 g) was prepared and filter sterilized using a 0.22 μm syringe filter (Millipore, Millex-GP, Cat. No. SLGPR25LS) into sterile 1.5 mL microcentrifuge tubes as 1 mL aliquots that were stored at 4°C. The pH of the 10 % D_2O solution was ~ 5.5 as determined using pH indicator strips (EM Science, colorpHast pH 4.0 - 7.0, 0.2 pH unit resolution, Cat. No. 9582). An aliquot of 10 % D_2O was incubated in the 55°C water bath for a minimum of 10 minutes prior to sample preparation. The pipette(s) to be used for sample preparation were calibrated immediately prior to use with water that had been equilibrated to temperature in the 55°C water bath.

The desired amount of 55°C 10 % D_2O was added to the melted lipid and the sample was immediately stirred with a flame-sealed 25 μL glass micropipette. The sample tube was sealed with laboratory film and placed in the custom centrifuge insert in the temperature controlled desktop centrifuge (Kendro Laboratory Products, Heraeus Biofuge Fresco, 7500 3325 rotor, fixed angle at 40°) where the sample was centrifuged at 21°C at the minimum speed required for the minimum length of time required to obtain a homogeneous sample. This was done to minimize the risk that the sample tube would crack since the quartz NMR tubes were not designed for centrifugation. The sample was wrapped in aluminum foil and equilibrated on the benchtop overnight at ambient temperature. Once a homogeneous sample had been obtained, the sample was carefully extruded into one of the hollow custom NMR tube inserts. This was done by slowly and steadily pushing the insert into the cut-off NMR tube. Once the insert had been pushed to the bottom of the cut-off NMR tube it was twisted a half-turn, or until the seal made by the cubic phase between the insert and the cut-off NMR tube had been broken. Then the insert was slowly removed from the cut-off NMR tube. The insert was wiped off with a tissue (Kimberly-Clark, Kimwipes EX-L, Cat. No. 34155) to remove any excess cubic phase that was stuck on the outside of the insert. The insert was then placed in a clean 5 mm NMR tube where it slowly sank to the bottom of the tube. The NMR tube was capped and sealed with laboratory film to prevent any water evaporation. NMR samples were stored upright at room temperature wrapped in aluminum foil to protect from light.

2.4 Results

2.4.1 Samples Prepared in 0.6 mL Microcentrifuge Tubes

When water was first added to the samples they appeared to be primarily 'dry' because the quantity of water that was added was not sufficient to 'wet' all of the lipid in the sample. After the samples had been left to equilibrate overnight at ambient temperature the surfaces of the samples where the water had been added had become slightly translucent. Following centrifugation for 2.5 hours at 10,000 × g, the three samples had transformed into stiff transparent gels containing some small cloudy regions. The 'stiffness' of the gels was readily apparent in that the sample 'meniscus' remained at the angle of the fixed angle rotor after the samples had been removed from the centrifuge.

When the samples were inspected the following day they appeared unchanged. In an attempt to obtain homogeneous samples without any small cloudy regions, the samples were centrifuged overnight. Following the 16 hour spin, the samples still had some cloudy regions, so they were heated in a waterbath at 54°C for 4 hours. Immediately following removal from the waterbath the

sample containing 60 % MO was completely cloudy, the sample containing 65 % MO was cloudy with a clear region at the surface of the sample and the sample containing 70 % MO was completely clear. After the samples had been allowed to equilibrate back to room temperature, the sample containing 60 % MO had a few cloudy regions, whereas the samples containing 65 % and 70 % MO were completely clear, with the exception of a small cloudy region at the surface of the samples. It was noticed that the samples had undergone some dehydration evident from the condensation that was observed on the inside of the outer tube. Following an additional 2 weeks of equilibration at room temperature, the samples containing 60 % and 65 % MO became completely clear and the sample containing 70 % MO became predominantly clear with a thin cloudy layer at the surface, and opaque speckles throughout the sample.

2.4.2 Samples Prepared Directly in 5 mm Solution NMR Tubes

Prior to incubation of the NMR samples in the 60°C water bath, the 20 % MO sample had two distinct regions: a clear upper layer consisting of excess water, and a translucent lower layer consisting of wetted lipid, the 70 % MO sample had three distinct regions: a clear upper water layer, a clear middle gel layer, and a white bottom layer of un-wetted lipid, and the 85 % MO sample had two distinct regions: an upper layer consisting of a clear gel, and a white lower layer consisting of un-wetted lipid. When the NMR tubes were placed in the water bath, the un-wetted lipid in the 70 % and 85 % MO samples became clear and melted, pooling at the bottom of the NMR tubes. The 20 % MO sample remained unchanged during heating. In an attempt to obtain samples containing a homogeneous lipid phase, all three NMR samples were mechanically mixed.

This approach was successful for obtaining a homogeneous *lipid phase* throughout each of the samples, however the *appearance* of each of the NMR samples was not homogeneous. In the 20 % MO sample there were still two distinct regions visible, the translucent cubic phases near the bottom of the tube, and excess water at the top of the tube. Numerous air pockets were present in both the 70 % and 85 % MO samples where clumps of lamellar phase (cloudy gooey gel) or cubic phase (clear stiff gel) lipids had become stuck to the sides of the NMR tubes and to the spatula used for mixing of the samples. When additional mixing of these samples was conducted the following day, no improvements could be made with respect to the homogeneity of the samples. In fact, a significant amount of the samples were lost during mixing when they became stuck to the mixing implements and could not be efficiently removed.

When the 70 % and 85 % MO samples were prepared a second time, much more homogeneous samples were obtained. The 70 % MO NMR sample prepared in D_2O or H_2O were clear and had some small bubbles in them. The 85 % MO sample was uniformly cloudy and it

also had some air bubbles in it.

2.4.2.1 NMR Spectra of Monoolein:Water Phases

^{1}H NMR spectra were collected on the newly prepared samples of 70 % and 85 % MO, on the sample of 20 % MO and on the sample of MO in deuterated chloroform. The spectrum collected on MO in deuterated chloroform is shown in Figure 2.6. The spectrum collected on 70 % MO in D_2O is shown in Figure 2.7. The spectrum of 20 % MO in D_2O was almost identical to the one obtained for 70 % MO in D_2O so it is not shown. The spectrum of 85 % MO in D_2O is in Figure 2.8 and the spectrum of 70 % MO in H_2O is shown in Figure 2.9.

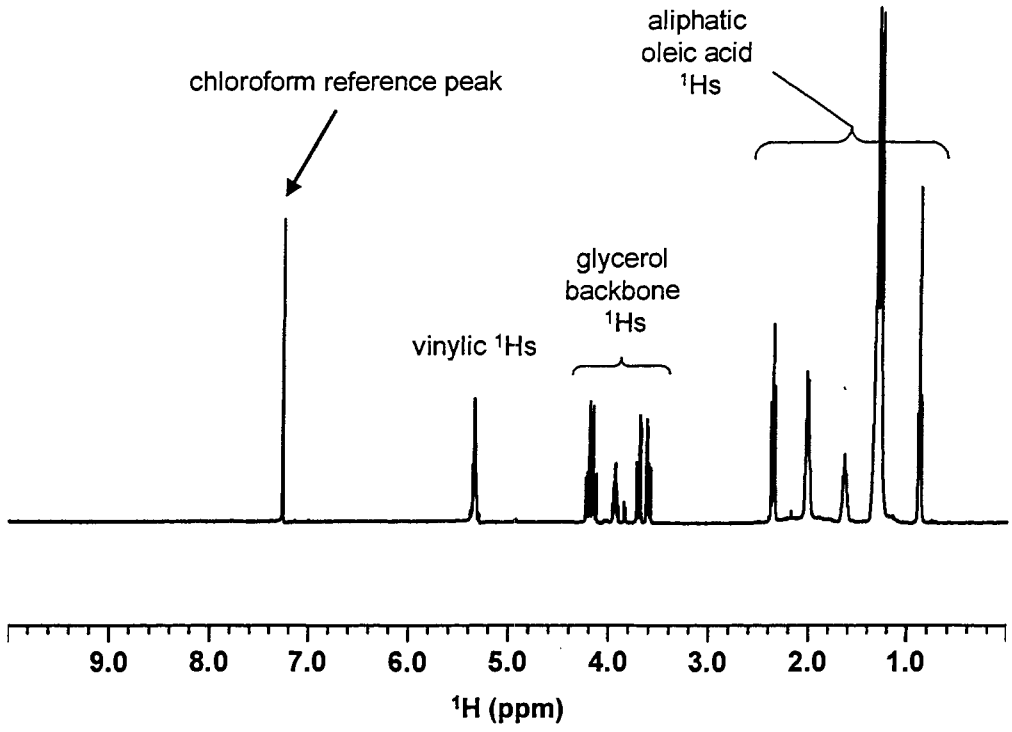


Figure 2.6. **^{1}H NMR Spectrum of Monoolein in Deuterated Chloroform at 20°C**

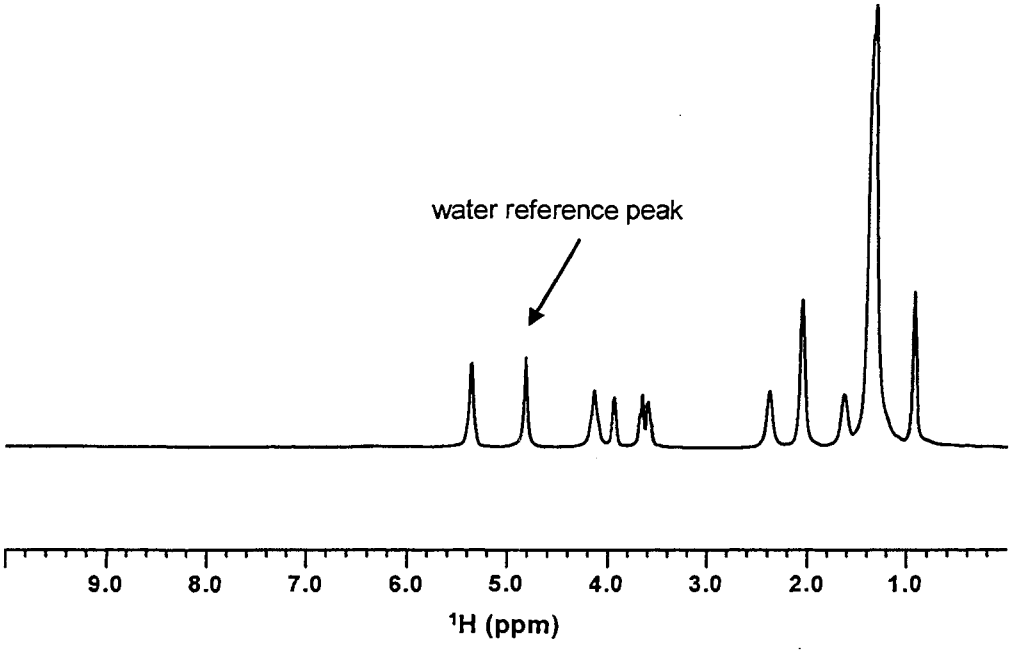


Figure 2.7. **^{1}H NMR Spectrum of 70 % Monoolein (wt/wt) in Deuterated Water at 20°C**

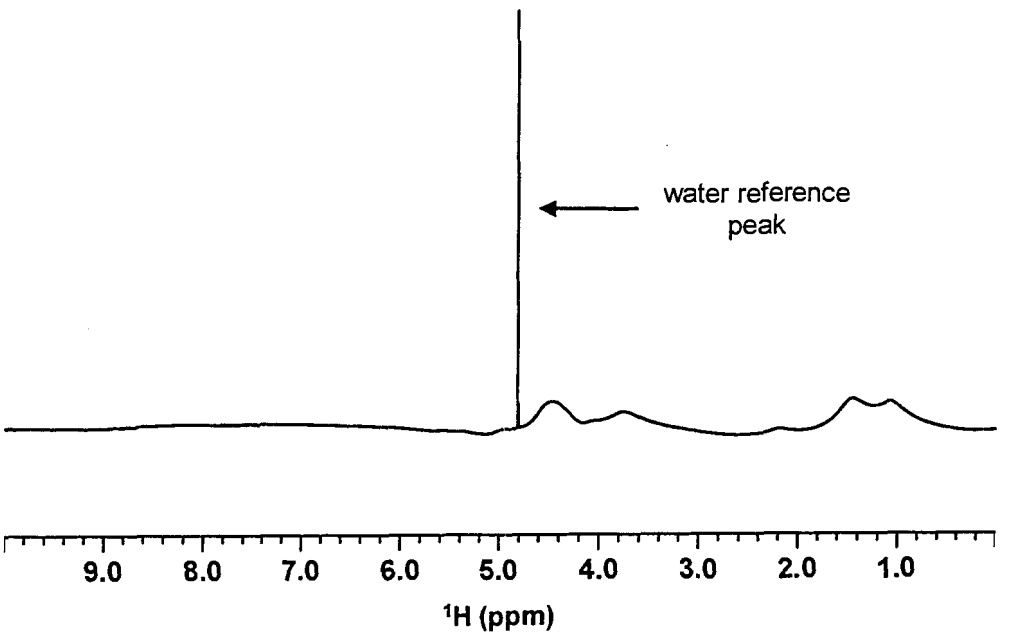


Figure 2.8. **^{1}H NMR Spectrum of 85 % Monoolein (wt/wt) in Deuterated Water at 20°C**

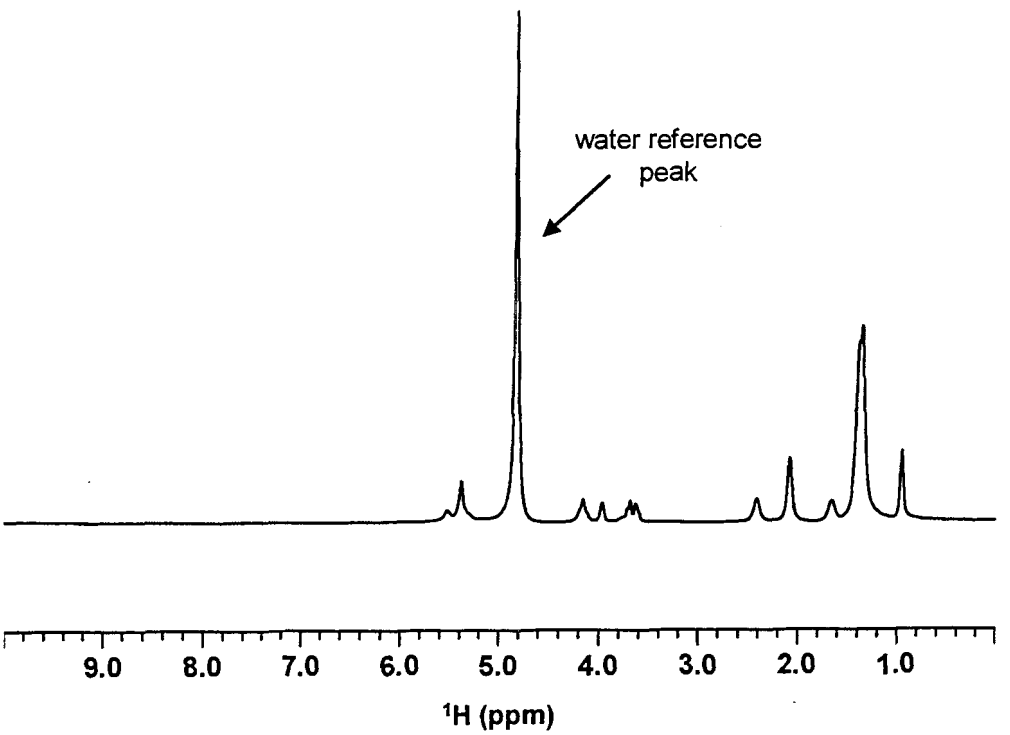


Figure 2.9. **^{1}H NMR Spectrum of 70 % Monoolein (wt/wt) in Water at 20°C**

2.5 Discussion

2.5.1 Sample Preparation in 0.6 mL Microcentrifuge Tubes

Homogeneous cubic phases formed by MO and water were successfully prepared using sample preparation information from the literature and the data provided in the phase diagram shown in Figure 2.1. It was found that the samples needed to be centrifuged in order to obtain proper mixing of the lipids and water. Using this procedure it was possible to obtain homogeneous cubic phases within a few days, instead of the weeks or months that might otherwise be required for the samples to come to equilibrium. The samples containing 60 % and 65 % MO (wt/wt) in water appeared to have formed homogeneous cubic phases evidenced by the transparency and stiff consistency of the samples. It appeared that the sample of 70 % MO (wt/wt) in water had also formed a cubic phase given the stiffness and general transparency of the sample. However, this sample was not homogeneous and contained small opaque speckles throughout the transparent gel which dominated the sample. It was likely that the 'speckles' in the sample were local regions of lamellar phase lipid since the sample had undergone some dehydration during heating. If enough water had been lost during heating of the sample its composition would have changed enough so that the sample would form a mixture of the lamellar (cloudy) and cubic (clear) phases instead of simply forming a homogeneous cubic phase. This explanation is consistent with the observation of some cloudy regions (presumably lamellar phase lipid), in an otherwise clear (cubic phase) sample.

It was found that the cubic phases formed by MO and water were very stiff and sticky, with a consistency similar to that of vacuum grease. This property of the cubic phases would make it very difficult to transfer the samples from one tube to another without significant sample loss. For this reason it would be preferable to prepare the cubic phases directly in the tube to be used for data collection, in this case a 5 mm solution NMR tube.

2.5.2 NMR Sample Preparation Directly in 5 mm Solution NMR Tubes

Preparation of samples containing 20 % and 70 % MO (wt/wt) in D_2O or H_2O directly in 5 mm solution NMR tubes met with moderate success. Initially there were some difficulties in obtaining adequate mixing of the lipids and water in the 70 % and 85 % MO samples because the 5 mm NMR tubes could not be centrifuged. It was found that because MO has a fairly low melting point (35.0°C), it was possible to melt the lipid *prior* to the addition of water and to then mechanically mix the sample (Lutton, 1965). If mechanical mixing was conducted only very briefly using a thin implement such as a wire, then a minimal number of sample defects (air bubbles) were introduced. After initial mixing it was necessary to equilibrate the sample for a

sufficient length of time (a couple of days) to achieve optimal sample homogeneity. Due to the reduced viscosity observed for the lamellar phase (85 % MO) sample compared with the cubic phase (70 % MO) sample, it was possible to prepare the 85 % MO sample using centrifugation, and to pipette the sample into the solution NMR tube. This approach was effective for obtaining a homogeneous lamellar phase sample, however there was still the problem of sample defects (air bubbles) that were introduced when the sample was transferred to the solution NMR tube. Because this sample was still quite viscous it was not possible to completely remove the air bubbles by mechanically mixing the sample.

Although it was possible to prepare cubic phase samples directly in 5 mm solution NMR tubes the method was not optimal because the samples contained defects which could potentially affect the quality of the NMR spectra. For this reason a reliable method for producing high quality NMR samples of MO cubic phases needed to be developed. Such a method would also need to be amenable to the preparation of high quality NMR samples of peptide or protein-containing cubic phases.

2.5.3 NMR Spectra of Monoolein:Water Phases

Despite the defects (air bubbles) in the 70 % and 85 % MO NMR samples, and the presence of two phase regions (water and cubic phase) in the 20 % MO NMR sample, the ^{1}H NMR spectra that were collected were of surprisingly good quality. The ^{1}H NMR spectrum of MO dissolved in deuterated chloroform is shown in Figure 2.6. The peak at 7.24 ppm corresponds to the ^{1}H peak of deuterated chloroform which was used to reference the spectrum. All of the other peaks in the spectrum originated from MO. This spectrum was collected for comparison with the ones collected on the samples of MO:water phases. In Figure 2.7 and Figure 2.9 the ^{1}H NMR spectra obtained for 70 % MO prepared in D_2O and H_2O are shown. The water peak at 4.82 ppm was used to reference these spectra. The other peaks visible in the spectra are from MO. All of the peaks are fairly sharp and well resolved indicating that the lipids and water are undergoing rapid isotropic reorientation in the MO G-type cubic phase. In the ^{1}H NMR spectrum collected on 70 % MO in water (see Figure 2.9), the lipid signals could still be clearly seen because of the extremely high lipid content of the sample. This indicated that cubic phase samples prepared using water could be studied using ^{1}H NMR techniques without the need for suppression of the water signal. This would be of particular importance in NMR studies of peptide and protein-containing samples which would need to be prepared in water, and not in D_2O, to avoid loss of signal from the exchangeable amide protons.

The NMR spectra collected on samples of 20 % MO in D_2O were almost identical to those obtained for the 70 % MO samples (spectra not shown). This result was surprising because the

samples appeared to have two distinct regions, a translucent region at the bottom of the NMR tube and a region of excess water in the remainder of the NMR tube. The only differences in the spectra of these D-type cubic phases, and those obtained for the G-type cubic phases, were the shapes of the peaks at ~ 3.6 ppm. The ^{1}H NMR spectrum of the sample of 85 % MO prepared in D_2O is shown in Figure 2.8. The sharp peak that is observed is from residual water in the D_2O used to prepare the sample, and the broad 'peaks' originate from MO. This spectrum clearly shows the anisotropy that would be expected for a lamellar phase sample.

The NMR spectra collected on cubic phase samples of MO in water were consistent with those previously published in the literature (Lindblom and Rilfors, 1989) and clearly demonstrated that the lipids and water in the samples were experiencing rapid isotropic reorientation. These results were very promising for the potential of 70 % MO cubic phases for solution NMR studies of incorporated peptides or proteins.

2.5.4 Optimized Solution NMR Sample Preparation of Cubic Phases

A method for the preparation of homogeneous defect-free solution NMR samples of MO G-type cubic phases was developed that could easily be applied to the preparation of peptide or protein-containing cubic phases. The method also minimized the amount of sample that was wasted during sample preparation which would be of particular importance when applying the technique to the preparation of cubic phase samples containing labelled peptides or proteins.

2.6 Conclusions

The cubic phases formed by mixtures of MO and water yielded NMR spectra with sharp well defined peaks. This result indicated that the lipid and water molecules in these cubic phases were undergoing rapid isotropic reorientation facilitating their study using solution NMR techniques. It was believed that by analogy it would be possible to study membrane peptides or proteins that were incorporated into the cubic phase using similar solution NMR techniques. Since a method had been established for the reproducible preparation of high quality cubic phase NMR samples, the next logical step was to identify and obtain sufficient quantities of a labelled membrane peptide or protein for solution NMR studies in the MO cubic phase.

Chapter 3: Rational Design, Recombinant Expression and Purification of ^{15}N- and ^{13}C-Labelled Model Transmembrane Peptides for NMR Studies

3.1 Objectives

To design, recombinantly express and purify ^{15}N- and ^{13}C-labelled transmembrane peptide models of integral membrane proteins for solution NMR studies in monoolein cubic phases.

3.2 Introduction

Lipid cubic phases were identified as potential membrane mimetic environments for solution NMR studies of membrane peptides and proteins. One of the major advantages of using cubic phases for the solubilization of membrane peptides and proteins is that very high lipid-to-peptide/protein ratios can be maintained within the samples, reducing the chances that peptide/protein aggregation will occur. The G-type cubic phase formed by mixtures of MO and water under ambient conditions was identified as the most suitable cubic phase for NMR studies of incorporated membrane peptides and proteins. These cubic phases have very high lipid concentrations, typically around 70% MO (wt/wt) in water (Lindblom and Rilfors, 1989). The high lipid content of these cubic phases could prove problematic for ^{1}H NMR studies of incorporated peptides and proteins which can probably *not* be incorporated into the cubic phase at concentrations comparable to those of the lipids without perturbing the phase behaviour of the system. Although membrane proteins had been successfully incorporated into cubic phases to relatively high concentrations, they had not been incorporated at the concentrations necessary to facilitate the observation of ^{1}H signals originating from the proteins, over the background ^{1}H signals originating from the lipids. Since MO was not available in its perdeuterated form, it was necessary to suitably label any peptides or proteins to be studied in order to facilitate the collection of NMR data where only the signals originating from the labelled peptide or protein would be observed.

3.2.1 Candidate Membrane Peptides or Proteins for NMR Studies

In order for a protein incorporated into the cubic phase to be studied using solution NMR techniques, it must experience rapid isotropic reorientation. The integral membrane proteins that have been studied thus far in the cubic phase were quite large, with multiple membrane-spanning domains, as well as extramembraneous loops. Since larger proteins will reorient more slowly than smaller proteins, it was desirable to select a smaller protein for initial studies. It was also important to study a protein with small, if any, extramembranous loops, to minimize the amount of protein in the aqueous pore. The most simple model of an integral membrane protein is a single membrane spanning α-helix. These peptides were identified as excellent candidates for study because the extramembranous regions of the peptide could be kept to a minimum. Based on these criteria, the ideal candidate peptides for solution NMR studies in the MO cubic phase were structurally well characterized model transmembrane peptides with a single membrane-spanning domain. At the time when this project was initiated, a number of studies had been conducted on membrane-active peptides, peptide fragments of the membrane-binding domains of integral and peripheral membrane proteins, and synthetic peptide models of the transmembrane regions of polytopic membrane proteins. The most thoroughly characterized of these peptides were the membrane-active peptides, the transmembrane region of the M13 bacteriophage coat protein and the polyleucine-based synthetic peptide models of the membrane-spanning regions of integral membrane proteins. These peptides will be discussed in detail below.

3.2.1.1 Membrane-Active Peptides

Hydrophobic and amphipathic membrane-active peptides are naturally produced by a number of organisms including: fungi, insects, amphibians and humans (Bechinger, 1997). These peptides interact with the membranes of living cells and can alter the spontaneous curvature of the membrane, and consequently the stability of the bilayer (Epand, 1998). At sufficiently high peptide concentrations, this disruption of the target cell membrane results in the loss of the transmembrane electrochemical gradient of the cell causing an inflow of excess water, cell swelling and eventual cell death (Bechinger, 1997). The exact mechanism by which these peptides elicit their activity is still the subject of much debate. In the case of some of these peptides, such as gramicidin and alamethicin, there is evidence to suggest that their cytotoxic activity is modulated through the formation of channels in the membrane, whereas for other peptides, such as magainin and melittin, a general disruption of membrane integrity is believed to modulate their cytotoxic effects (Weaver *et al.*, 1992; Yee and O'Neil, 1992; Vogt *et al.*, 1994; Ludtke *et al.*, 1995). The common feature of these peptides is that they adopt helical structures

with amphipathic characteristics upon lipid association (Epand, 1998).

Magainin

The magainins are a family of basic peptides isolated from the skin of the African clawed frog, *Xenopus laevis,* which inhibit the growth of bacteria, fungi, protozoa and cancer cells (Hirsh *et al.*, 1996). These peptides are water soluble and vary in length from 21 to 26 residues. In aqueous solution they assume random coil conformations, however they adopt an α-helical conformation upon lipid binding (Bechinger, 1997). In studies in lipid bilayers it has been shown that magainin binds in an orientation parallel to the bilayer surface and causes a decrease in the thickness of the bilayer (Ludtke *et al.*, 1995). At sufficient concentrations these peptides cause membranes to become permeable to ions. The exact mechanism by which this occurs is not known, but it has been hypothesized to occur through the transient formation of ion channels at sufficiently high peptide concentrations (Hirsh *et al.*, 1996).

Gramicidin A

Gramicidin A is a 15 residue peptide composed of alternating D- and L- amino acids that is produced by *Bacillus brevis* (Prosser *et al.*, 1994). It has antibiotic activity against gram-positive bacteria and it is available commercially as a topical bacteriostatic treatment (Woolley and Wallace, 1992). The amino acid sequence of gramicidin is very hydrophobic and does not contain any charged residues, it is also formylated at the N-terminus and contains a C-terminal ethanolamine group. These features make the peptide *very* insoluble in water (Woolley and Wallace, 1992). Gramicidin A forms ion channels in membranes as a symmetrical β-helical dimer with its formylated amino termini at the bilayer center (Vogt *et al.*, 1994; Epand, 1998). Due to its transmembrane orientation and stability, Gramicidin A is often studied as a model for integral membrane proteins (Lindblom and Rilfors, 1989). This peptide has also been shown to convert stable bilayers of DOPC to the hexagonal phase and it can convert micelles of lysoPC into bilayers (Vogt *et al.*, 1994).

Melittin

Melittin is a water soluble 26 residue cationic peptide that is the main component of honey bee venom (Dempsey, 1990). This peptide is known to spontaneously bind to cell membranes, and at sufficiently high concentrations melittin acts as a strong lytic agent of both native and model lipid membranes, an activity that is believed to be modulated by its ability to alter lipid organization (Brown *et al.*, 1982). It has been extensively studied using a variety of techniques, in both aqueous and membrane-mimetic environments (methanol, detergent micelles and multilamellar vesicles), in attempts to elucidate its structure and mode of action, which are still not well

understood. (Brown *et al.*, 1980; Lauterwein *et al.*, 1980; Brown and Wüthrich, 1981; Brown *et al.*, 1982; Bazzo *et al.*, 1988; Dempsey, 1990; Weaver *et al.*, 1992). In membrane-mimetic environments the peptide has been shown to adopt an amphipathic α-helical conformation with a bend (or hinge) involving residues Thr^{10} - Gly^{12} (Bechinger, 1997). Studies in DPC micelles revealed that the majority of the residues in the peptide were located near the surface of the micelle indicating that it was unlikely that the peptide was located within the micelle interior (i.e. in a 'transmembrane' orientation) (Brown *et al.*, 1982). The structure and aggregation state of melittin are dependent on peptide and salt concentrations, temperature and pH (Brown *et al.*, 1980; Stankowski and Schwarz, 1990). The interactions of melittin with lipid membranes is complex, and it has been shown that melittin can induce the micellization of bilayers, as well as membrane fusion (Dempsey, 1990; Monette and Lafleur, 1996).

Alamethicin

Alamethicins are 20 residue non-ribosomally synthesized antibiotic peptides containing the amino alcohol phenylalaninol (also referred to as Phol or O) and the amino acid α-aminoisobutyric acid (often referred to as Aib or B), that are produced by the soil fungus *Trichoderma viride* (Yee and O'Neil, 1992). They are hydrophobic and interact with cell membranes where they adopt an amphipathic α-helical structure and form voltage-gated ion channels (Esposito *et al.*, 1987; Epand, 1998). It is believed that the peptides insert into membranes in a transbilayer orientation even in the absence of a transmembrane potential (Franklin *et al.*, 1994). Alamethicin has been extensively studied as both a structural and functional model of α-helical membrane proteins with transport or channel functions (Woolley and Wallace, 1993; Keller *et al.*, 1996).

3.2.1.2 M13 Bacteriophage Coat Protein

The 50 residue coat protein of the M13 bacteriophage is an integral membrane protein that is inserted into the inner membrane of *E. coli* during the reproductive cycle of the phage (O'Neil and Sykes, 1988). This peptide has been extensively studied as a model monotopic integral membrane protein due to its small size and ease of isolation in gram quantities (Wickner, 1988; Deber *et al.*, 1993). Both wild type and mutant forms of this protein have been characterized using CD, molecular modelling and NMR techniques (Henry *et al.*, 1986; O'Neil and Sykes, 1988; Henry and Sykes, 1990; Deber *et al.*, 1993; Papavoine *et al.*, 1998). When bound to detergent micelles the protein adopts an α-helical conformation and is believed to be self-associated as a dimer, whereas in aqueous solution the protein is insoluble and highly aggregated, exhibiting some characteristics of β-sheet structure (O'Neil and Sykes, 1988; Wickner, 1988; Henry and Sykes, 1990; Deber *et al.*, 1993; Papavoine *et al.*, 1994).

Peptide Mimics of the Transmembrane Segment

Peptides derived from the 19 residue hydrophobic membrane-spanning domain of the M13 bacteriophage coat protein have been studied to gain insights into the sequence-specific non-covalent helix-helix interactions that occur between the transmembrane segments of membrane proteins (Wang and Deber, 2000). Transmembrane peptides with sequences corresponding to the wildtype (Tyr^{24}-Phe^{42}) and two mutant M13 coat proteins (where Val → Ala substitutions had been made at positions 29 and 31), were synthesized with flanking lysine residues and an N-terminal cysteine residue to have the general sequence CKKK-(transmembrane sequence)-KKK. The incorporation of terminal lysine residues into the peptide sequences were necessary to increase the solubility of the peptides thus facilitating their chemical synthesis, however the peptides were still sufficiently hydrophobic to spontaneously insert into membranes. Mutagenesis studies conducted on the full length M13 coat protein had shown that various Val → Ala mutations in the transmembrane region of the protein had position-dependent affects on helix stability (Deber *et al.*, 1993). It was believed that this was related to the involvement of select residues in the oligomerization of the protein in membrane environments. Helix stability could be assessed by the propensity of the proteins to aggregate upon heating which was known to induce an α-helix → β-sheet transition in the protein. The peptides were incorporated into lysoPC micelles and CD was used to measure the secondary structure of the peptides upon heating from 25°C → 95°C. It was found that the V31A substitution increased the thermal stability of the peptide whereas the V29A substitution did not change the stability of the peptide compared to the wild type peptide sequence.

3.2.1.3 Peptide Models of Integral Membrane Proteins

Polyleucine and Poly(leucine-alanine) Peptides

A number of studies have been conducted on synthetic polyleucine peptides as models of the membrane-spanning regions of integral membrane proteins (Davis *et al.*, 1983; Morrow *et al.*, 1985; Huschilt *et al.*, 1989; Roux *et al.*, 1989; Zhang *et al.*, 1992b; Zhang *et al.*, 1992a; Zhang *et al.*, 1995a; Liu *et al.*, 2002). These peptides were designed to form stable α-helices of the appropriate length to span a bilayer once. Lysine residues were placed at the ends of the peptides in order to maintain a transmembrane orientation for the peptides in membrane environments. Three polyleucine peptides were studied (L16, L20 and L24) with hydrophobic lengths of 24 Å, 30 Å and 36 Å, respectively. These helices were shown to be extraordinarily stable in studies investigating the effects of mismatch between the hydrophobic length of the peptide and the hydrophobic thickness of the bilayer.

It was believed that a peptide which was more sensitive to the lipid environment would be a

better model of protein transmembrane domains. So flexibility was introduced by modifying the peptide sequence to be one of alternating leucine and alanine residues. Subsequent studies were conducted on a poly(leucine-alanine) peptide LA12, and it was found that the conformation of this peptide was more sensitive to the composition of the surrounding lipid environment (Zhang *et al.*, 1995b; Harzer and Bechinger, 2000; Zhang *et al.*, 2001). It was believed that this peptide would be a good model of the hydrophobic membrane-spanning helices of integral membrane proteins.

WALP Peptides

Peptides containing interfacial tryptophan residues were designed based on the poly(leucine-alanine) peptides as models of the transmembrane segments of integral membrane proteins for the purposes of investigating whether or not α-helical peptides could induce the formation of nonbilayer structures in PC systems in response to conditions of hydrophobic mismatch (Killian, 1996; Morein *et al.*, 1997; de Planque *et al.*, 1998). At the time of this study, gramicidin was the only peptide known to induce the formation of nonbilayer structures in PC bilayers (Vogt *et al.*, 1994). Based on the structure and transbilayer orientation of gramicidin, it was believed that the tryptophan residues located near the lipid/water interface of the peptide were essential for its lipid phase modulating properties and was the reason that the poly(leucine-alanine) peptides were designed with interfacially located tryptophan residues (Aranda *et al.*, 1987). These peptides, called 'WALP' peptides, had the sequence Ac-GWW$(LA)_n$LWWA-ethanolamine.

Initially three such peptides were studied, WALP16, WALP17 and WALP19, with total calculated lengths of 25.5 Å, 27 Å and 30 Å respectively (including the terminal ethanolamine which was given the length of an additional amino acid) (Killian, 1996; Morein *et al.*, 1997). These peptides were very difficult to handle due to their extremely hydrophobic nature and it was found that the peptides aggregated over time when dissolved in various organic solvents. For this reason a special sample preparation protocol was developed where peptides dissolved in TFE were added to preformed lipid bilayers, mixed, lyophilized and rehydrated to yield homogeneous peptide/lipid dispersions. It was shown that the peptides formed *transmembrane* α-helices, indicating that the interfacially located tryptophan residues could act as membrane anchors since the peptides did not contain end charges to secure a transmembrane orientation of the peptides in bilayers. The peptides were shown to induce nonlamellar phases in PC model membranes in response to hydrophobic mismatch, providing the first examples of synthetic peptides with lipid phase modulating properties. Subsequent studies have shown that tryptophan residues in membrane proteins are preferentially located at the membrane interface (Yau *et al.*, 1998; Lew *et*

al., 2000; de Planque *et al.*, 2003).

TMX-1

A synthetic transmembrane peptide was designed to investigate the feasibility of engineering a peptide that would spontaneously insert across bilayers *and* also have monomeric water solubilty (Wimley and White, 2000). The peptide that was designed had a 21 residue hydrophobic core and the sequence Ac-WNALAAVAAALAAVAAALAAVAAGKSKSKS-NH_2. It was designed to incorporate N- and C-terminal 'caps' to favour stable helix formation. A charged C-terminus was employed to increase the water solubility and directionality of insertion of the peptide into vesicles. The peptide was designed to have bulky leucine and valine residues clustered on one face of the helix to favour helix formation and membrane insertion. The peptide was found to exist as soluble aggregates in solution. It was believed that the peptide might self-associate as an antiparallel α-helical dimer in solution and when bound to micelles due to the presence of an alacoil motif (Ala at *i*, *i* + *7*, at positions 3, 10 and 17) that was unintentionally present in the peptide. The peptide bound reversibly to POPC vesicles, but irreversibly to POPG vesicles (Wimley and White, 2000).

3.2.2 Stable Isotope Labelling of Peptides for NMR Studies

Due to their size, biological macromolecules have very complex 1D ^{1}H NMR spectra containing many overlapping peaks. In order to resolve some of this overlap, spectra are often resolved in additional dimensions using multidimensional NMR experiments that employ the use of heteronuclei. In these types of experiments nuclei such as ^{13}C and ^{15}N are used to 'filter' some of the ^{1}H signals. For example, it is possible to collect spectra where only ^{1}Hs directly bonded to ^{15}N are observed. In this type of experiment the number of peaks that are observed in the spectrum are equal to the number of residues in a polypeptide chain (the amide NH protons), plus the number of protons in nitrogen-containing side chains. The nuclei which are typically used for labelling are ^{13}C and ^{15}N, the solution NMR active isotopes of carbon and nitrogen. Luckily peptides and proteins contain many carbon and nitrogen atoms. However, at natural abundance only 1.07 % of carbon atoms and 0.368 % of nitrogen atoms in a peptide or protein are the NMR active isotopes of these nuclei, in contrast to hydrogen where 99.9885 % of the atoms in a polypeptide are of the NMR active isotope (Harris *et al.*, 2001). Fortunately it is possible to enrich peptides and proteins in these isotopes of carbon and nitrogen, or to incorporate other nuclei such as ^{19}F (which have similar NMR properties) using chemical or biosynthetic approaches.

3.2.2.1 Chemical Synthesis

One method of obtaining labelled peptide is to chemically synthesize the peptide using amino acids containing the label of interest. This approach can be used to incorporate labelled amino acids at selected positions in the peptide, or to uniformly label the entire peptide. This strategy is appropriate when the peptides of interest are of a moderate size (less than 50 amino acids in length), and do not require stable isotope-labelling at more than a handful of sites (Thai *et al.*, 2005). For long peptides, or those that require uniform stable isotope-labelling, the costs for chemical synthesis of labelled peptides in sufficient quantities for NMR studies are prohibitive.

3.2.2.2 Recombinant Peptide Expression in Bacteria

The most cost-effective method of incorporating stable isotope labels into a protein is to recombinantly express the protein to high levels in a host organism that can be grown in labelled media (McIntosh and Dahlquist, 1990; Emerson *et al.*, 1994; Jansson *et al.*, 1996; Kim *et al.*, 1997; Cai *et al.*, 1998). This strategy has been successfully employed in the study of many soluble and membrane-associated proteins, but has been of limited success for transmembrane peptides and integral membrane proteins. For peptides, the biggest challenge has been to avoid degradation of the overexpressed peptide by host proteases, and for membrane proteins the biggest problems have been associated with misfolding and aggregation of the overexpressed protein (Wood and Komives, 1999; Lajmi *et al.*, 2000; Lee *et al.*, 2000).

Another problem that can be encountered when overexpressing hydrophobic peptides is a high rate of cell mortality due to the sequestration of the overexpressed hydrophobic peptide in host cell membranes at concentrations which disrupt the structural integrity of the cell membranes leading to cell death (Majerle *et al.*, 2000; Thai *et al.*, 2005). One method of avoiding this problem is to express the peptide of interest as a fusion to a larger more stable 'carrier' protein (Kuliopulos and Walsh, 1994; Riley *et al.*, 1994; LaBean *et al.*, 1995; Pilon *et al.*, 1996; Drevet *et al.*, 1997; Ottleben *et al.*, 1997; Haught *et al.*, 1998; Kohno *et al.*, 1998; Lee *et al.*, 1998; Ottavi *et al.*, 1998; Zhang *et al.*, 1998; Gavit and Better, 2000; Jones *et al.*, 2000; Lee *et al.*, 2000; Majerle *et al.*, 2000; Therien *et al.*, 2002; Thai *et al.*, 2005).

The 'carrier protein' can be carefully selected to optimize the expression and purification of the peptide of interest by employing tags that: i) facilitate affinity column purification of the overexpressed protein, ii) selectively target the overexpressed protein to a particular compartment of the host cell (thus limiting toxicity), and iii) act as specific enzymatic or chemical cleavage sites to release the peptide product of interest from the carrier protein. The only disadvantage of this type of strategy is that a large amount of the cell's resources go into the production of the 'carrier' protein and not towards the production of target peptide, which is of particular concern

when producing isotopically labelled peptides and proteins for NMR studies. One method of resolving this problem has been the use of expression vectors where multiple copies of the peptide of interest are produced in tandem (Kuliopulos and Walsh, 1994; Lee *et al.*, 1998; Ottavi *et al.*, 1998).

3.2.3 Membrane Peptide Modulation of Lipid Phase Behaviour

A number of membrane peptides have been shown to affect the phase behaviour of lipid bilayers including: signal peptides, peptides derived from the transmembrane regions of integral membrane proteins and viral fusion peptides (Killian *et al.*, 1990; Morein *et al.*, 1997; Siegel and Epand, 2000). Due to their effects on biological membranes, it is not surprising that the membrane-active peptides nisin, gramicidin S and gramicidin A are also among this group of peptides capable of modulating lipid phase behaviour (Vogt *et al.*, 1994; de Planque *et al.*, 1998; Jastimi and Lafleur, 1999). In addition to the modulation of lipid phase behaviour by *naturally* occurring peptides, it has been demonstrated that the *synthetic* poly-leucine and WALP peptides can also induce changes in lipid phase behaviour (Killian, 1996; Morein *et al.*, 1997; de Planque *et al.*, 1998; Morein *et al.*, 2000; Rinia *et al.*, 2000; Liu *et al.*, 2001). It is believed that the modulation of lipid phase behaviour by synthetic α-helical transmembrane peptides is due to mismatch between the hydrophobic thickness of the lipid bilayer and the hydrophobic length of the peptide (Killian, 1996; Morein *et al.*, 1997; de Planque *et al.*, 1998).

3.2.3.1 Hydrophobic Matching

Hydrophobic matching is a term used to refer to the matching of the hydrophobic length of transmembrane peptides or proteins, to the hydrophobic thickness of biological membranes (Mouritsen and Bloom, 1984). It is reasonable to expect that the condition of hydrophobic matching would be met in biological membranes because the exposure of hydrophobic residues of proteins, or the acyl chains regions of lipids, would be energetically unfavourable (Dumas *et al.*, 1999). To investigate the validity of this theory experimentally, a series of synthetic peptides were studied in different bilayer-forming lipid mixtures to determine the effects of hydrophobic matching on lipid phase behaviour. The effects of WALP peptides 16, 17 and 19 residues in length on PC phase behaviour was studied using CD and ^{31}P solid state NMR (Killian, 1996; Morein *et al.*, 1997; de Planque *et al.*, 1998). It was found that the peptides could induce the formation of isotropic (later confirmed to be cubic) and hexagonal phases in PC bilayers of different hydrophobic thicknesses and that the propensity of the peptides to induce a certain phase could be predicted based on the degree of hydrophobic mismatch. This effect was believed to in part be modulated by the tryptophan resides located at the termini of the peptide. The induction

of changes in lipid phase behaviour is believed to be one mechanism by which peptides can 'compensate' for the condition of hydrophobic mismatch.

3.3 Materials and Methods

The membrane peptide and proteins that had been studied thus far in lipid cubic phases were either isolated from their host organism, or were commercially available. Since it was not possible to limit the contribution of lipid resonances to the ^{1}H NMR spectrum of MO cubic phases (due to the unavailability of perdeuterated MO), it was necessary to isotopically label any peptide to be studied with ^{15}N-, ^{13}C- or ^{19}F-. One of the primary criteria for the selection of a peptide for solution NMR studies in MO cubic phases was its amenability to isotopic labelling. The possibility of having the peptide chemically synthesized using labelled amino acids was investigated. In 1999 the cost of such a synthesis was very high, approximately $4600 for *one* ^{15}N-labelled amino acid in a peptide 20 amino acids in length on the 0.025 mM synthesis scale from the UBC Peptide Synthesis Facility. Since it would be necessary to incorporate more than one labelled amino acid *and* it would probably be necessary to study more than one peptide in order to properly characterize the system, the chemical synthesis option was deemed to be too costly to be feasible. The only viable option for obtaining labelling transmembrane peptides for study was through recombinant peptide expression.

3.3.1 Rational Design of Peptide Sequences for Study

Since labelling of peptides is quite expensive and time consuming, the peptide(s) for study needed to be carefully selected. The membrane-active peptides, although well characterized, were not deemed suitable as candidate peptides for study because: i) they contained non-coded amino acids (so they could not be recombinantly expressed and purified from *E. coli*), ii) their mode of membrane association was variable, iii) they induced changes in lipid phase behaviour and iv) they had a propensity to self-associate. To avoid any problems with the conformational averaging of a membrane surface-associating peptide that would be both dissolved in the water phase (random coil), *and* be associated with the cubic phase (α-helical), transmembrane peptides were identified as the ideal model peptides for study (Bechinger, 1997). Peptides with a single membrane-spanning domain and minimal flanking sequences were selected for study in order to favour rapid isotropic reorientation of the peptides within the cubic phase. Since the peptides were to be expressed and purified from bacterial sources, it was decided that at least one peptide of biological origin should be selected for study. Three synthetic model peptides (WALP, poly(leucine-alanine) and TMX-1), and one biologically-derived peptide (M13 coat protein-

derived peptides), were selected as the basis for the rational design of transmembrane peptides for NMR studies in the MO cubic phase.

The peptides were designed such that the hydrophobic length of the peptides matched the hydrophobic thickness of the MO 'bilayer' to avoid any problems with peptide-modulated alterations in the phase behaviour of the system. The hydrophobic core of the MO cubic phase is approximately 24 Å in thickness (see Figure 2.4), so the peptides were designed to have a hydrophobic membrane-spanning region 16 residues in length which was calculated using a length estimate of 1.5 Å per amino acid residue in an α-helical conformation (Davis *et al.*, 1983; Chung and Caffrey, 1994). In addition to the transmembrane residues, approximately 3 hydrophilic residues were required to flank the hydrophobic sequence at each end of the peptide. A number of peptide sequences were designed for study. Some of the peptides incorporated the use of N- and/or C-terminal caps, while others employed the use of interfacial tryptophan residues and/or lysine residues to: i) anchor the peptide in a transmembrane orientation upon membrane insertion, ii) increase the aqueous solubility of the peptides, and iii) modulate the hydrophobic length of the peptide using the 'snorkelling' property of the lysine side chains (Presta and Rose, 1988; Richardson and Richardson, 1988; Morein *et al.*, 1996; Stellwagen and Shalongo, 1997; Aurora and Rose, 1998; Park *et al.*, 1998; Tomich *et al.*, 1998; Penel *et al.*, 1999a; Penel *et al.*, 1999b; Tang and Deber, 2004). The proposed sequences were run through the secondary structure prediction program, *nnpredict* to confirm that the peptides would have an α-helical conformation. An alignment of the potential peptide sequences for study and their *nnpredict* results are shown in Figure 3.1.

Based on the results of the secondary structure predictions for the proposed sequences, and on information in the literature for the 'parent' peptides, three different peptides were chosen for expression in *E. coli* to facilitate the stable isotope labelling of the peptides for NMR studies in the MO cubic phase. These peptides are shown in Figure 3.2. Since it was not known how well a 'synthetic' or 'designed' peptide sequence would be expressed in *E. coli*, peptides of both biological *and* synthetic origins were selected for expression. Another consideration in the selection of transmembrane peptides for study was their potential modulation of the phase behaviour of the MO:water system. In order to decrease the chances that this would occur, the peptides that were selected for study had different properties with respect to their amino acid sequences, length and charge distributions.

3.3.1.1 WALK

The 'WALK' peptide was based on the WALP peptides, however in addition to the N- and C-

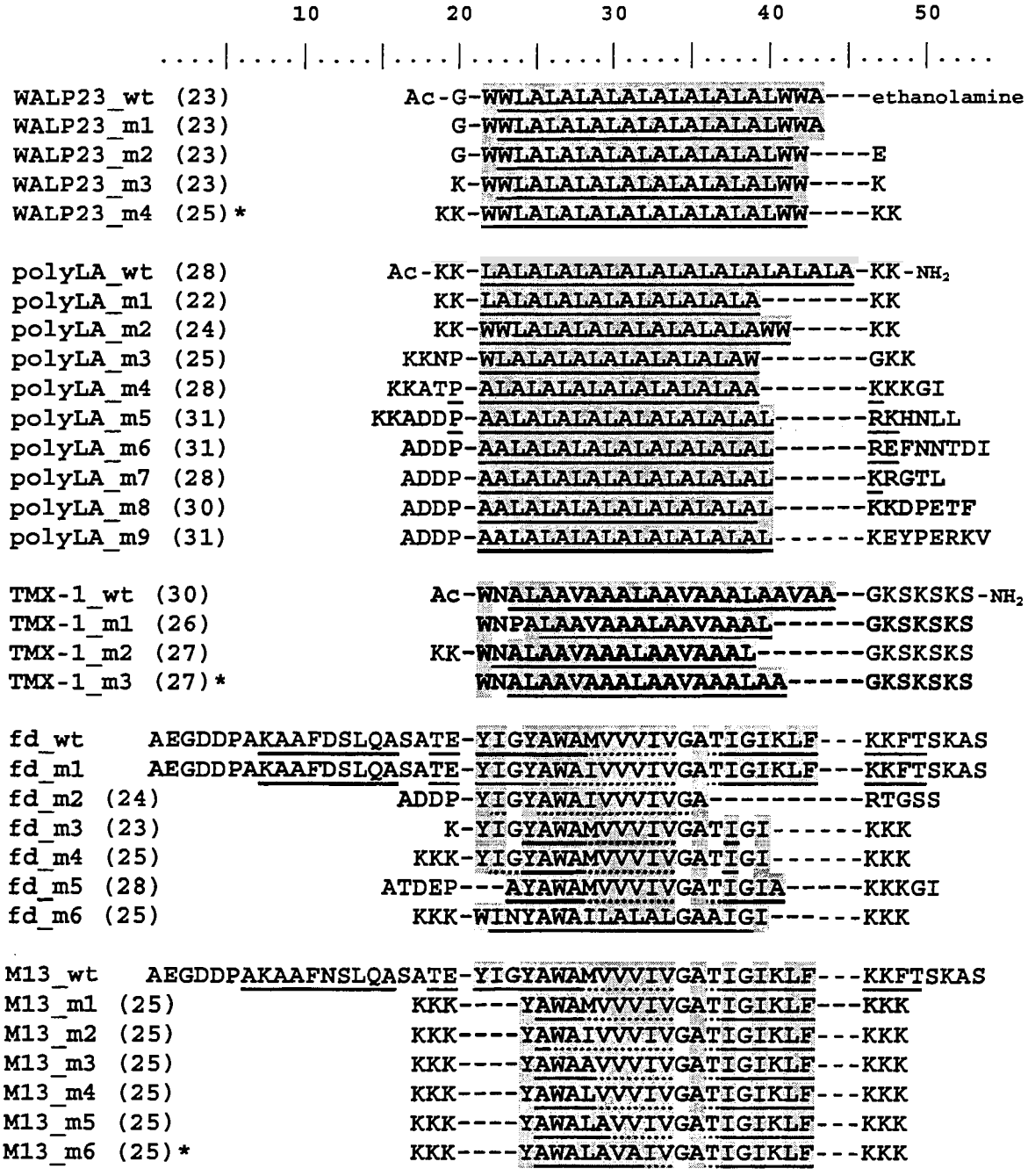


Figure 3.1. An Alignment of Transmembrane Peptide Sequences Designed for Recombinant Expression in *E. Coli*

The parent peptide sequences are denoted as *wt* and the modified peptide sequences are denoted as *m1, m2,* and so on. The total length of each peptide is listed in brackets following the peptide name. An asterisk indicates which peptides were selected for further study. Gaps in the alignment are indicated by dashes in the amino acid sequence. Hydrophobic residues in the putative membrane spanning regions of the peptides are coloured in grey. The program *nnpredict* was used to predict the likely secondary structure type for each residue in the peptide sequences. A solid underline indicates a predicted helix element, a dotted underline a predicted beta strand element and no underlining indicates no predicted secondary structure.

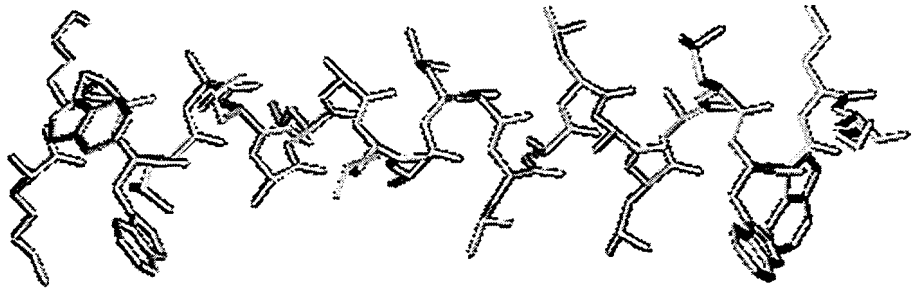
terminal tryptophan anchors, the peptide was end-capped with lysine residues having the sequence KKWWLALALALALALALALWWKK. This peptide had characteristics of both the poly(leucine-alanine) peptides (model transmembrane helices) and the WALP peptides (which employed interfacial tryptophan residues to anchor helices in a transmembrane orientation). However, these peptides had been shown to perturb lipid phase behaviour as a function of the hydrophobic mismatch between the hydrophobic length of the peptide and the hydrophobic thickness of the bilayer in which they resided. Another concern with the WALP peptides was their tendency to aggregate as a function of time. In order to minimize these potential problems a pair of lysine residues were added to each end of a WALP peptide of the appropriate hydrophobic length. These lysine residues would: i) increase the solubility of the peptide by increasing its polarity, ii) reduce the propensity of the peptide to self aggregate through the repulsion of like charges, iii) decrease the chances that the peptide would perturb the lipid phase behaviour of the system (since the acyl side chains of the lysine residues have the ability to 'snorkel', thus compensating for any differences in the hydrophobic length of the peptide and the hydrophobic thickness of the bilayer), and iv) reduce the overall hydrophobicity of the peptide, which would hopefully improve the bacterial expression and purification of the peptide.

3.3.1.2 TMK

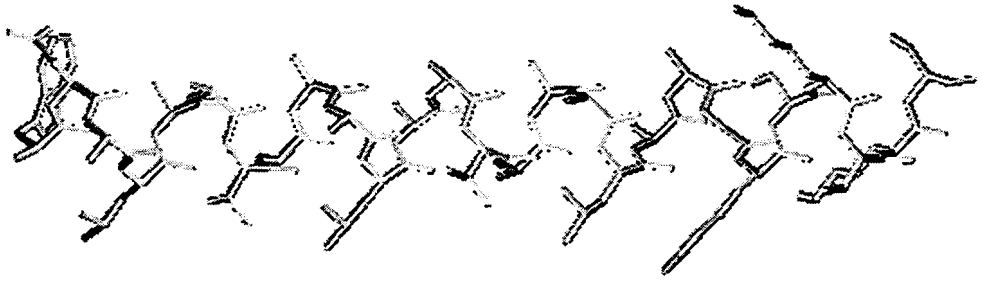
The 'TMK' peptide was based on the TMX-1 model peptide that was designed to spontaneously insert into lipid vesicles. The peptide needed to be truncated in order to maintain hydrophobic matching with the MO 'bilayer'. Three residues (VAA) were removed from the C-terminal portion of the transmembrane region of the peptide. This modification in peptide sequence was not expected to affect the structure of the peptide. The peptide had the sequence WNALAAVAAALAAVAAALAAGKSKSKS. It had the potential to be a particularly good candidate peptide for study because it had been engineered with N- and C-terminal caps and it was not as hydrophobic as the other peptides. It also contained a tryptophan residue so, if necessary, its aggregation state and lipid binding could be monitored using tryptophan fluorescence.

3.3.1.3 M13-LA

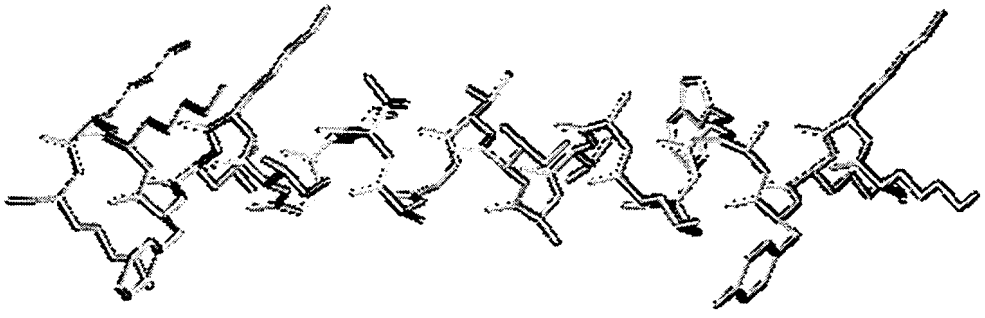
The M13-LA peptide was based on the lysine-capped transmembrane peptide derived from the M13 bacteriophage coat protein and had the sequence KKKYAWALAVAIVGATIGIKLFKKK. Two modifications to the peptide sequence were made to increase the α-helical propensity (and consequently the stability of the α-helical conformation of the peptide), and to disfavour the



WALK



TMK



M13-LA

Figure 3.2. Three Rationally Designed Model Transmembrane Peptides for Solution NMR Studies in the Monoolein:Water G-Type Cubic Phase

formation of β-sheet sheet structure (which was know to form when the protein aggregated). The two modifications were valine → alanine substitutions at positions 29 and 31. Full length protein containing this double mutation had been previously characterized and it was found to form helices that were more stable than those of the protein with wild type sequence (Deber *et al.*, 1993). Another modification (a M28L substitution), was made to the peptide sequence to make the peptide amenable to recombinant expression using an expression system that employed the use of cyanogen bromide (CNBr) to cleave the recombinantly expressed peptide from the carrier protein. The M13-LA peptide was expected to be more easily overexpressed and purified than the WALK peptide because it was less hydrophobic. If NMR spectra were successfully obtained, this peptide would have been amenable to assignment of chemical shifts and possible structure determination due to its more varied amino acid sequence (Brown and Wüthrich, 1981).

3.3.2 Bacterial Expression System Selection

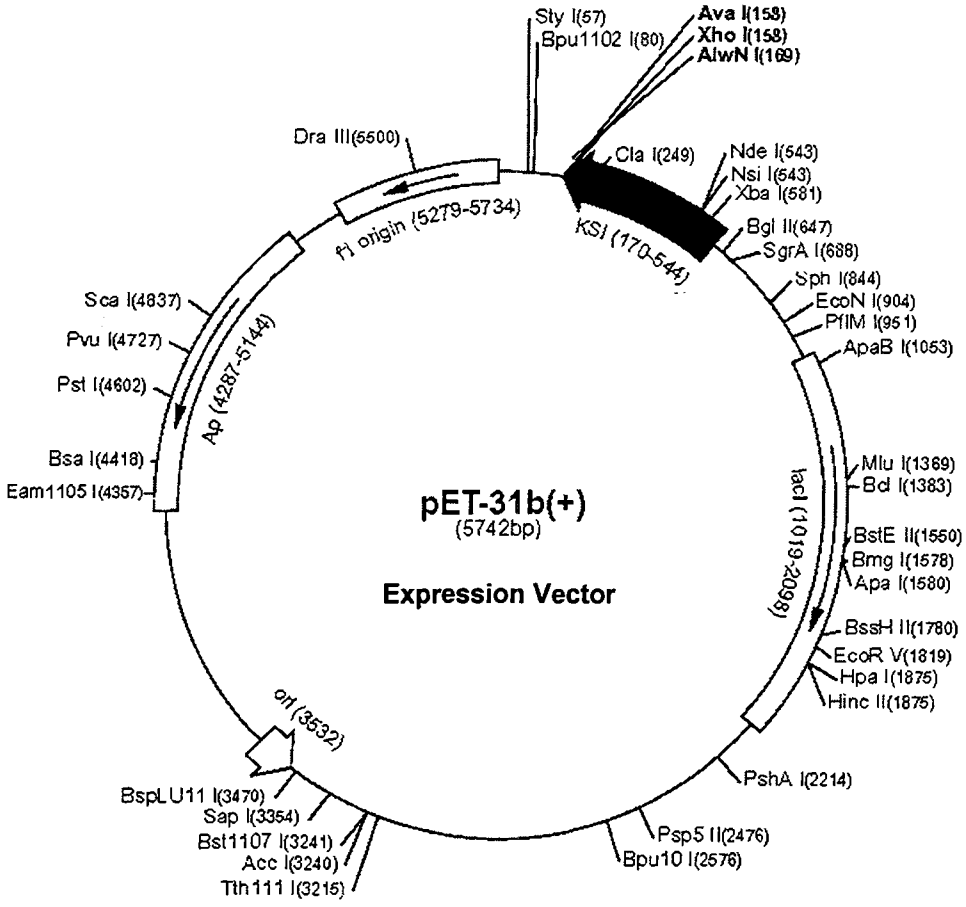
3.3.2.1 Expression Vector

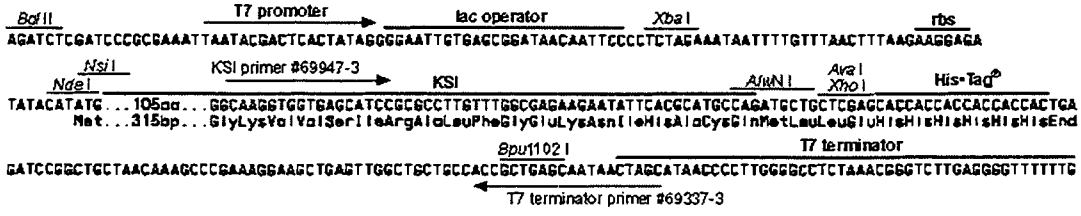
A good method for purification of recombinant proteins is through the use of Ni^{2+}-affinity chromatography to purify histidine-tagged proteins. Expressing the peptide as a His-tag fusion protein allows the tag to be removed by chemical cleavage using CNBr providing that the sequence of the expressed peptide(s) does not contain any internal methionine residues. Chemical cleavage is less complicated than enzymatic cleavage in that it is not necessary to engineer in the appropriate cleavage site or to maintain buffer conditions where enzyme activity is maintained and solubility of the hydrophobic peptides is maintained.

A plasmid designed specifically for the cloning and high level expression of peptides in *E. coli* was commercially available from Novagen. The vector map and cloning region for this plasmid are shown in Figure 3.3. The system was designed for the expression of peptides in *E. coli* as His-tagged fusion proteins to ketosteroid isomerase (KSI). The vector contains an ampicillin resistance gene. This system facilitates the production of high yields of small peptides (10-25 amino acids in length), in particular when cloned as tandem repeats, or larger peptides (25-75 amino acids in length). The fusion protein can be expressed at high levels as inclusion bodies and subsequently purified using Ni^{2+}-affinity chromatography. The peptides are cleaved from the KSI carrier protein using CNBr.

3.3.2.2 Expression Host

The expression host recommended for use with the pET31b(+)® expression vector was *E. coli* strain BLR(DE3)pLysS. This strain cannot undergo homologous recombination which may help to stabilize plasmids containing repetitive sequences, such as tandem repeats of protein-encoding





pET-31b(+) Cloning Region

Figure 3.3. Novagen pET 31b(+)® Expression Vector and Cloning Region

DNA sequences. BLR is a derivative of the protease deficient BL21 strain of *E. coli* designed for protein overexpression and purification. The 'DE3' designation indicates that this host strain carries a chromosomal copy of the T7 RNA polymerase gene under the control of the *lac* promoter making them suitable for the production of target proteins cloned into T7 expression vectors. The 'pLysS' designation indicates that the host strain carries the gene encoding for T7 lysozyme, a natural inhibitor of T7 RNA polymerase. This ensures that basal expression of T7 RNA polymerase does not occur prior to induction which is important when working with target proteins whose expression may negatively affect host cell growth and viability. The gene encoding for T7 lysozyme is on the chloramphenicol resistance plasmid. This strain also contains the tetracycline resistance transposable element.

3.3.3 DNA Oligonucleotide Design

3.3.3.1 Codon Usage

DNA oligonucleotides were designed to optimize the levels of protein expression by incorporating the highly used codons from *E. coli* (de Boer and Kastelein, 1986).

3.3.3.2 DNA Oligonucleotide Sequences

DNA sequences were designed using the codons associated with highly expressed proteins in *E. coli* in a manner so as to avoid tandem repeats in the DNA sequence as much as possible. In order to cleanly clone the DNA sequences into the pET31b(+)® vector, both the sense and antisense oligonucleotides were synthesized with 3 base pair overhangs to generate sticky ends for cloning the double stranded DNA into the plasmid vector which had been cut with the restriction enzyme *AlwN1*.

3.3.3.3 Oligonucleotide Synthesis

Phosphorylated DNA oligonucleotides were synthesized and gel purified by the Calgary DNA Synthesis Center.

3.3.4 Expression Vector Preparation

The pET31b(+)® plasmid DNA (Novagen, Cat. No. 69952-3, 10 μg) was supplied as a frozen (-20°C) stock solution with a concentration of 0.48 μg/μL. Since the DNA concentration of the stock solution was much greater than required for bacterial transformation a more dilute solution of plasmid DNA was prepared. The plasmid stock solution was thawed on ice, vortex mixed and 0.5 μL (Gilson, Pipetman, P20) of the solution was transferred to a sterile 1.5 mL microcentrifuge tube containing 1.2 mL of sterile TE Buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) using sterile technique. The solution was vortex mixed and both plasmid solutions were returned to the -20°C for storage. The final plasmid DNA concentration of the dilute solution was 0.2 ng/μL.

3.3.4.1 Bacterial Transformation

LB agar plates containing: 50 μg/mL ampicillin for the plating of transformed cloning host cells (Life Technologies GibcoBRL, Subcloning Efficiency DH5α Competent Cells, Cat. No. 18265-017, 2 mL), or 50 μg/mL ampicillin and 34 μg/mL chloramphenicol for the plating of transformed expression host cells (Novagen, BLR(DE3)pLysS Competent Cells, Cat. No. 69956-4, 5 × 200 μL), were pre-warmed in a 37°C incubator for a couple of hours prior to use to evaporate off any excess moisture. To avoid repeated freezing and thawing of competent cells they were frozen in ready-to-use aliquots. Competent cells were removed from the -80°C, thawed on ice, mixed by gently flicking the tube, quickly transferred to sterile 0.5 mL microcentrifuge tubes in 20 μL aliquots, flash frozen in a dry ice-ethanol bath and returned to the -80°C freezer.

A 20 μL aliquot of competent *E. coli* cells were thawed on ice just prior to use. Approximately 0.2 ng of plasmid DNA was added to the cells using sterile technique. The pipette tip was used to gently stir the cells before incubation on ice for 5 minutes. The cells were then heat shocked in a 42°C water bath for 30 seconds before being placed back on ice for 2 minutes. Using sterile technique, 80 μL of room temperature SOC media was added to the cells incubating on ice. The cells were then vortex mixed and placed in a 37°C incubator shaking at 250 rpm for 60 minutes to allow the cells to recover. Two dilutions of the transformation mixture were plated onto pre-warmed LB agar plates containing the appropriate antibiotics. A 40 μL pool of SOC media was added to the plate first when small volumes (< 20 μL) of transformation mixture was plated to ensure even plating of the cells. For expression host transformations typically 25 μL and 75 μL of cells were plated onto the appropriate plates. A few microliters of transformed cells were always plated onto plates that did not contain any antibiotics to use as a positive control to check for cell viability following transformation. Plates were left to sit for ~10 minutes on the benchtop to allow any excess liquid to be absorbed into the agar before being inverted and placed in a 37°C incubator overnight.

3.3.4.2 Plasmid Mini Preps

Single bacterial colonies grown on LB agar plates following transformation with pET31b(+)® plasmid DNA, were picked with a flame-sterilized innoculating loop and used to innoculate 2 mL aliquots of LB broth containing 50 μg/mL ampicillin and 34 μg/mL chloramphenicol in sterile 17 × 100 mm polystyrene culture tubes (Simport, Cat. No. T406-2). The cultures were incubated at 37°C overnight (~16 hours) while shaking at 250 rpm. The next day the absorbance of the cultures was measured at 600 nm to make sure that no more than 20 A_{600} units of culture would

be processed on a single DNA purification column. A commercially available plasmid DNA purification kit (Amersham Pharmacia Biotech Inc., GFX™ Micro Plasmid Prep Kit, Cat. No. 27-9601-02, 250 purifications) was used to isolate and purify pET31b(+)® plasmid DNA from transformed *E. coli* cultures. Plasmid DNA was isolated using the purification protocol ('Procedure B') that was provided with the kit. A few modifications to the protocol needed to be made.

Modifications Made to 'Procedure B' of the GFX™ Micro Plasmid Prep Kit

Because the centrifuge used for plasmid isolation only had a top speed of 11,500 × g (not 13,000 × g, as specified in *Procedure B*), the length of some of the spins were increased as follows: the cells were spun at full speed for 45 seconds (not 30 seconds) to pellet the cells, the samples were spun at full speed for 8 minutes (not 5 minutes) to pellet the cell debris, the samples were spun at full speed for 1 minute (not 30 seconds) to load the supernatant onto the GFX columns, the columns were spun for 2 minutes (not 1 minute) to remove the wash buffer and dry the matrix prior to DNA elution, and the plasmid DNA was eluted from the column by spinning at full speed for 2 minutes (not 1 minute). To avoid the chance of accidentally aspirating the cell pellet or DNA pellet, supernatants were removed from samples using a glass pasteur pipette instead of aspiration.

When time permitted, the GFX columns that had been loaded with plasmid DNA and washed with wash buffer, were allowed to dry overnight before the DNA was eluted. The additional drying time ensured that any residual ethanol on the column had evaporated prior to DNA elution. This extra step helped to improve the yields of plasmid DNA recovered from the columns. Water was not used to elute the DNA from the columns because the pH of pure water can be quite variable and DNA is more stable over the long term if it is stored in a buffered solution. DNA was eluted from the columns using TE buffer at pH 8.0 containing 0.1 mM EDTA, not 1 mM EDTA. This was done to avoid any potential interference of higher concentrations of EDTA on the functioning of any of the enzymes that would be used in subsequent steps involving the purified plasmid samples. It was however important to have a low level of EDTA present in the buffers chelate any contaminating metal ions that could interfere with enzyme function in subsequent reactions.

3.3.4.3 DNA Precipitation

If DNA solutions are too dilute for use in enzymatic reactions the DNA can be precipitated and re-dissolved in an appropriate amount of buffer. DNA was precipitated by measuring the volume of the DNA solution and adding a volume of 3M sodium acetate pH 5.2 that would yield a final

sodium acetate concentration of 0.3 M (*after* two sample volumes of ethanol had been added). After the addition of sodium acetate the sample was vortex mixed and two sample volumes of 95% ethanol were added. The sample was vortex mixed again and then it was frozen at -20°C for a minimum of 1 hour (typically 4 hours). After the sample was removed from the freezer it was centrifuged at 16,000 × g for 15 minutes and the supernatant was carefully removed and saved in a 1.5 mL microcentrifuge tube. An additional 750 µL of 70% ethanol was added to the sample to rinse the DNA pellet and the sample was centrifuged at 16,000 × g for 15 minutes. The supernatant was carefully removed and saved, and the DNA pellet was allowed to air dry overnight on the benchtop covered with aluminum foil. The dried DNA pellet was then dissolved in an appropriate volume of TE Buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0).

3.3.4.4 Restriction Enzyme Digest of Purified Plasmid DNA

Purified pET31b(+)® plasmid DNA was digested with the restriction enzyme *AlwNI* to create the 'sticky ends' required for ligation to the peptide-encoding DNA oligonucleotides. In a sterile 0.5 mL microcentrifuge tube approximately 3 µg of plasmid DNA was mixed with 3 µL of 10 × Buffer (New England Biolabs, NEBuffer 4, at 1 × contains 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate and 1 mM DTT at pH 7.9), 1 µL of the restriction enzyme *AlwNI* supplied at a concentration of 10U/µL (New England Biolabs, Cat. No. 514S, 500U), and enough sterile distilled water to yield a final reaction volume of 30 µL. The reaction was incubated at 37°C for 4 hours. The restriction enzyme was then inactivated by heating of the sample tube at 65°C for 25 minutes. The digested plasmid DNA was stored at -20°C.

3.3.4.5 Agarose Gel Electrophoresis

Agarose gels were cast and run using the Sub-Cell GT Mini Agarose Gel Electrophoresis System (BioRad, Cat. No. 170-4466). Agarose gels were prepared by melting agarose powder (Bio/Can Scientific, Eclipse LE Agarose, Cat. No. EMBGA-LE-100, 100 g) in an appropriate volume of 1 × TAE Buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) in the microwave, cooling the solution to ~ 60°C, adding ethidium bromide to yield a final concentration of 0.5 µg/mL, before pouring the melted agarose into the gel-casting tray. The ethidium bromide was added to facilitate the visualization and quantitation of the DNA under UV illumination. After the gel was cooled (~ 45 minutes), the comb was removed and the appropriate volume of 1 × TAE Buffer was added to the gel box.

Prior to loading of DNA samples onto the agarose gels, an appropriate amount of 6 × Agarose Gel Loading Buffer (15% aqueous Ficoll™ containing various combinations of 0.25% bromophenol blue, 0.25% xylene cyanol FF and 0.01% tartrazine) was added to the samples,

along with an appropriate amount of water to ensure that all of the samples had the same total volume. The dyes that were present in the loading buffers were used to indirectly monitor the migration of the DNA samples in the gel. The mobility of the dye was dependent on the agarose concentration of the gel. Dyes that *did not* have the same mobility as the DNA molecules that were being run on the gel were used for sample prep. This was to avoid problems with the dye absorbing the fluorescence of the DNA bands that lay under it when the gel was UV illuminated. Samples were vortex mixed to ensure complete mixing of the loading buffer and the DNA sample so the samples would sink in the wells when loaded on the agarose gel.

DNA ladders of various sizes were run on the agarose gels as standards to correspond to the sizes of DNA fragments being resolved. The ladders that were used as size standards contained: 50 bands 20 - 1000 bp in size in exact 20 bp increments (BioRad, 20 bp Molecular Ruler, Cat. No. 170-8201, 250 µL), 10 bands 100 - 1000 bp in size in exact 100 bp increments (BioRad, 100 bp Molecular Ruler, Cat. No. 170-8202, 250 µL) or 16 bands 500 - 8000 bp in size in exact 500 bp increments (BioRad, 500 bp Molecular Ruler, Cat. No. 170-8203, 200 µL). Ladders were prepared for running on the gels in the same manner as the DNA samples were. Prior to loading of the samples on the agarose gel they were centrifuged briefly to bring all of the liquid to the bottom of the tube. Samples were loaded into the wells of the agarose gel using a pipette (Gilson, Pipetman, P20). Before the gel was run, enough ethidium bromide to yield a final concentration of 0.5 µg/mL was added directly to the running buffer at the anode (red electrode) end of the gel box . Gels were visualized under UV illumination after they were run.

0.7 % Agarose Gels

Agarose gels used for running out restriction enzyme digested plasmid for subsequent purification were 0.75 cm thick, 7 × 7 cm in size and had an agarose concentration of 0.7 %. Samples to be run on the gel were prepared using 6 × Agarose Gel Loading Buffer containing bromophenol blue and tartrazine. A blank well was loaded with 1 × Agarose Gel Loading Buffer containing bromophenol blue, xylene cyanol and tartrazine to use as a reference for DNA migration within the gel. Bromophenol blue and xylene cyanol are known to migrate at the same rate as DNA molecules 400 - 500 bp and 4000 - 5000 bp and in size respectively, in 0.5 - 1.5 % agarose gels (BioRad, 2000). The gel was run at 75 V for ~ 70 minutes.

2.0 % Agarose Gels

Agarose gels 0.75 cm thick, 7 × 10 cm in size, with a composition of 2.0 % agarose, were used to resolve the tandem repeats of ligated DNA oligonucleotides. Samples were prepared using 6 × Agarose Gel Loading Buffer containing tartrazine. Two standards were routinely run on gels

used to resolve the tandem repeats. These were the 20 bp ladder which was prepared using 6 × Agarose Gel Loading Buffer containing xylene cyanol and tartrazine and the 500 bp ladder prepared using 6 × Agarose Gel Loading Buffer containing bromophenol blue and tartrazine. Gels were run at 75 V for ~ 2.5 hours. DNA bands were visualized using 305 nm UV light and were cut out using a low intensity 360 nm UV light box to avoid nicking of the DNA.

2.5 % Agarose Gels

Agarose gels 0.75 cm thick, 7 × 7 cm in size, with a composition of 2.5 % agarose, were used to resolve the PCR products from screens for insert-containing clones. Samples were prepared by mixing 5 μL of the PCR reaction mixture with an appropriate amount of 6 × Agarose Gel Loading Buffer containing tartrazine and distilled water (depending on the final sample volume to be prepared). Gels were run at 75 V for ~ 1.5 hours. The DNA ladders and loading buffers that were run on the gels were the same as those used for running out the ligated tandem repeats on 2.0 % agarose gels.

3.3.4.6 Extraction of DNA from Agarose Gel Slices

A kit was commercially available for the purification of DNA from agarose gel slices. This kit (Qiagen Inc., QIAquick™ Gel Extraction Kit, Cat. No. 28706, 250 purifications) was used for the purification of both plasmid and oligonucleotide DNA from agarose gel slices. The standard purification protocol provided with the kit was followed without the use of any additional wash steps. Purified DNA was eluted in 30 μL of the supplied buffer after a 5 minute incubation of the buffer on the column prior to centrifugation.

3.3.4.7 Dephosphorylation of DNA

DNA was dephosphorylated using shrimp alkaline phosphatase (Roche, Cat. No. 1758250, 1000U) supplied at a concentration of 1U/μL. In a sterile 0.5 mL microcentrifuge tube approximately 3 μg of linearized plasmid DNA (corresponding to ~ 2 pmol of DNA ends) was mixed with 3 μL of 10 × Buffer (Roche, Dephosphorylation Buffer at 1 × contains 50 mM Tris-HCl and 5 mM magnesium chloride at pH 8.5), 2 μL of shrimp alkaline phosphatase and enough sterile distilled water to yield a final reaction volume of 30 μL. Samples were incubated at 37°C for 20 minutes. The phosphatase was then heat inactivated by heating of the sample at 65°C for 30 minutes.

3.3.4.8 Annealing of DNA Oligonucleotides

The DNA oligonucleotides that were synthesized for the construction of peptide-encoding pET31b(+)® plasmids were received as dried pellets so they needed to be dissolved prior to use. The number of OD_{260} units of DNA in each vial was indicated on the label. This information was used to calculate the number of moles of DNA in each tube. The oligonucleotides were dissolved

in TE Buffer (10 mM Tris-HCl 0.1 mM EDTA pH 8.0) in the tubes that they were received in by adding a volume of buffer that would yield a final DNA concentration of 200 pmol/μL. After buffer was added the tubes were incubated at room temperature with periodic vortex mixing for ~ 3.5 hours. The dissolved oligonucleotides were stored at 4°C.

Since the DNA oligonucleotides were single stranded, it was necessary to anneal them prior to ligation into the pET31b(+)® plasmid. The oligonucleotides were annealed in duplicate by mixing 10 μL (~ 1500 - 2000 pmol) *each* of the sense and anti-sense DNA oligonucleotides for each peptide sequence, 20 μL of 10 × Buffer (GibcoBRL, React 2, at 1 × contains 40 mM Tris-HCl, 10 mM $MgCl_2$ and 50 mM NaCl at pH 8.0) and 160 μL of sterile distilled water in a sterile 0.4 mL microcentrifuge tube. The tubes were vortex mixed and centrifuged at 5000 × g for 1 minute to ensure that all of the liquid was at the bottom of the tube. The oligonucleotides were annealed by heating to 99°C and holding for 10 minutes, followed by gentle cooling to 30°C at a rate of ~ 4°C per minute in a PCR machine (Eppendorf, Mastercycler Gradient PCR Machine). The annealed oligonucleotides for each peptide sequence were pooled and then precipitated following the procedure outlined in Section 3.3.4.3. They were re-dissolved in 20 μL of TE Buffer to give a final concentration of 100 pmol/μL.

3.3.4.9 Preparation of Oligonucleotide Tandem Repeats

Tandem repeats of the peptide-encoding oligonucleotides were prepared for ligation into the pET31b(+)® expression vector. Since the T4 DNA ligase was supplied at concentration that was greater than needed for ligation, a dilute stock solution was prepared by mixing 3.35 μL of enzyme (New England Biolabs, Cat. No. 202S, 20,000U) supplied at a concentration of 400U/μL with 3 μL of 10 × Buffer (New England Biolabs, T4 DNA Ligase Buffer, at 1 × contains 50 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP and 25 μg/mL BSA at pH 7.5) and adding 23.65 μL of sterile distilled water in a sterile 0.5 mL microcentrifuge tube. In a sterile 0.5 mL microcentrifuge tube 2.5 μL (~ 250 pmol) of phosphorylated annealed oligonucleotides were mixed with 2.5 μL of 10 × T4 DNA Ligase Buffer, 3 μL of the diluted stock solution of T4 DNA ligase, and enough sterile distilled water to yield a final reaction volume of 25 μL. The samples were vortex mixed and incubated at 16°C for 1, 2 and 3 hours to determine which conditions gave the best yields of tandem repeats. Following ligation, the oligonucleotides were run out on a 2.0 % agarose gel as described in Section 3.3.4.5. DNA bands were visualized using 305 nm UV light and were cut out using a low intensity 360 nm UV light box to avoid nicking of the DNA. The DNA was recovered from the agarose gel slices as described in 3.3.4.6.

3.3.4.10 Preparation of Peptide-Encoding pET31b(+)® Plasmid

The concentrations of plasmid and peptide-encoding DNA samples were estimated from the intensity of the bands observed on agarose gels. Lambda DNA digested with the restriction enzyme *HindIII* was used as a standard since it has fragments of varying sizes of known concentration. In a sterile 0.5 mL microcentrifuge tube ~ 0.026 pmol of *AlwNI* digested, dephosphorylated pET31b(+)® plasmid DNA was mixed with 0.13 - 0.26 pmol of tandem repeats, 3.0 µL of 10 × T4 DNA Ligase Buffer, 3 µL of the diluted stock solution of T4 DNA ligase, and enough sterile distilled water to yield a final reaction volume of 30 µL. The samples were vortex mixed and incubated at 16°C for 19 hours. Following ligation, the enzyme was heat inactivated by incubating the reaction mixture at 65°C for 30 minutes. The ligation mixture was used for transformation of NovaBlue™ *E. coli* cells following the procedure described in Section 3.3.4.1.

3.3.4.11 PCR Screening for Successful Transformants

PCR was used to screen for bacterial colonies containing plasmids where tandem copies of peptide-encoding DNA had been successfully incorporated. Single bacterial colonies that had grown on the LB agar plates following transformation were picked with a flame-sterilized innoculating loop and used to innoculate 2 mL aliquots of LB broth containing 50 µg/mL ampicillin and 34 µg/mL chloramphenicol in sterile 17 × 100 mm polystyrene culture tubes (Simport, Cat. No. T406-2). The cultures were incubated at 37°C overnight (~16 hours) while shaking at 250 rpm. The next day 1 µL of the culture was removed and mixed with 9 µL of sterile distilled water in a sterile microcentrifuge tube. The solution was boiled for 15 minutes and placed on ice.

The primers that were used for PCR screening were a primer to the KSI fusion protein with a 5' - 3' sequence of GGCAAGGTGGTGAGCATC, and a primer to the T7 Terminator with a 5' - 3' sequence of GCTAGTTATTGCTCAGCGG. These primers were commercially available from Novagen, however it was more economical to order them from Life Technologies GibcoBRL. Approximately 275 µg of the KSI, and ~ 298 µg of the T7 Terminator primers were received as desalted lyophilized pellets. Primers were dissolved in 1 mL of sterile distilled water overnight at 4°C. A 'Master Mix' of the reactants required for the PCR reaction was prepared and added directly to 1 µL samples of boiled bacterial culture. The 'Master Mix' was prepared so that the equivalent of 2.5 µL of 10 × Buffer (Amersham Pharmacia Biotech, PCR Buffer, at 1 × contains 10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$ and 0.1 % Triton X-100 at pH 8.6), 4 µL of 1.25 mM dNTP mix (Amersham Pharmacia Biotech, Ultrapure dNTP Set, Cat. No. 27-2035-01,

25 μmol), 1 μL of 10 μM T7 Terminator primer, 1 μL of 10 μM KSI primer and 0.2 μL Taq polymerase (Amersham Pharmacia Biotech, Cat. No. T0303Y, 50 units) would be added to each sample diluted to a final volume of 25 μL. A positive control consisting of cells containing the pET31b(+)® plasmid and a negative control of sterile distilled water were always included with each PCR run.

Samples were vortex mixed and then briefly centrifuged to ensure that all of the liquid was at the bottom of the tube before 30 μL of paraffin oil was added to the samples to prevent evaporation during the thermal cycling. The sample tubes were put in a GTC-2 Genetic Thermal Cycler, GL Applied Research that was pre-heated to 80°C (called a 'hot start'). The 'hot start' was done to prevent extension from any primers that might have annealed at an incorrect spot on the vector. The PCR cycle involved initial denaturing of the DNA at 94°C for 3:00 minutes, followed by 30 cycles of: denaturing at 94°C for 30 seconds, annealing at 55°C for 30 seconds and extending at 72°C for 30 seconds, and a finishing step of heating at 72°C for 5 minutes to make sure all templates had been extended. The PCR products (5 μL of the reaction mixture) were run out on 2 % agarose gels (as described in Section 3.3.4.5) to identify insert-containing plasmids.

Cultures that were identified as containing plasmids with peptide-encoding inserts were used to innoculate fresh 2 mL cultures in duplicate that were incubated at 37°C overnight (~16 hours) while shaking at 250 rpm. The following day plasmid DNA from these cultures was isolated as outlined in Section 3.3.4.2, and glycerol freezer stocks of the cultures were prepared (see Section 3.3.4.12). The cultures were also streaked onto pre-warmed LB agar plates containing the appropriate antibiotics. Plates were placed in a 37°C incubator overnight. The following day the plates were removed from the 37°C incubator and sealed with laboratory film (or placed in a sealed plastic), before being stored at 4°C. Colonies from these plates could used to innoculate fresh cultures for up to a month.

3.3.4.12 Preparation of Glycerol Freezer Stocks

Freezer stocks of bacterial cultures were prepared by taking 900 μL of an actively growing culture (OD_{600} of 0.6 - 0.8), placing it in a 1.5 mL sterile screw cap vial (Nalge Company, Cryogenic Vials, Cat. No. 5000-1020), and adding 100 μL of a sterile 80% glycerol:20% LB broth solution using sterile technique. The tubes were vortex mixed and frozen at -70°C.

3.3.4.13 DNA Sequencing

The sequences of the inserts were verified prior to the transformation of expression host cells for the overexpression and purification of labelled peptides. Sequencing was conducted using the

same primers used for PCR screening (see Section 3.3.4.11). Approximately 0.65 μg of purified plasmid DNA was mixed with 3.2 pmol of the KSI primer (only one primer is needed for sequencing) in a total volume of 12 μL of sterile distilled water. The samples were sent to the University Core DNA & Protein Services DNA Sequencing Laboratory at the University of Calgary for sequencing.

3.3.5 Recombinant Peptide Expression

The expression host, *E. coli* strain BLR(DE3)pLysS, was transformed with sequence verified plasmid following the procedure outlined in Section 3.3.4.1. The presence of insert-containing plasmid in transformed colonies was verified using PCR (see Section 3.3.4.11). Cells from these colonies were used to innoculate cultures for the overexpression of fusion protein.

3.3.5.1 Induction of Fusion Protein Expression

Overnight cultures of plasmid containing cells were diluted 40-fold into fresh media containing 50 μg/mL ampicillin and 34 μg/mL chloramphenicol. Cultures were grown in sterile glass erlenmyer flasks containing a volume of media equal to ~ 20 % of the volume of the flask. The cultures were grown until they reached an OD_{600} of between 0.3 - 0.5 (approximately 2 - 3 hours) before fusion protein production was induced by the addition of sterile Isopropyl-β-D-thiogalactopyranoside (IPTG) (MW 238.31, GibcoBRL, UltraPURE grade, Cat. No. 15529-019, 1 g) to a final concentration of 1 mM. Cells were harvested when the culture reached an OD_{600} of ~ 2.0 (approximately 3 - 6 hours after induction). The cultures were placed on ice for 5 minutes before being centrifuged at 5,000 × g (Eppendorf, Centrifuge 5804, F35-6-38 rotor for 6 x 85 mL tubes) for 5 minutes in a rotor pre-chilled to 4°C. The supernatant was carefully poured off and the cell pellets were resuspended in 0.25 × the culture volume of pH 7.4 phosphate buffered saline (PBS) solution (contains 8 g NaCl, 0.2 g KCl, 1.44 g Na_2PO_4, 0.24 g KH_2PO_4 per 1 L). The cells were spun a second time at 5,000 × g for 5 minutes, the supernatant was carefully removed and the cell pellet was stored frozen at -80°C.

3.3.5.2 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) of Proteins

The SDS-PAGE (Laemmli) Buffer System was used for running polyacrylamide gels (Laemmli, 1970). Gels were cast and run using the Mini-Protean® 3 Cell gel apparatus (Biorad, Cat. No. 165-3301). Once the gels and samples had been prepared the gel-running apparatus was assembled and 1 × Electrode Running Buffer (prepared from a 10 × stock solution containing 30.3 g Tris Base, 144.0 g Glycine, 10.0 g SDS per 1 L) was added. A Prestained Protein Marker (New England Biolabs, Cat. No. 7708S, 1050 μL) containing proteins ranging in size from 6.5 - 175 Da was run on all gels to so protein migration could be monitored while the gel was

being run. An unstained Protein Marker (New England Biolabs, Cat. No. P7702S, 1125 μL) containing proteins ranging in size from 2.3 - 212 Da was also run on all protein gels in order to facilitate accurate estimation of protein sizes. Protein markers were prepared for running on gels in the same manner as the protein samples were. Samples and standards were loaded and the gels which were for ~ 1 hour at 200V.

Gel Casting

The gels that were run had a 5 % stacking gel and a 15 % resolving gel. The resolving gel was prepared by mixing 3.65 mL of distilled water, 3.75 mL of a 40 % acrylamide/Bis solution, 2.5 mL of 1.5 M Tris-HCl pH 8.8 and 0.1 mL of a 10 % SDS solution in a 125 mL side arm erlenmyer flask. The solution was degassed for 15 minutes. Immediately before pouring the gel, 50 μL of a 10 % ammonium persulfate (APS) solution (prepared fresh daily) and 5 μL of N,N,N',N'-tetramethylethylenediamine (TEMED) were added to the solution. The flask was gently swirled to mix the solutions without introducing bubbles. The solution was then carefully poured between two glass plates and *t*-amyl alcohol was overlaid above the resolving gel solution. After 45 - 60 minutes the gel had polymerized and the *t*-amyl alcohol was poured off, and the gel was rinsed well with distilled water and blotted dry with filter paper. The stacking gel was prepared by mixing 6.125 mL of distilled water, 1.275 mL of a 40 % acrylamide/Bis solution, 2.5 mL of 0.5 M Tris-HCl pH 6.8 and 0.1 mL of a 10 % SDS solution in a 125 mL side arm erlenmyer flask. The solution was degassed for 15 minutes. Immediately before pouring the gel, 50 μL of a 10 % APS solution and 5 μL of TEMED were added to the solution. The flask was gently swirled to mix the solutions before it was carefully poured above the resolving gel. The comb was carefully inserted and the gel was allowed to polymerize for 30 - 40 minutes.

Sample Preparation

Protein gels were run to monitor the induction of fusion protein production in cells. Samples were prepared by resuspending frozen cell pellets in 0.1 × the culture volume of cold PBS. The cells were disrupted by sonicating in an ice-water bath until the samples were no longer viscous. Aliquots of these samples were run out on gels to show the total cellular protein (TCP) present in the cells. Prior to sample preparation, the amount of sample to be run on the protein gels was standardized for the OD_{600} of the culture to allow for relative comparisons of protein expression levels. Samples were prepared for running on protein gels by adding 2 × Protein Gel Sample Buffer (80 mM Tris-HCl pH 6.8, 2 % SDS, 0.006 % bromophenol blue and 15 % glycerol to which 100 mM DTT is added *immediately* prior to use), vortex mixing and placing the samples in a 90°C water bath for 5 minutes. Samples were placed on ice for a minimum of 5 minutes prior

to being centrifuged (to ensure that the complete sample volume was at the bottom of the tube) and were then loaded onto the gel. If samples were not going to be analyzed immediately they were frozen at -20°C until needed.

Coomassie Staining

The protein gels were stained in Coomassie Stain (500 mL H_2O, 400 mL methanol, 100 mL glacial acetic acid, 1 g Coomassie Brilliant Blue per 1 L) for ~ 30 minutes while gently shaking. Gels were then destained in Destain (500 mL H_2O, 400 mL methanol, 100 mL glacial acetic acid per 1 L) for ~ 60 minutes while gently shaking. Gels could also be destained overnight by using diluted (1:5) Destain. Gels were rinsed well with deionized water and air dried using a Gel Drying Kit (Promega, Cat. No. V7120) following the protocol provided with the kit.

3.3.6 Labelled Peptide Production

3.3.6.1 Minimal Media

Initially cells transferred from LB broth cultures were cultured on minimal media for 3 - 4 days before the cultures started to grow. For this reason cells were *always* grown on minimal media, which included passaging of the cells on minimal media agar plates. These plates contained 0.4 % glucose and cells grew more slowly than on LB agar plates. Plates were usually grown for 2 - 3 days 37°C. Cultures were grown at 37°C shaking at 250 rpm.

3.3.6.2 Optimization of Protein Overexpression

Protein overexpression was optimized prior to the production of labelled peptides. Some of the factors that were investigated were the: culture density at the time of induction, concentration of IPTG used for induction, media composition, temperature of incubation and the culture density at which cells were harvested. The optimal conditions for fusion protein overexpression in minimal media were determined and the composition of minimal media cultures and agar plates are included in Appendix A.

Protein Overexpression Protocol

Cultures for the overexpression of fusion protein were innoculated from fresh overnight seed cultures grown in minimal media. The cells were spun down, the media was decanted and the cell pellets were resuspended in fresh M9 media that did not contain NH_4Cl or glucose. Protein overexpression was conducted using 400 mL cultures grown in sterile 2 L Erlenmyer flasks. Cultures were innoculated with the equivalent of a 1:20 dilution of a seed culture with an OD_{600} = 1.5, followed by incubation at 36°C while shaking at 250 rpm. Fusion protein production was induced by the addition of 1 mM IPTG to the cultures when their OD_{600} reached ~ 0.4. This took approximately 4 hours. The cells were harvested 6 hours after induction by centrifugation at

3000 × g for 30 minutes in a rotor pre-chilled to 4°C. The supernatant was carefully decanted and the centrifuge jars were inverted on paper towels to drain any excess media before the cell pellets were frozen at -80°C

3.3.6.3 Incorporation of ^{15}N and ^{13}C Labels

For peptide labelling, bacterial cultures were grown on minimal media where the only nitrogen source was ^{15}N-labelled ammonium chloride (CIL, 98 %+ Ammonium-^{15}N Chloride, Cat. No. NLM-467, 10 g) and in some cases the only carbon source was ^{13}C-labelled glucose (Isotec, 99 atom % D-Glucose-$^{13}C_6$, Cat. No. 389374, 10 g). Peptides were expressed with uniform labelling with ^{15}N-, or with both ^{15}N and ^{13}C.

3.3.7 Peptide Purification

3.3.7.1 Isolation of Inclusion Bodies

Cell pellets were resuspended in 0.1 × of the original culture volume by sonicating in 1 × Binding Buffer. 20 mL volumes of resuspended cell pellets were disrupted using a French Press that was pre-chilled to 4°C. To ensure thorough disruption of the cells, each sample was passed through the French Press three times. Inclusion bodies were separated from soluble and low molecular weight cellular debris by centrifugation at 3000 × g for 30 minutes in rotor prechilled to 4°C. The supernatant was decanted and the inclusion bodies were stored frozen at -80°C.

3.3.7.2 Nickel-Affinity Column Purification of His-tagged Fusion Protein

Isolated inclusion bodies were solubilized in 6M guanidine hydrochloride (500 mM Tris-HCl pH 7.9). Dissolved inclusion bodies were centrifuged at 12,000 × g for 10 minutes to remove any insoluble debris and were then syringe filtered through a 0.45 μm syringe filter prior to loading onto a gravity flow Ni^{2+}-affinity column (Novagen, Ni-NTA HisBind Resin, Cat. No. 70666-4, 25 mL) for purification. Fractions containing fusion protein were pooled and dialyzed using 12,000 - 14,000 molecular weight cut-off dialysis tubing (Fisher Scientific, Fisherbrand regenerated cellulose, 25 mm, Cat. No. 21-152-16) against 2 × 4L of distilled water. The precipitated protein was pelleted by centrifugation at 12,000 × g for 10 minutes and frozen as a wet pellet at -80°C. Some samples of column purified fusion protein were lyophilized in order to determine a mass of 'crude' fusion protein produced, however it was not possible to completely solublize these samples in 70% formic acid for peptide cleavage with CNBr, so this procedure was not repeated.

3.3.7.3 Cyanogen Bromide Cleavage

Peptides were cleaved from the N-terminal KSI fusion protein and the C-terminal histidine tag by solubilizing the fusion protein in 6 mL of 80 % formic acid in a 50 mL round bottom flask

containing a micro magnetic stir bar. The solution was purged with argon gas for ~ 15 minutes to remove any dissolved oxygen gas prior to the addition of ~ 0.2 g of solid CNBr. The reaction vessel was sealed with a latex stopper and wrapped in aluminum foil. The reaction was stirred overnight at ambient temperature for ~ 20 hours in the fumehood. When the reaction was complete the solution was evaporated to dryness while heating in a waterbath at ~ 30°C. The peptide was resuspended in a minimum volume (2 - 4 mL) of aqueous acetonitrile (60:40:0.1 acetonitrile:water:TFA) and stirred for ~ 1 hour. The solution was centrifuged to pellet the KSI crystals that had formed. The supernatant was removed and filtered prior to injection onto an HPLC column for peptide purification.

3.3.7.4 HPLC Peptide Purification

^{15}N-labelled TMK was HPLC purified on a 9.4 mm × 250 mm Vydac C-8 semi-prep column using an acetonitrile (0.01% TFA):water (0.01% TFA) solvent system. Various flow rates and gradients were used for purification of the peptide samples, however an optimized protocol was never established due to restrictions in the flow rates that could be used due to leaky HPLC fittings.

3.3.7.5 MALDI Mass Spectrometric Peptide Characterization

MALDI-TOF mass spectrometry (PerSeptive Biosystems, Voyager-DE Mass Spectrometer) was used to identify the HPLC fractions containing purified peptide (which had an expected mass of ~ 2600). The matrix that was used for sample preparation was α-cyano-4-hydroxycinnamic acid dissolved in an aqueous solvent mixture (11.9 mg α-cyano-4-hydroxycinnamic acid, 495.4 μL distilled water, 123.9 μL 3 % TFA and 619 μL acetonitrile).

3.4 Results

3.4.1 Preparation of Peptide-Encoding Plasmids

Plasmids for the expression of the TMK peptide as a single or as 2 tandem repeats were successfully prepared, as were plasmids for expression of the WALK peptide as a single copy, or as 3 and 4 tandem repeats, and plasmids for the expression of the M13-LA peptide as a single copy, or as 2 or 7 repeats. It was found that the bacteria transformed with the plasmid encoding 7 copies of the M13-LA peptide in tandem grew *extremely* slowly in culture and on agar plates as compared with the bacteria transformed with the other peptide-encoding plasmids.

3.4.2 Peptide Expression and Purification

Protein gels were run to identify the optimal conditions for protein expression and purification.

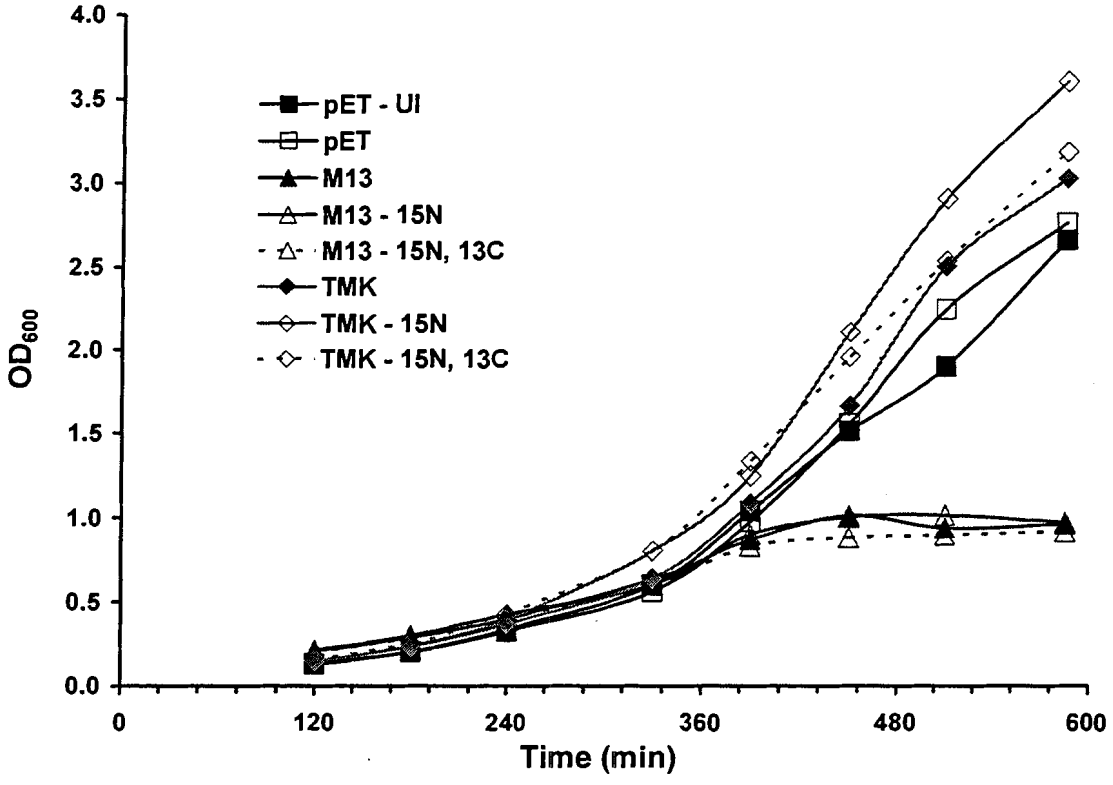


Figure 3.4. **Induction Experiment for the Production of Unlabelled and ^{15}N-, or ^{15}N and ^{13}C-Labelled TMK and M13-LA Peptides Conducted Using 400 mL Cultures Grown in Optimized M9 Minimal Media**

Where *pET* and *pET - UI* denote cultures of cells containing an 'empty' vector (i.e. one that does not contain a peptide-encoding insert). The *pET* culture was used as a positive control, so protein production was induced in the same manner as for the *TMK* and *M13-LA* cultures. The *pET - UI* culture was not induced. It was used as a control for cell growth in the absence of protein overexpression.

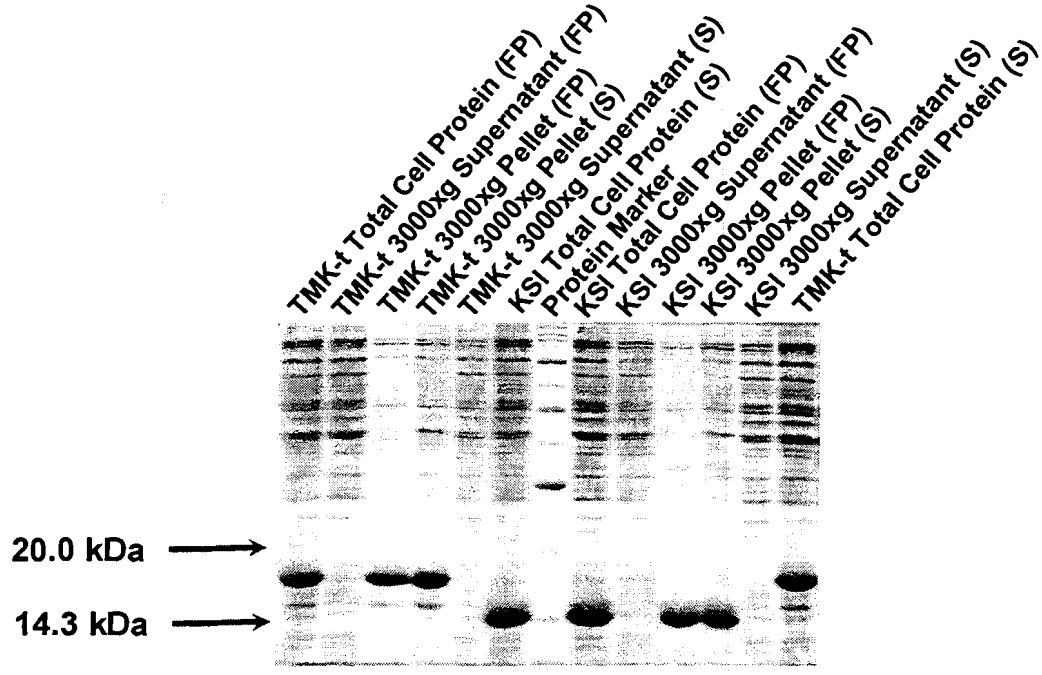


Figure 3.5. Protein Gel Showing the Expression of the TMK Peptide

The lanes labelled 'KSI' contain the carrier protein alone (from the 'positive control' sample), the lanes labelled 'TMK' contain the TMK peptide fused to the carrier protein KSI. 'FP' is used to indicate cells disrupted using the French Press and 'S' is used to indicate cells disrupted using sonication. All of the lanes in this gel were equally loaded based on the OD_{600} of the culture at the time of harvest.

3.4.2.1 Fusion Peptide Overexpression

IPTG was used to induce fusion peptide expression. Growth curves for induced *E. coli* cells expressing the M13-LA and TMK peptides is shown in Figure 3.4. The size of the induced protein was ~ 17 kDa, the expected size for the carrier protein fused to a single peptide. As a positive control a culture of cells containing an 'empty' pET31b(+)® expression vector were also induced. In these cultures a band ~ 14 kDa in size corresponding to His-tagged-KSI was observed. Generally, very high levels of protein expression were achieved which can be clearly seen in the gel shown in Figure 3.5 for the induction of TMK peptide production. It was however found that protein expression levels were dependent on both the nature of the peptide being expressed, as well as on the DNA sequence encoding for the peptide. For cells expressing a fusion protein with one copy of the M13-LA peptide it was found that the amount of fusion protein produced in relation to the total cellular protein was much lower than for the fusion protein alone. It was determined that a mutation had occurred in the plasmid DNA sequence encoding for the peptide and this was causing the lower levels of peptide expression.

3.4.2.2 Peptide Purification

Inclusion Body Isolation

It was found that cells were more thoroughly disrupted using the French Press than by sonication (see Figure 3.5). A lower speed centrifugation (3000 × g) was used to pellet the inclusion bodies in order to minimize the amount of cell membranes that were pelleted with the inclusion bodies. For the two samples of TMK inclusion bodies that had been lyophilized it was possible to estimate yields of crude fusion protein to be 600 mg/L.

HPLC Purification and Characterization

A 1.5 mg quantity of lyophilized ^{15}N-labelled TMK peptide was obtained following HPLC purification and MALDI-TOF mass spectrometric characterization. A sample HPLC trace and the MALDI-TOF spectrum of purified ^{15}N-labelled TMK are shown in Appendix A.

3.5 Discussion

Transmembrane peptides labelled with ^{13}C- and ^{15}N- were successfully produced in *E. coli* strain BLR(DE3)pLysS using the pET31b(+)® expression system. Although it was possible to prepare plasmids containing tandem repeats of peptides, it did not appear that this strategy resulted in increased yields of peptides. When multiple copies of the transmembrane peptides were expressed in tandem with the KSI carrier protein, the overall protein expression level decreased with the increasing number of peptide repeats. So it was not possible to improve peptide yields using this technique. Similar observations have been made by other researchers who have tried to

express tandem copies of hydrophobic peptides and short hydrophobic proteins to improve yields. Even if the peptides could be highly expressed as tandem repeats, the purification of the peptides would likely be difficult because mixtures of products would result from incomplete CNBr cleavage at all of the methionine residues in the fusion protein. It would be difficult to purify the products of incomplete cleavage because the various species would have very similar chemical properties.

In order to obtain optimal expression levels of recombinant protein in *E. coli* strain BLR(DE3)pLysS the cells were grown on minimal media at *all* times. This allowed the cells to become accustomed to growing on minimal media so there was no lag time in growth when the cells were innoculated into the larger volume cultures for labelled peptide production. It was actually found that the levels of protein expression that could be achieved in minimal media were *higher* than the protein expression levels that could be achieved using nutrient media. The protein was also more easily purified from the cells grown in minimal media. The minimal media that was used for growing the cells had increased amounts of phosphate buffer compared with the 'standard' minimal media used to grow *E. coli* cells.

It was found that there was a strong correlation between the hydrophobicity of the peptides and the protein expression levels that could be achieved. The *least* hydrophobic of the three peptides was TMK, and it was the most highly expressed. The *most* hydrophobic of the peptides was the WALK peptide and it was expressed at the lowest levels. The M13-LA peptide had a hydrophobicity that was intermediate between the TMK and WALK peptides, and it was expressed at intermediate levels. An interesting observation was made with one of the M13-LA clones where a mutation had occurred in the plasmid DNA encoding for the peptide. A TTC codon had become a TTT codon. Both encode for phenylalanine, however the TTT codon is the less frequently used of the two codons. When a new clone was prepared which had the optimized sequence for protein expression, increased levels of protein expression were observed.

3.6 Conclusions

High yields of rationally designed ^{15}N-, and $^{13}C,^{15}N$-labelled transmembrane peptides based on both naturally occurring and synthetic peptide sequences can be successfully produced using the pET31b(+)® expression system from Novagen for NMR studies. A recent article has reported the successful use of this expression system for the production of a labelled 81 residue membrane protein MerF for solution NMR studies (Howell *et al.*, 2005). However, the strategy of employing the use of tandem repeats to increase peptide yields does not appear to be effective.

Chapter 4: NMR Studies of Transmembrane Peptides Incorporated into the Monoolein Cubic Phase

4.1 Objectives

To collect multidimensional heteronuclear solution NMR spectra of labelled-transmembrane peptides incorporated into MO:water cubic phases for the purposes of peptide structure determination.

4.2 Introduction

The cubic phases formed by mixtures of MO and water were identified as suitable membrane mimetic environments for solution NMR studies of model transmembrane peptides. The peptides that were designed for study had a hydrophobic length that was tailored to match the hydrophobic thickness of the cubic phase 'bilayer' in order to minimize the chances that the incorporated peptides would alter the phase behaviour of the system. Of the three peptides that were designed, expressed and purified (as described in Chapter 3), one isotopically labelled peptide was obtained in sufficient quantities for NMR studies in the cubic phase. This was the TMK peptide, a 27 residue peptide with a polar tryptophan-containing N-terminus and a charged lysine and serine containing C-terminus. Heteronuclear multidimensional ^{1}H NMR spectra were collected on a sample of cubic phase-bound ^{15}N-labelled TMK. Based on the NMR spectra that were obtained for this peptide, a second labelled peptide was studied in the MO cubic phase. This peptide was the antibiotic peptide alamethicin that is naturally produced by some species of fungus. Solution NMR techniques were used to measure lipid, water and peptide diffusion within peptide-containing MO cubic phases in order to better understand the interactions and behaviour of these transmembrane peptides within the cubic phase.

4.2.1 Alamethicin

Alamethicin belongs to a group of small hydrophobic peptides with antibiotic activity known as peptaibols (Bechinger *et al.*, 2001). These peptides are 20 residues in length and are often produced as a mixture of closely related compounds by the soil fungus *Trichoderma viride* (Yee and O'Neil, 1992). Due to the hydrophobic nature of these peptides, and the absence of any charged residues, they have limited aqueous solubility that is dependent on both temperature and ionic strength (Woolley and Wallace, 1992). The peptides form weakly amphipathic helices that are involved in complex interactions with lipid membranes (He *et al.*, 1996). The structures of the alamethicin peptides have been extensively studied as models of α-helical membrane proteins because they insert into lipid membranes where they form voltage-gated ion channels (Esposito *et al.*, 1987; Woolley and Wallace, 1993; Franklin *et al.*, 1994; Dempsey and Handcock, 1996). Solid state NMR studies of alamethicin in POPC bilayers have shown that the peptide adopts a transmembrane orientation in the absence of an electric field (Bechinger *et al.*, 2001).

4.2.1.1 Interactions of Alamethicin With Lipids

Alamethicin has been shown to affect the properties of the lipid bilayers with which it associates in studies conducted in the *absence* of electric fields. In one study using freeze-fracture EM and X-ray diffraction experiments, it was found that alamethicin altered the morphology of DPPC bilayers (McIntosh *et al.*, 1982). Subsequent studies have shown that the incorporation of alamethicin into phospholipid bilayers can alter the phase behaviour of the system. It has even been shown that alamethicin can induce the formation of cubic phases.

In one study it was shown that as little as 1 % alamethicin could induce the formation of bicontinuous cubic phases in 1,2-dielaidoyl-*sn*-glycero-3-phosphoethanolamine:water (DEPE:H2O) mixtures that would normally form lamellar or inverse hexagonal phases, depending on the temperature of the system (Keller *et al.*, 1996). The samples for these studies were prepared by first dissolving lyophilized alamethicin in methanol, followed by mixing of the dissolved peptide with DEPE dissolved in benzene. The peptide:lipid mixture was frozen and lyophilized and then hydrated to the desired level of hydration by stirring. The cubic phases formed by mixtures of DEPE and alamethicin are formed in the presence of excess water (~ 1:1 wt/wt lipid:water). X-ray diffraction and solid-state ^{31}P NMR studies were used to characterize the phase behaviour of the system. It was found that cubic phases formed over the temperature range of 40°C - 85°C in DEPE:water mixtures containing as low as 1 % (wt/wt) alamethicin to DEPE. In these cubic phases the peptide:lipid ratio was ~ 1:200, if one assumes that alamethicin is uniformly dispersed throughout the lipid. The lattice spacing of the cubic phases did not change over an alamethicin concentration range of 1 % - 10 %, as determined using X-ray

diffraction. The authors postulated that the effect of adding alamethicin was saturated above 1 % suggesting that perhaps phase separation of the lipid and alamethicin was occurring. The cubic phases that were formed were found to be stable for at least 4 months.

The activity and membrane binding of protein kinase C (PKC) has been studied in the bicontinuous cubic phases formed by mixtures of MO, POPS and water, and DEPE, alamethicin and water since this enzyme was known to be activated by nonlamellar forming lipids (Giorgione *et al.*, 1998). It was found that the enzyme was active in both cubic phases where it catalyzed the phosphorylation of histone, even though only a minor fraction of PKC was bound to the 'membrane'. D-type MO:water cubic phases containing up to 10 mol % PS were successfully prepared. The cubic phases were characterized using X-ray diffraction and ^{31}P NMR techniques. Samples were prepared by mixing the lipids in a mixture of chloroform and methanol followed by drying under vacuum. The lipid films were resuspended in the buffer (100 mM KCl, 5 mM $MgCl_2$, 20 mM Tris-HCl at pH 7.0) and subjected to 5 freeze-thaw cycles. The cubic phases were used for activity assays in PKC-containing sucrose buffer in a procedure that involved the use of centrifugation (100,000 × g for 45 minutes at 25°C) to separate membrane-bound enzyme. Diacylglycerol could be added to MO:POPS:water cubic phases up to 1 mol % without altering the cubic phase. It was found that incorporating more than 10 mol % POPS into the MO:water cubic phases altered the phase behaviour of the system at 25°C. At first coexistence of the lamellar and cubic phases were observed at around 15 mol % POPS, and then above about 18 mol % POPS only the lamellar phase was observed. As a 'control' sample for the studies conducted on PKC activation in alamethicin:DEPE:water cubic phases, 4 mol % alamethicin was incorporated into MO:water cubic phases containing 10 mol % POPS. The addition of 4 mol % alamethicin did not disrupt the MO:POPS:water cubic phases.

4.2.2 Protein Structure Determination Using NMR

The application of NMR to the study of protein structure relies on the assignment of the peaks observed in an NMR spectra to specific residues in the polypeptide sequence. In the case of proteins, the 1D ^{1}H spectra are very complex, containing many overlapping peaks. These problems of chemical shift overlap can be resolved by increasing the dimensionality of the spectra and by employing the use of selective filtering of signals from ^{13}C- and/or ^{15}N-labelled protein. The first step in assigning the peaks in an NMR spectra is to identify the residue associated with each of the peaks in the spectrum. This is done through the use of 2D or 3D correlation spectroscopy where the through-bond connectivities (spin systems) of the non-labile (non-exchangeable) protons in the residues comprising the polypeptide chain are identified. The standard nomenclature for atoms in a polypeptide chain are shown in Figure 4.1, and an example

of the spin-system of an amino acid is shown in Figure 4.2. In order to obtain sequence-specific resonance assignments for all of the residues in a polypeptide chain it is necessary to collect nuclear Overhauser enhancement (NOE) data. NOE interactions are observed between two protons if they are located at a distance of less than ~ 5 Å from one another in the protein structure (Wüthrich, 1986). So NOE interactions will be observed between the protons of the *same* amino acid residue (intraresidue NOEs) as well as between protons from *different* amino acid residues (interresidue NOEs). The NOE interactions that are observed between protons of adjacent amino acid residues are used for obtaining sequence-specific resonance assignments for all of the residues in a polypeptide chain. Once this process is completed the 3D structure of a protein can be determined using interproton distance constraints derived from the NOE interactions observed between protons of sequential and in particular, non-sequential amino acid residues. The distance constraints derived from non-sequential NOEs are used to define regions of a protein that are close together in space in the folded conformation of the protein. This information is then used to search conformational space for spatial arrangements of the polypeptide chain that are compatible with the experimental data.

4.2.2.1 Chemical Shifts

It has been recognized for many years that the NMR chemical shift of a nucleus can be a sensitive indicator of the molecular environment of a particular nucleus. Extensive studies of the ^{1}H, ^{13}C and ^{15}N chemical shifts of proteins demonstrated that chemical shifts were very sensitive to protein secondary structure and that trends correlating the dispersion of chemical shifts to protein secondary structure could be identified (Wishart, 1991; Wishart *et al.*, 1992; Wishart *et al.*, 1995). The analysis of chemical shift data makes it possible to determine the location of secondary structural elements in a protein sequence from the ^{1}H, ^{13}C and ^{15}N chemical shift data for the protein. For α-protons, amide protons and amide nitrogens, an upfield shift is observed when these nuclei are in α-helical conformations, whereas a downfield shift is observed when these nuclei are in β-strand, or extended conformations. For α-carbons and carbonyl carbons a downfield shift is observed when these nuclei are located in α-helices, and an upfield shift when they are located in β-strands, or in extended conformations. A chemical shift index (CSI) has been developed to use the information provided by the observed ^{1}H and ^{13}C chemical shifts for a protein to determine what the most likely secondary structure would be (Wishart *et al.*, 1992; Wishart and Sykes, 1994).

4.2.3 Pulsed Field Gradient NMR Measurement of Molecular Diffusion

NMR has been used to study translational diffusion in lipid-water systems since the 1970s (Rilfors *et al.*, 1986). The primary advantage of using pulsed field gradient NMR to measure

molecular diffusion is that measurements can be made directly on the molecules of interest without the need for the use of probes. A pulsed field gradient NMR experiment designed to measure molecular diffusion involves the use of two radio frequency (rf) pulses, one 90° and one 180°, which are used to generate an echo whose amplitude is directly proportional to the lateral diffusion of the molecule of interest (Lindblom and Rilfors, 1989). The time during which diffusion is measured is set using pulsed magnetic field gradients. For a molecule that is freely diffusing within the sample, the amplitude of the echo will decrease with increasing gradient strength. On the other hand, the amplitude of the echo will remain constant for a molecule whose diffusion is restricted irrespective of the gradient strengths used in the experiment.

4.2.3.1 Diffusion Measurements in Lamellar Phases

Measurement of lipid lateral diffusion coefficients in lamellar phases cannot be conducted using standard NMR techniques because narrow peaks are not observed in the spectra due to the anisotropic nature of these phases. In lamellar phases the presence of static dipolar couplings results in an effective relaxation time (T_2) that is too short for the observation of narrow well resolved peaks in the NMR spectra. It is however possible to study lamellar phases using an NMR method in which the samples are macroscopically oriented at the 'magic' angle (54.7°) in a magnetic field because under these circumstances the static dipolar couplings vanish and sharp well resolved peaks are observed. This technique has been successfully used to measure lipid lateral diffusion coefficients in lamellar phases.

4.2.3.2 Diffusion Measurements in Cubic Phases

Some assumptions are made in the measurement and analysis of diffusion coefficients in the cubic phase. First of all it is assumed that the local lipid diffusion coefficient is not affected by the structure of the cubic phase and is thus equal to the measured lateral diffusion coefficient in lamellar phases. This assumption is corroborated by the observation that the lipid diffusion coefficients for a large number of bicontinuous cubic phases have been measured and were found to be of the same order of magnitude as those obtained for the corresponding lamellar phase. In contrast, the lipid diffusion coefficients measured for closed aggregate cubic phases were found to be one to two orders of magnitude smaller than for the corresponding lamellar phase (Rilfors *et al.*, 1986).

Figure 7.2 from Kurt Wüthrich (1986) *NMR of Proteins and Nucleic Acids* John Wiley & Sons Inc.

Figure 4.1. Diagram Showing Standard Nomenclature for Atoms in a Polypeptide Chain

Figure 8.1 from Kurt Wüthrich (1986) *NMR of Proteins and Nucleic Acids* John Wiley & Sons Inc.

Figure 4.2. Diagram Showing the Identification of Spin Systems in a Polypeptide Chain

When diffusion measurements are made in lamellar phases, the lateral diffusion is measured directly as the translational diffusion of a lipid along the plane of the bilayer. However, when diffusion measurements are made in cubic phases, the measured diffusion coefficient is an average of the local diffusional motion over the entire cubic phase structure. So it is not surprising that the structure of the cubic phase will have an appreciable influence on the measured lipid and water diffusion coefficients. In the case of bicontinuous cubic phases with a local 'bilayer' structure, the local diffusion coefficient will actually be 1.5 times the measured diffusion coefficient (Lindblom and Rilfors, 1989). It has been found that the water diffusion coefficients in the cubic phase are a factor of three to five less than those observed for pure water in solution depending on the temperature and the system studied (Lindblom and Rilfors, 1989).

Monoolein:Water Cubic Phases

For mixtures of MO and water the lipid diffusion coefficients have been measured in the lamellar and the cubic phases as a function of temperature since this system can undergo the lamellar $\rightarrow$ cubic phase transition at constant composition. Water diffusion has also been studied in these cubic phases. In one study pulsed field gradient NMR was used measure the diffusion coefficients for MO:water cubic phases of varying compositions at 25°C and 34°C (Eriksson and Lindblom, 1993). For MO cubic phases with a composition of 30.4 % wt/wt water, diffusion coefficients determined at 25°C were 2.30×10^{-5} cm^2/s for pure water, 5.45×10^{-6} cm^2/s for water in the cubic phase and 1.51×10^{-7} cm^2/s for lipid in the cubic phase, and at 34°C the diffusion coefficients were 2.83×10^{-5} cm^2/s for pure water, 6.76×10^{-6} cm^2/s for water in the cubic phase and 2.22×10^{-7} cm^2/s for lipid in the cubic phase.

4.3 Materials and Methods

4.3.1 NMR Sample Preparation

4.3.1.1 TMK

Uniformly ^{15}N-labelled TMK that had been expressed and purified from a bacterial host (as described previously in Chapter 3) was used for the preparation of a peptide-containing MO cubic phase sample for solution NMR studies. The NMR sample was prepared following the procedure outlined in Section 2.3.3.2. The desired amount of MO was melted in a 55°C waterbath for ~ 30 minutes prior to the addition of 1.5 mg of lyophilized ^{15}N-labelled TMK peptide. The peptide was thoroughly dissolved in the molten lipid prior to the addition of water. The sample had a final peptide concentration of 1.15 mM, and a peptide:lipid molar ratio of 1:1800.

4.3.1.2 Alamethicin

A sample of uniformly ^{15}N- and ^{13}C-labelled alamethicin was obtained from Dr. Joe O'Neil in the Department of Chemistry at the University of Manitoba. This sample contained a mixture of uniformly ^{15}N-, and uniformly ^{15}N- and ^{13}C-labelled alamethicins that had been expressed and purified from the fungus *Trichoderma viride* (Yee and O'Neil, 1992; Yee *et al.*, 1997). The sample was of alamethicin with a primary sequence of Ac-Aib^{1}-Pro^{2}-Aib^{3}-Ala^{4}-Aib^{5}-Ala^{6}-Gln^{7}-Aib^{8}-Val^{9}-Aib^{10}-Gly^{11}-Leu^{12}-Aib^{13}-Pro^{14}-Val^{15}-Aib^{16}-Aib^{17}-Gln^{18}-Gln^{19}-Phol^{20}.

Prior to sending of the sample, Dr. O'Neil collected a ^{1}H-^{15}N HSQC spectrum of the peptide dissolved in methanol at 27°C on a Bruker AMX500 NMR spectrometer equipped with a 5 mm triple resonance probe. Proton and ^{15}N chemical shifts were directly or indirectly referenced relative to methanol. The spectrum is shown in Figure 4.5. Heterogeneity was observed in the ^{1}H-^{15}N HSQC spectrum indicating that de-amidation of some of the glutamine residues of the peptide had occurred. A solution NMR sample of MO cubic phases containing labelled alamethicin was prepared following the procedure outlined in Section 2.3.3.2. The desired amount of MO was melted in a 55°C waterbath for ~ 30 minutes prior to the addition of 4.6 mg of lyophilized ^{15}N- and ^{13}C-labelled alamethicin. The peptide was thoroughly dissolved in the molten lipid prior to the addition of water. This sample contained a final peptide concentration of 7 mM, and the sample had a peptide:lipid molar ratio of 1:270.

4.3.2 Multidimensional Heteronuclear NMR Spectra

NMR experiments were conducted at the University of Alberta in the Department of Biochemistry with the assistance of Dr. Brian Sykes. NMR data were collected on a Varian Unity 600 MHz spectrometer equipped with a 5 mm triple resonance z-pulse field gradient probe. All ^{1}H and ^{15}N chemical shifts were directly or indirectly referenced relative to an external DSS solution in the case of the TMK sample, or to internal DSS in the case of the alamethicin sample. Samples were allowed to equilibrate in the spectrometer for ~ 40 minutes at the experimental temperatures before data were acquired. NMR data were processed using *VNMR version 5.1* software (Varian) on a Sun workstation.

4.3.2.1 TMK

A 2D sensitivity-enhanced gradient ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectrum of the TMK cubic phase sample was acquired for 18 hours at 40°C (Farrow *et al.*, 1994). A 3D ^{15}N-edited total correlation spectroscopy (TOCSY)-HSQC spectrum was acquired for 109 hours at 40°C (Marion *et al.*, 1989a).

4.3.2.2 Alamethicin

Two-dimensional sensitivity-enhanced gradient ^{1}H-^{15}N heteronuclear HSQC spectra of the alamethicin sample were acquired for 21 hours at 25°C, 40°C and 50°C (Farrow *et al.*, 1994).

4.3.3 Diffusion Data Collection

NMR experiments to measure lipid, water and peptide diffusion within the peptide-containing cubic phase samples were conducted by Dr. Mark Okon in the Laboratory for Molecular Biophysics at the University of British Columbia using a Varian Unity 500 MHz spectrometer equipped with a 5 mm triple resonance z-pulse field gradient probe. Lipid diffusion was measured using an adaptation of the water-sLED pulse sequence used to measure protein self-diffusion in aqueous solutions (Altieri *et al.*, 1995). Peptide diffusion was measured using a pulse sequence that employed ^{15}N-editing so that the signals originating from the peptide could be selectively observed. A pulse sequence for measuring ^{15}N-labelled peptide diffusion was derived by Dr. Okon from the one developed to measure lipid diffusion, in combination with additional information in the literature (Kay *et al.*, 1992; Zhang *et al.*, 1994). In the experiments that were conducted to measure peptide diffusion the gradients were used to suppress the water signal. Gradients from 6.4 G/cm to 64.4 G/cm, increasing in increments of 3.05G/cm were used for all of the diffusion experiments. The gradients were calibrated using the standard Varian calibration on water.

4.3.4 Diffusion Data Processing and Analysis

4.3.4.1 Processing of NMR Data

The spectral data (the 'fid' files) were imported into *MestReC* and processed using a Fourier transformation (FT). Prior to application of the FT, the spectral data (free induction decay) was zero filled by a factor of two and multiplied by an exponential function using a line-broadening value close to the digital resolution of the spectrum. Exponential multiplication with a line broadening of 10 Hz was applied to all data sets collected to measure lipid and water diffusion. For data sets collected to measure peptide diffusion, exponential multiplication using a line broadening of 30 Hz, in conjunction with multiplication by a 90° shifted sine squared function, were applied prior to Fourier transformation. The spectra were plotted as stacked plots so that all of the spectra collected during an experiment could be viewed simultaneously. The lipid spectra were phased using the automatic phase correction function in *MestReC* and when necessary, manual adjustments of the phasing were made. In order to obtain reliable values for the diffusion coefficients a 3rd order polynomial baseline correction was applied prior to data analysis. Whenever possible, the spectra were referenced to the water peak.

4.3.4.2 Analysis of NMR Spectra

The module provided with *MestReC* for the analysis of arrayed experiments was used to analyze the processed spectral data so that the lipid, water and peptide diffusion coefficients could be calculated. The following equation was used by *MestReC* for diffusion data analysis:

$$G(Y) = I = I_o \exp\{-[(\gamma G\delta)^2 \times (\Delta - \delta/3) \times D)]\} \quad [4.1]$$

Where: G(Y) = echo amplitude
I = peak intensity
γ = ^{1}H gyromagnetic ratio (26751 rad/Gs)
G = gradient strength (G/cm)
δ = duration of the pulsed field gradient pulse (s)
Δ = time between pulsed field gradient pulses (s)
D = diffusion coefficient (cm^2/s)

The spectral data were analyzed using peak intensity because this gave more reliable and reproducible results than peak integrals. A plot of the peak intensities (normalized by dividing by $(\gamma\delta)^2(\Delta - \delta/3)$) as a function of the square of the gradient strength was expected to yield an exponential decay curve. The exponent of the equation describing this decay curve would correspond to the diffusion coefficient (D) for the peak of interest in cm^2/s. If the natural Log of the peak intensities (normalized by division by $(\gamma\delta)^2(\Delta - \delta/3)$) were plotted as a function of the square of the gradient strength, a linear plot would result. The slope of this line would correspond to the diffusion coefficient for the peak of interest in cm^2/s.

4.4 Results

4.4.1 Cubic Phase Sample Preparation

Stable homogeneous samples of TMK- or alamethicin-containing MO cubic phases could be prepared for solution NMR studies. The formation of homogeneous cubic phases were verified by viewing the samples under crossed polarizers. The samples were completely dark. It was found that the viscosity of the cubic phase samples containing TMK were qualitatively similar to the cubic phase samples of MO prepared in the absence of peptide. However, for the alamethicin samples it was found that the viscosity of the samples was much less than that of 'normal' MO cubic phases. These cubic phases appeared to be more 'fluid', at least macroscopically. The samples were tolerant of the prolonged periods of heating at elevated temperatures (40°C) without significant changes in phase behaviour. It was however necessary to add additional water to the samples following prolonged heating to rehydrate the cubic phase as a small cloudy region started to form at the surface of the cubic phase since the NMR tube insert was not sealed.

4.4.2 Correlation Spectra Collected on TMK

The 2D ^{1}H-^{15}N HSQC correlation spectra of cubic phase-incorporated ^{15}N-labelled TMK is shown in Figure 4.3. Eight backbone amide cross peaks were visible in the spectrum, along with the side chain N1H of tryptophan and the terminal NHs of the lysine side chains. Strips of the 3D ^{15}N-edited TOCSY spectra containing peaks are shown in Figure 4.4. These spectra were used to identify the spin systems associated with each of the amide backbone cross peaks seen in the HSQC spectrum. It was possible to unambiguously identify the amino acids associated with each spin system due to the short sequence of the TMK peptide. It was not however possible to make sequence specific assignments of the residues.

4.4.3 Correlation Spectra Collected on Alamethicin

Two-dimensional ^{1}H-^{15}N correlation spectra were collected on uniformly ^{15}N-, ^{13}C-labelled alamethicin in the MO cubic phase at 25°C, 40°C and 50°C. The spectra collected at 25°C and 40°C are shown in Figure 4.6. No backbone amide peaks were observed in these spectra, only the glutamine side chain NH_2 protons were visible. As the temperature of the experiments was increased, more peaks started to appear in the spectra at intensities that were just above baseline. It appeared that the linewidths of the peaks in the spectra were decreasing with increasing temperature. The exact reason for this was not known. It was however noticed that when the sample was removed from the spectrometer after the 50°C experiment it had undergone a phase change. For this reason the results obtained at 50°C cannot be used for comparison with those collected at 25°C and 40°C.

4.4.4 Lipid, Water and Peptide Diffusion Measurements

The NMR data collected to measure lipid, water and peptide diffusion in MO cubic phases containing the peptides TMK or alamethicin were processed and analyzed. The experimentally determined lipid, water and peptide diffusion coefficients are summarized in Table 4.1.

4.4.4.1 Determination of Lipid and Water Diffusion Coefficients

The spectra collected for the measurement of lipid and water diffusion coefficients contained sharp well resolved peaks that underwent exponential decay as a function of increasing gradient strength. A stacked plot showing the ^{1}H NMR spectra collected on the TMK sample at 20°C to measure lipid and water diffusion coefficients is shown in Figure 4.7. The water peak can be seen near the center of the spectrum at approximately 4.8 ppm. All of the other peaks that were visible in this spectrum originated from MO. Similar spectra were collected on the alamethicin sample at 20°C and 40°C. The spectra collected at 40°C are shown in Figure 4.9. Water diffusion coefficients were calculated for all of the data sets using the water peak intensities from the first

five spectra before the water signal had decayed completely. Figure 4.11 shows the linear plot, and Figure 4.12 the normalized natural Log plot, of the water peak intensity as a function of gradient strength for the data collected on the TMK sample at 20°C. The data from the spectra collected at all twenty gradients strengths were used for the determination of the lipid diffusion coefficients. The lipid diffusion coefficients for *each* experiment were calculated as an average of the diffusion coefficients determined for the lipid peaks at 0.9 ppm, 1.3 ppm and 2.0 ppm because the decay behaviour of these peaks was the most reproducible. The linear, and the normalized natural Log plots of the intensity of the MO peak at 1.3 ppm for the spectra collected on the TMK sample at 20°C are shown in Figure 4.13 and Figure 4.14, respectively. The data collected to measure water and lipid diffusion in alamethicin samples yielded exponential decay curves that were similar to those obtained for the TMK sample (plots not shown).

4.4.4.2 Determination of Peptide Diffusion Coefficients

The level of signal-to-noise in the spectra collected to measure peptide diffusion was extremely low, even after data processing had been optimized. A stacked plot of the spectral data obtained for the TMK peptide at 20°C is shown in Figure 4.8. Overlapped peaks from the amide -NHs would be expected in the range of 8.0 - 8.5 ppm. A very weak broad peak was observed in this region of the spectrum. The intensity of this peak was plotted as a function of the gradient strength in Figure 4.15 and the normalized natural Log of the peak intensity was plotted in Figure 4.16. A stacked plot of the spectral data obtained for the alamethicin peptide at 40°C is shown in Figure 4.10. These spectral data had an even lower signal-to-noise ratio than the spectra collected on the TMK sample. The Glu-NH_2 peaks were expected to have chemical shifts between 6.5 - 7.5 ppm. A broad weak peak was observed at ~ 6.5 ppm. The intensity of this peak was plotted as a function of gradient strength in Figure 4.19 and the normalized natural Log of the peptide peak intensity was plotted as a function of the gradient strength in Figure 4.20. The diffusion coefficients for both of these peptides in the MO cubic phase could be calculated from the data that were collected. As a 'negative control' a data set was taken from a region of the spectrum without any visible peaks (~ 10 ppm) for both peptides. These data were plotted in Figure 4.17 for the TMK sample and in Figure 4.18 for the alamethicin sample to demonstrate the difference between random data points and the experimentally derived data points collected to measure peptide diffusion.

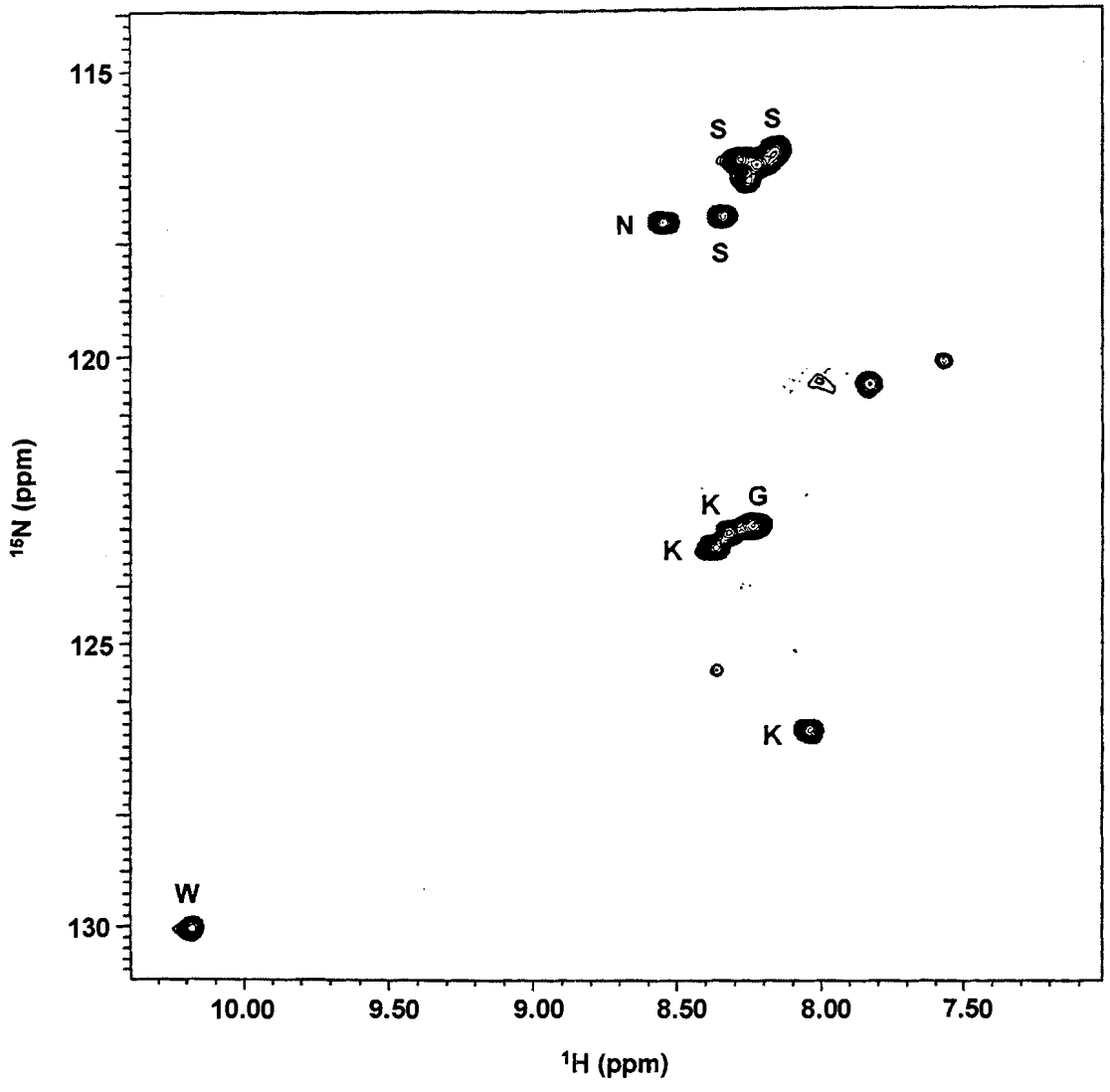


Figure 4.3. **^{1}H-^{15}N HSQC Spectrum of TMK in Monoolein Cubic Phases at 40°C**

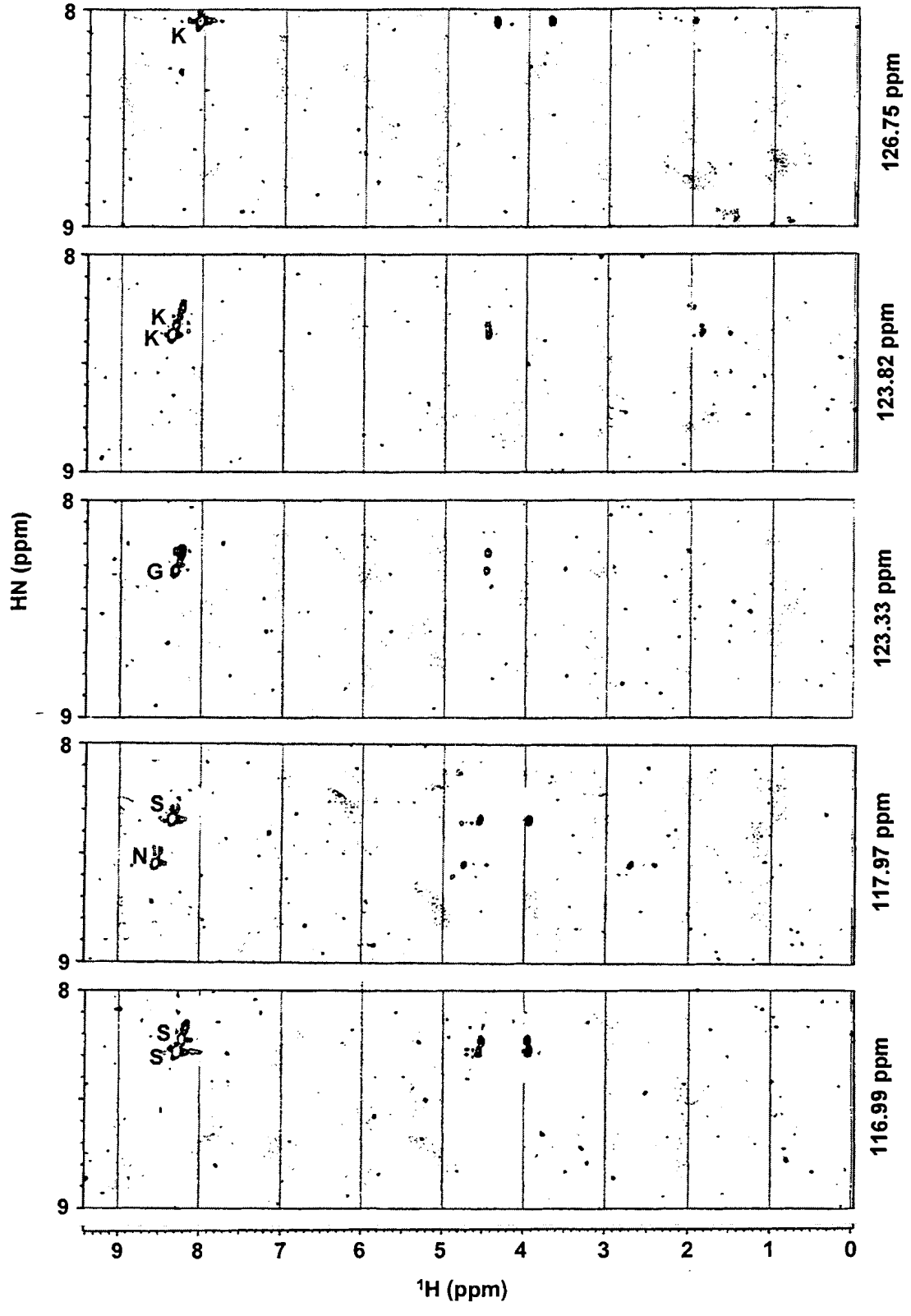


Figure 4.4. Slices Along the ^{15}N Dimension of a TOCSY Spectrum of TMK in Monoolein Cubic Phases at 40°C

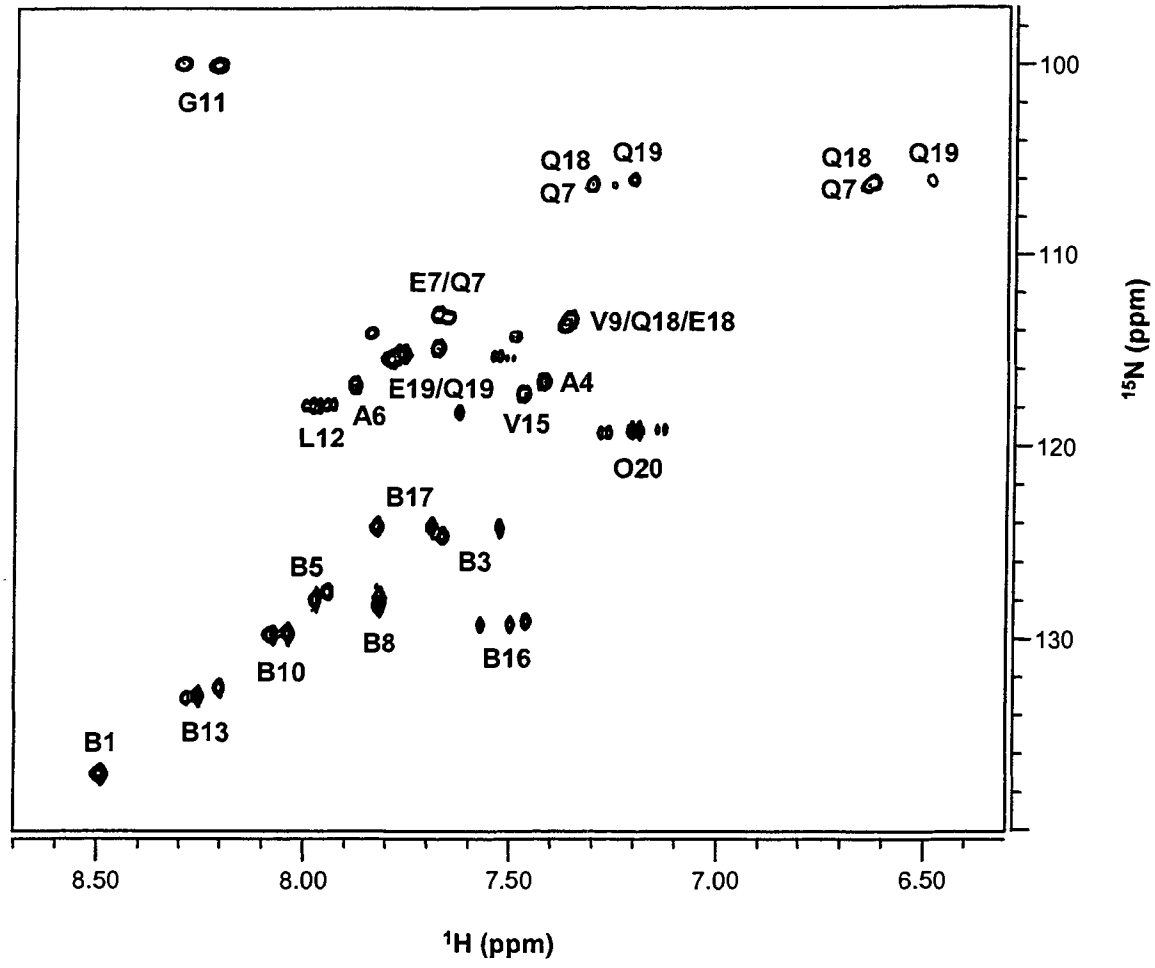


Figure 4.5. ^{1}H-^{15}N HSQC Spectrum of Alamethicin in Methanol at 27°C

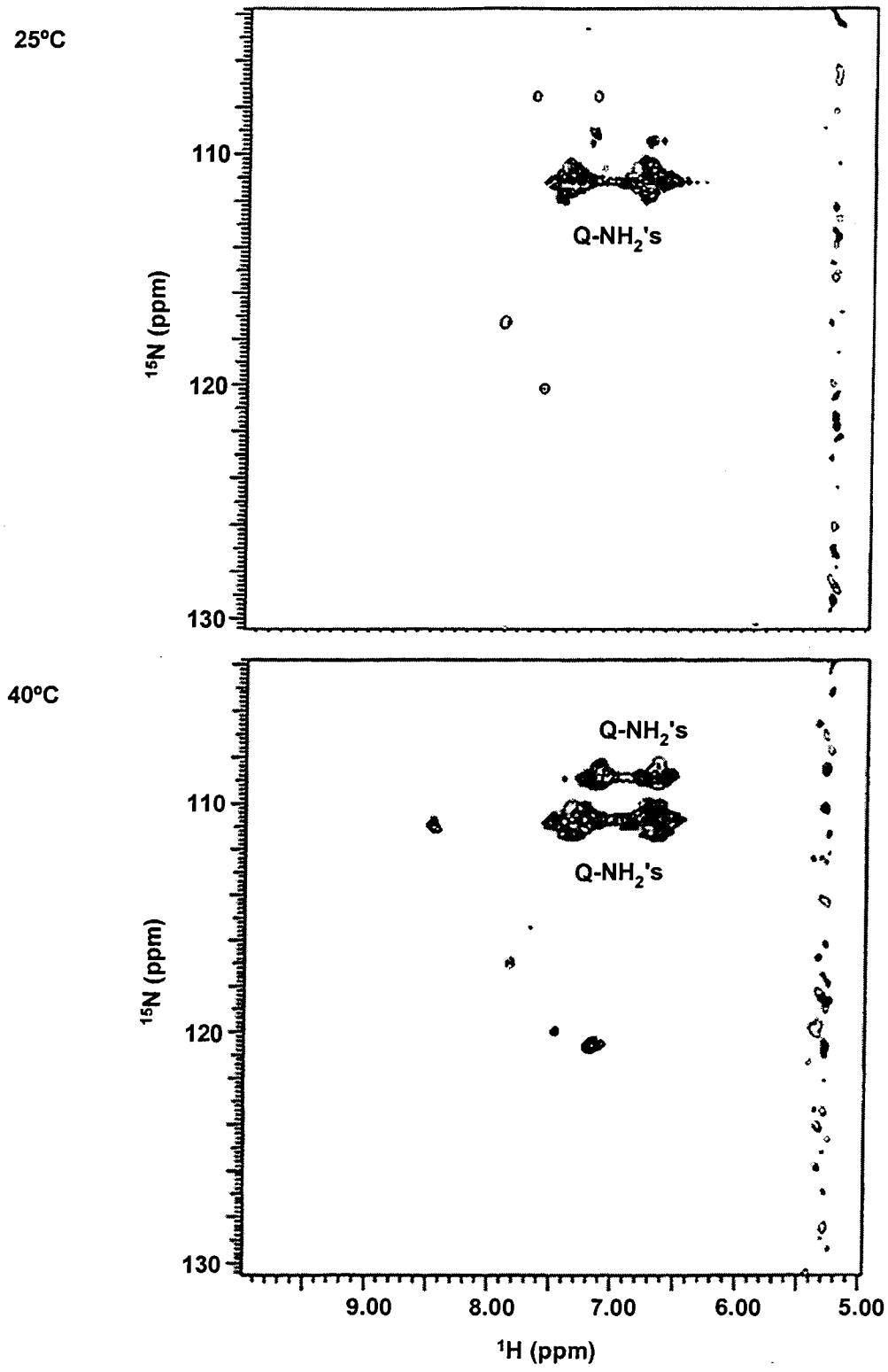


Figure 4.6. ^{1}H-^{15}N HSQC Spectra of Alamethicin in Monoolein Cubic Phases at 25°C and 40°C

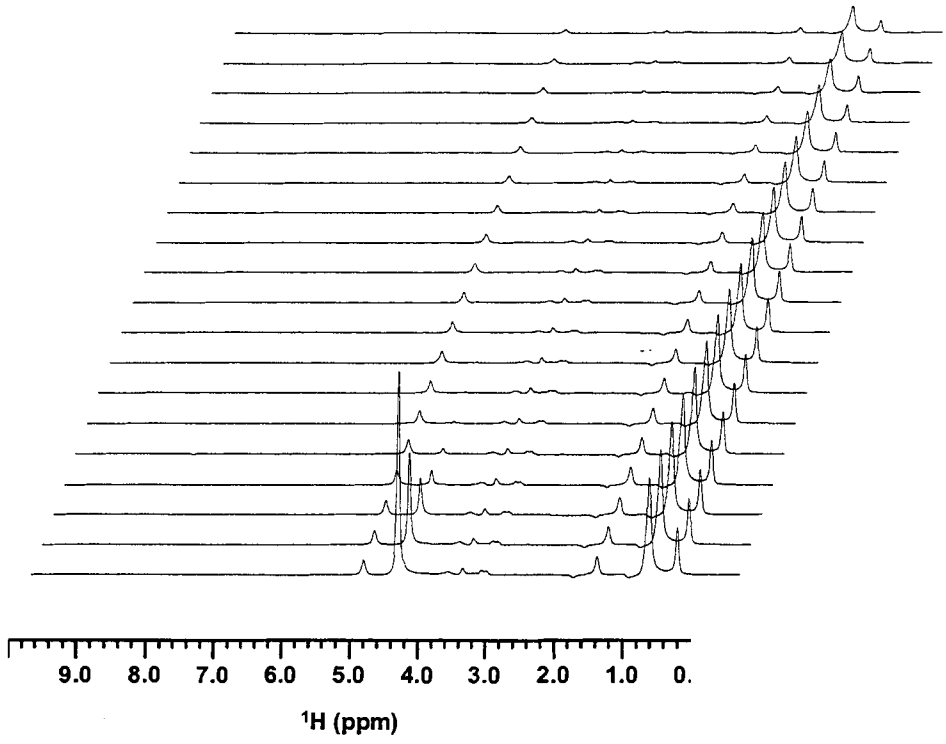


Figure 4.7. **^{1}H NMR Spectra of TMK-Containing Monoolein Cubic Phases Collected to Measure Lipid and Water Diffusion at 20°C**

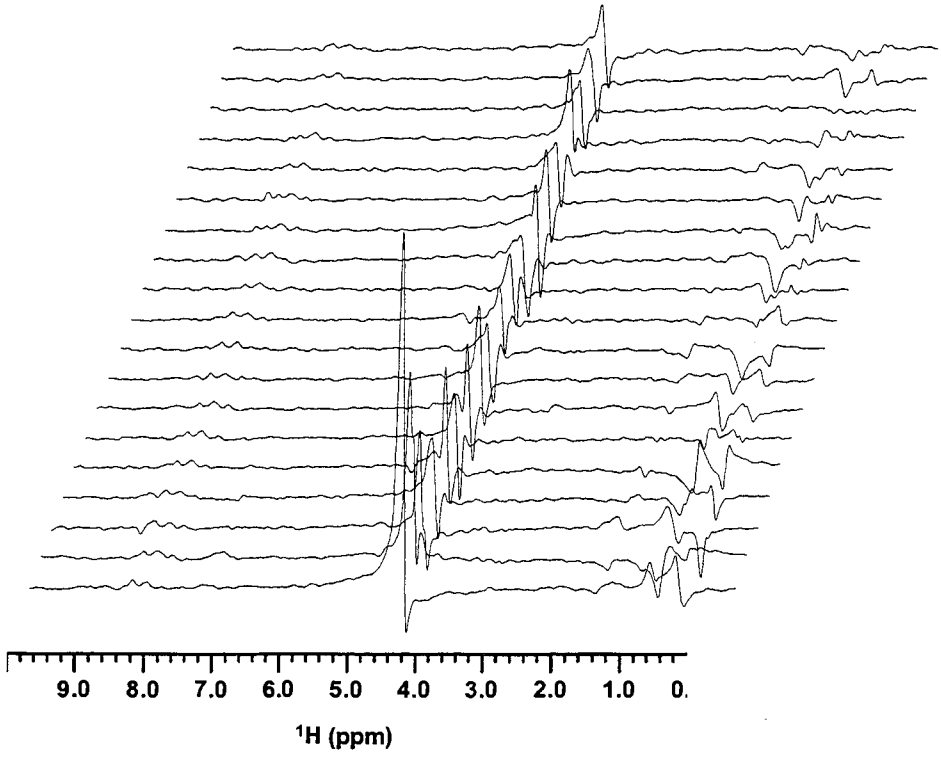


Figure 4.8. **^{15}N-edited ^{1}H NMR Spectra of TMK-Containing Monoolein Cubic Phases Collected to Measure Peptide Diffusion at 20°C**

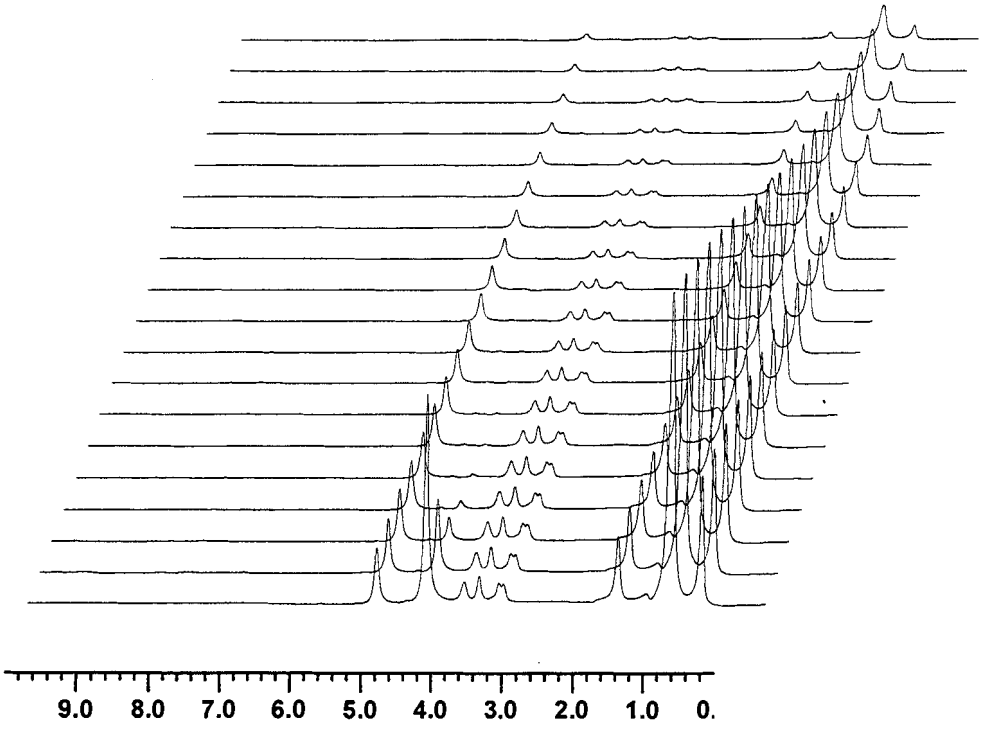


Figure 4.9. **^{1}H NMR Spectra of Alamethicin-Containing Monoolein Cubic Phases Collected to Measure Lipid and Water Diffusion at 40°C**

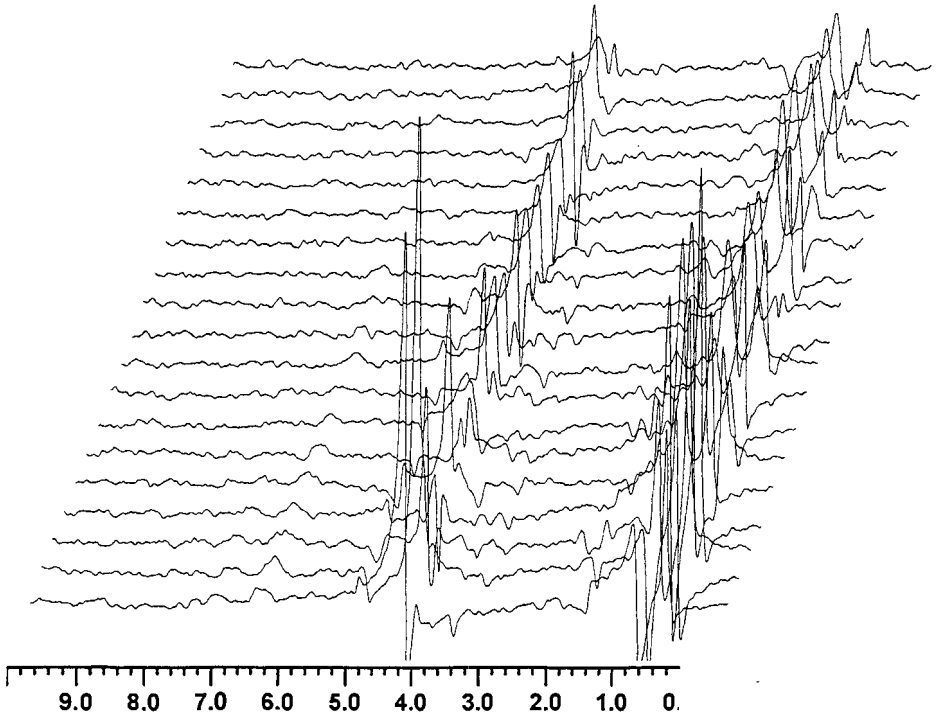


Figure 4.10. **^{15}N-edited ^{1}H NMR Spectra of Alamethicin-Containing Monoolein Cubic Phases Collected to Measure Peptide Diffusion at 40°C**

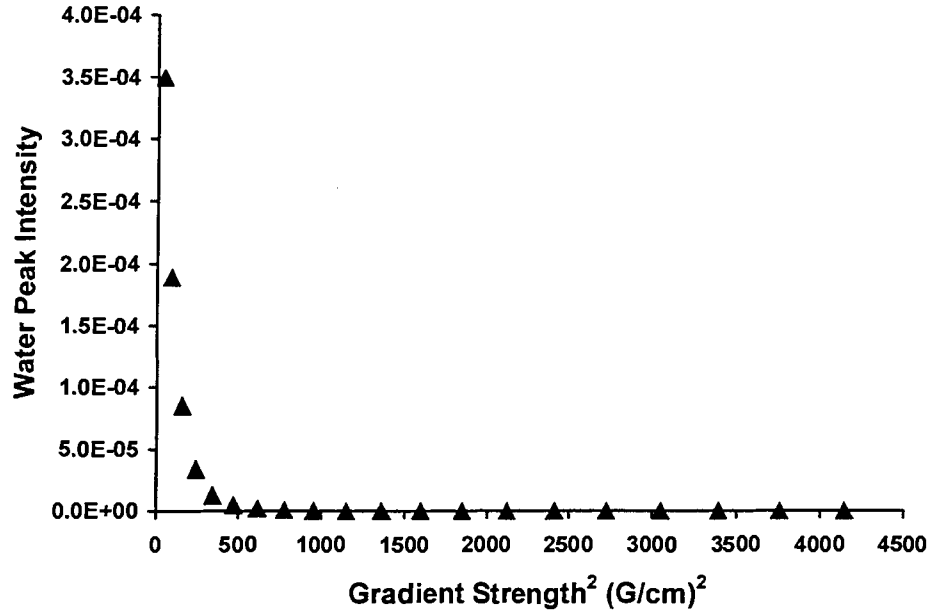


Figure 4.11. **Plot of Water Peak Intensity versus Gradient Strength for the TMK-Containing Monoolein Cubic Phase Sample**

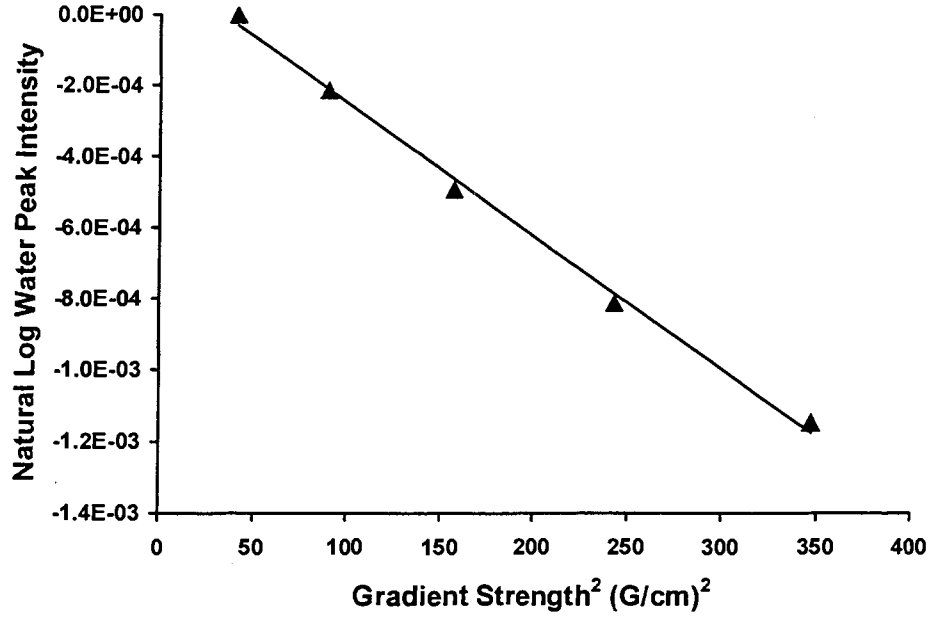


Figure 4.12. **Natural Log Plot of Water Peak Intensity versus Gradient Strength for the TMK-Containing Monoolein Cubic Phase Sample**

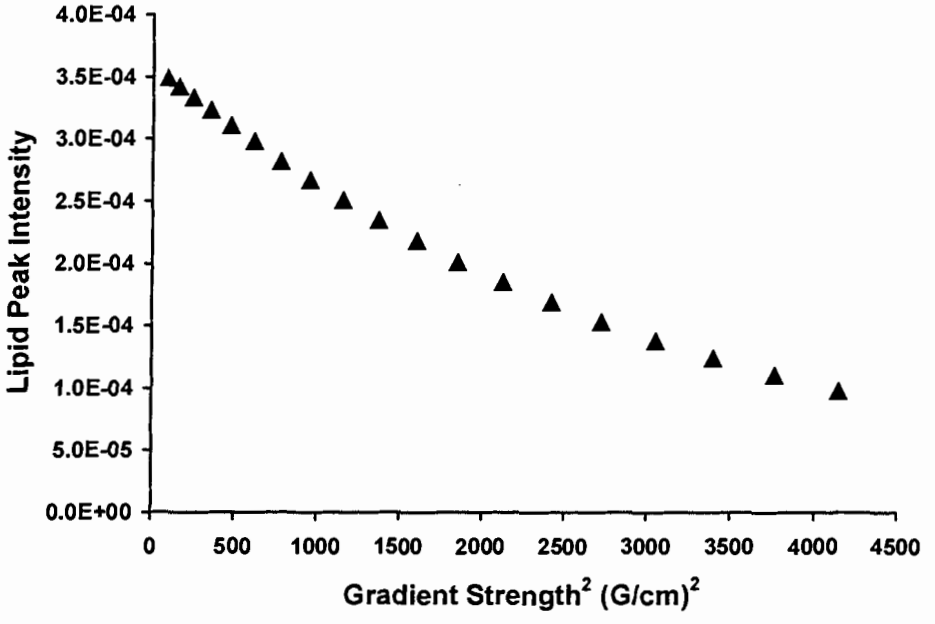


Figure 4.13. **Plot of Lipid Peak Intensity versus Gradient Strength for the TMK-Containing Monoolein Cubic Phase Sample**

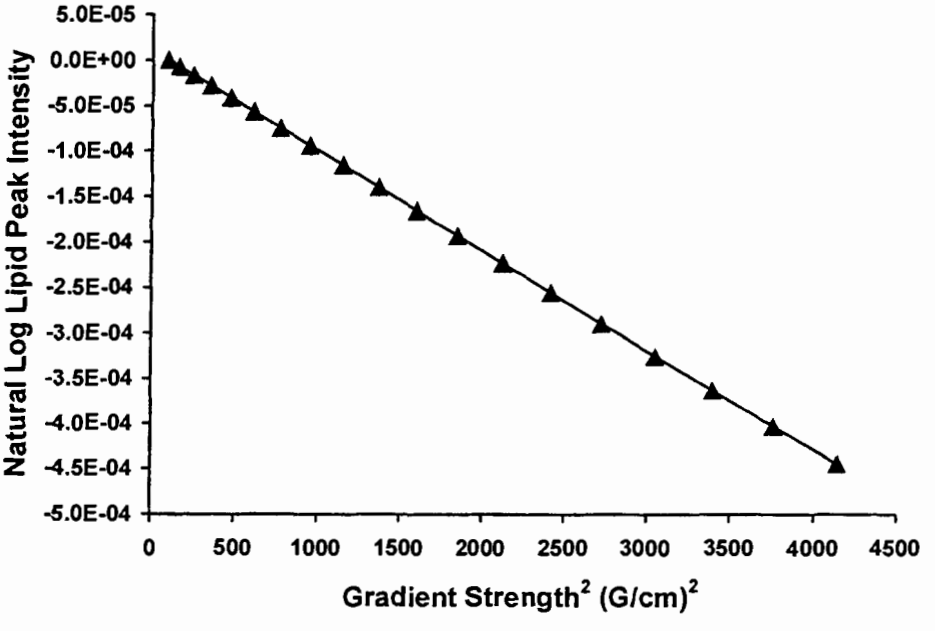


Figure 4.14. **Natural Log Plot of Lipid Peak Intensity versus Gradient Strength for the TMK-Containing Monoolein Cubic Phase Sample**

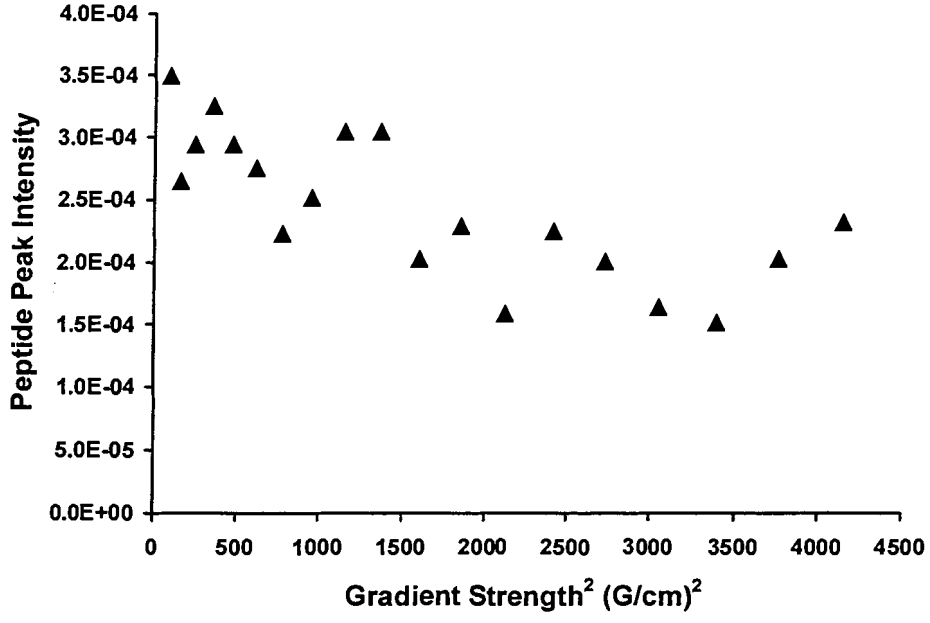


Figure 4.15. Plot of Peptide Peak Intensity versus Gradient Strength for the TMK-Containing Monoolein Cubic Phase Sample

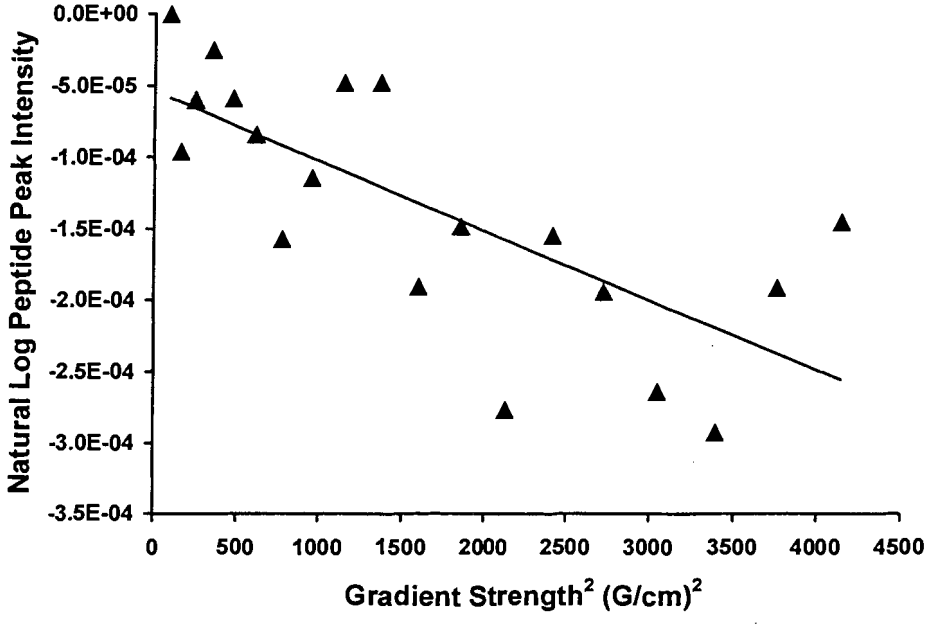


Figure 4.16. Natural Log Plot of Peptide Peak Intensity versus Gradient Strength for the TMK-Containing Monoolein Cubic Phase Sample

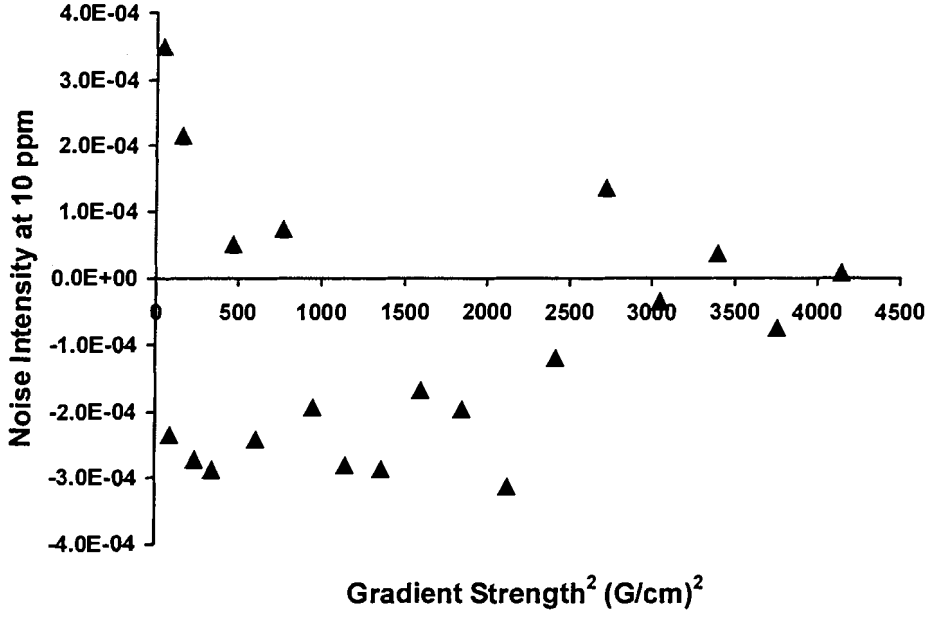


Figure 4.17. Plot of Noise at 10 ppm versus Gradient Strength for the TMK-Containing Monoolein Cubic Phase Sample at 20°C

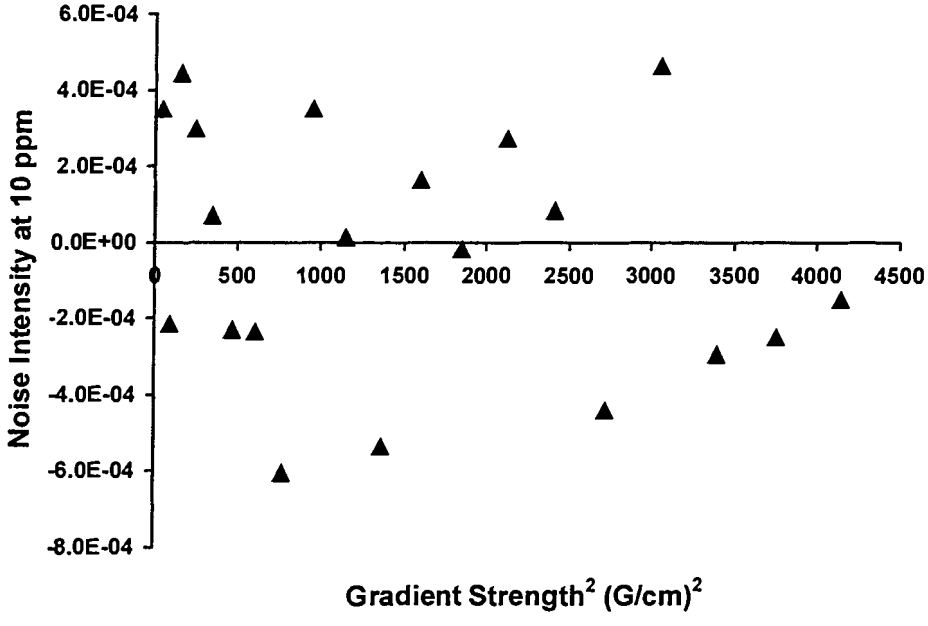


Figure 4.18. Plot of Noise at 10 ppm versus Gradient Strength for the Alamethicin-Containing Monoolein Cubic Phase Sample at 40°C

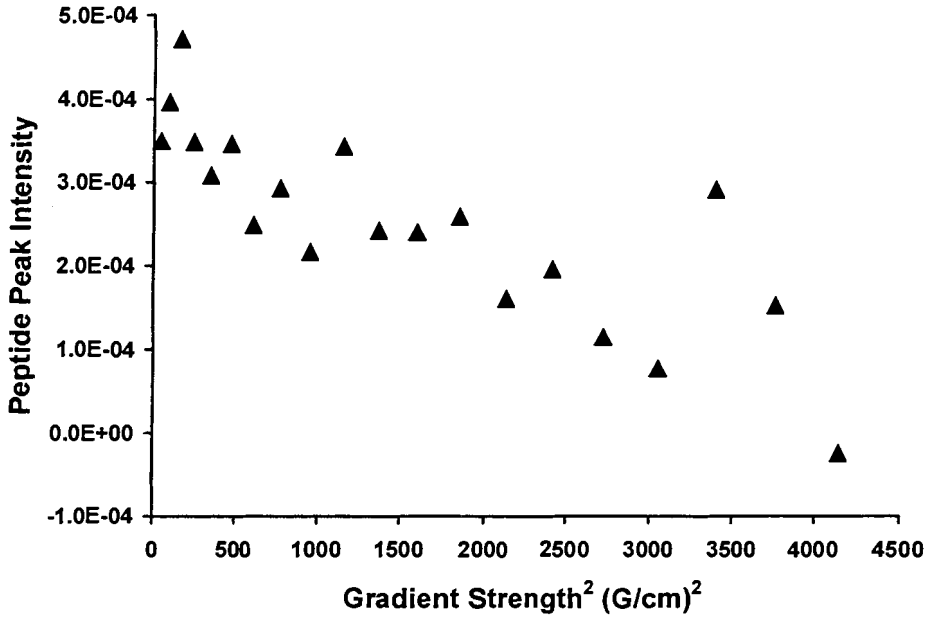


Figure 4.19. Plot of Peptide Peak Intensity versus Gradient Strength for the Alamethicin-Containing Monoolein Cubic Phase Sample

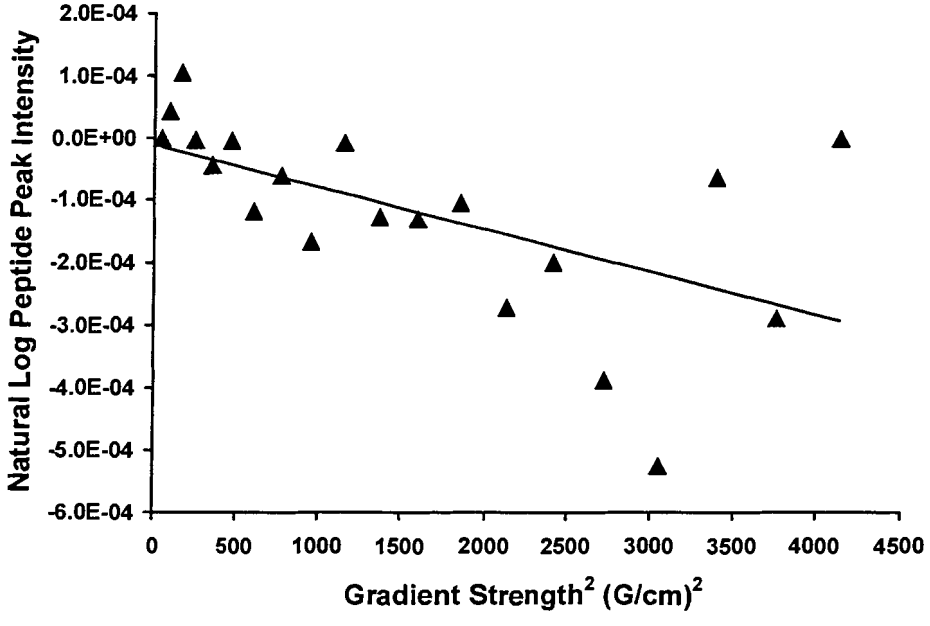


Figure 4.20. Natural Log Plot of Peptide Peak Intensity versus Gradient Strength for the Alamethicin-Containing Monoolein Cubic Phase Sample

Table 4.1. Summary of Diffusion Data Collected on the TMK and Alamethicin Peptides in Monoolein Cubic Phases

Experiment Description	Diffusion Coefficient (cm^2/s)	Temperature (°C)	δ (s)	Δ (s)
Alamethicin Lipid & H_2O Diffusion at 20°C	$D_{LIPID} = 1.0 \times 10^{-7}$ $D_{WATER} = 3.3 \times 10^{-6}$	20	0.002	1
Alamethicin Lipid & H_2O Diffusion at 40°C	$D_{LIPID} = 1.9 \times 10^{-7}$ $D_{WATER} = 4.6 \times 10^{-6}$	40	0.002	1
Alamethicin Peptide Diffusion at 40°C	$D_{PEPTIDE} = 6.8 \times 10^{-8}$	40	0.002	1
TMK Lipid & H_2O Diffusion at 20°C	$D_{LIPID} = 1.1 \times 10^{-7}$ $D_{WATER} = 3.8 \times 10^{-6}$	20	0.002	1
TMK Peptide Diffusion at 20°C	$D_{PEPTIDE} = 5.1 \times 10^{-8}$	20	0.002	1

4.5 Discussion

4.5.1 Cubic Phase Sample Preparation

It appears that the transmembrane peptide TMK can be incorporated into MO cubic phases without perturbing the phase behaviour of the system. Incorporation of alamethicin into the MO cubic phase did not perturb the phase behaviour of the system, however this was not unexpected as previous studies had shown that alamethicin could be incorporated into the MO cubic phase at fairly high concentrations without causing perturbations of the cubic phase (Giorgione *et al.*, 1998). It was noticed that the alamethicin-containing cubic phases were less viscous than the TMK-containing cubic phases. However, no significant differences in the lipid diffusion coefficients were found between these two samples. In addition to qualitative observations of the stiff gel-like nature of the cubic phase samples, the isotropic nature of the cubic phase samples was confirmed by observation of samples under cross polarizers. The samples were completely dark, as was expected.

4.5.2 NMR Correlation Spectra

4.5.2.1 TMK

In the ^{1}H-^{15}N correlation spectra of TMK incorporated into the MO cubic phase one could theoretically see up to 26 peaks corresponding to the backbone amide ^{1}Hs of the peptide. Since the peptide had a highly repetitive sequence, and the predicted secondary structure was an α-helix, it was anticipated that there would be extensive cross-peak overlap and chemical shift degeneracy of the observable peaks. When the spectra were collected, far fewer peaks were observed than was expected. Only 8 peaks were visible in the spectrum and a number of them overlapped. In order to identify which residues were being observed, a ^{15}N-edited TOCSY experiment was conducted to identify the spin systems associated with each of the cross peaks in the HSQC spectrum. It was found that the residues that were observable were located at the N- and C-termini of the peptide. This indicated that the peptide was likely inserted into the MO cubic phases in a transmembrane orientation and that only the residues located in the interfacial region of the cubic phase were visible. A schematic diagram of the proposed location of TMK within the cubic phase based on this data is shown in Figure 4.21. The chemical shifts for the NMR observable residues at the termini of the peptide were not characteristic of α-helical secondary structure indicating that the ends of the peptide are likely 'frayed' (unstructured) (Wishart, 1991). It was not possible to obtain sequence specific assignments for the observed amide resonances because of the weak signal-to-noise in the spectrum (it would not have been

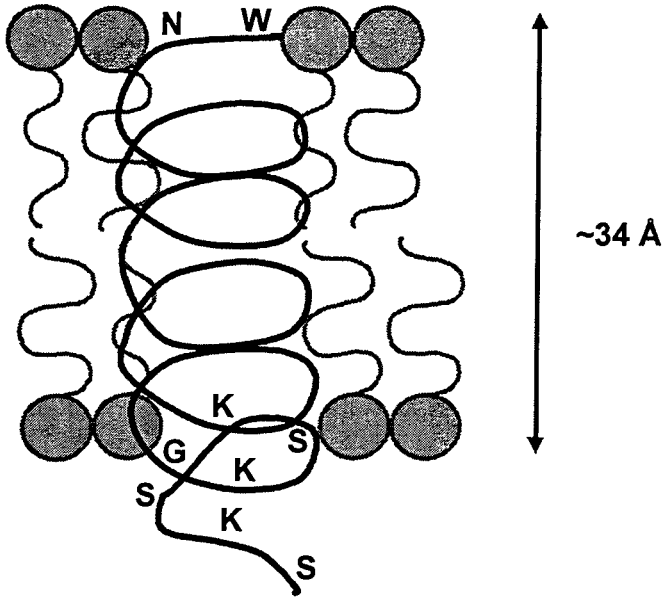


Figure 4.21. **Schematic Diagram of TMK Localization Within the Monoolein Cubic Phase**

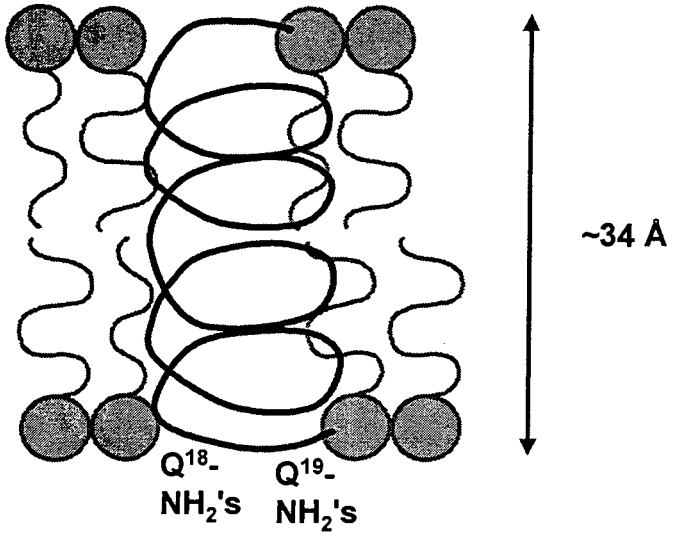


Figure 4.22. **Schematic Diagram of Alamethicin Localization Within the Monoolein Cubic Phase**

feasible to collect a NOESY spectrum), and even if these data were available, the repetitive nature of the peptide would likely have precluded sequence-specific assignment of the peaks.

4.5.2.2 Alamethicin

It was expected that the 18 backbone amide cross peaks observable in the ^{1}H-^{15}N HSQC correlation spectra collected on alamethicin in methanol (shown in Figure 4.5), would also be observable for alamethicin that had been incorporated into the MO cubic phase. However, none of the backbone amide cross peaks were observable. Only the glutamine side chain NH_2s were visible. With increasing temperature the peaks in the spectrum became more sharp and some weak peaks started to become visible just above the baseline threshold.

The identity of the emerging peaks could not be determined. It was however noted that the phase behaviour of the sample had changed during the experiment conducted at 50°C, when the greatest number of emerging peaks were observed. The reason that more peaks became visible at higher temperatures could have been because of increased diffusion rates of the peptides or thinning of the MO 'bilayer' such that more of the peptide protruded from the bilayer giving it a greater range of motion. Although the reasons for the temperature dependent improvements in the alamethicin spectra are not known, these results indicated that the spectra obtained for TMK were not due to an inherent property of the peptide sequence studied, but were rather due to a more generalized phenomenon. Since alamethicin is a few residues shorter than TMK, and it does not form an ideal α-helix, it would not protrude from the 'bilayer' of the MO cubic phase. A schematic diagram of the proposed location of alamethicin within the cubic phase based on the NMR data is shown in Figure 4.22.

4.5.3 Diffusion Data

In the stacked plots of the spectral data collected for measurement of the diffusion coefficients, it can be seen that the signal intensity of the water peak (~ 4.8 ppm) decays to zero much more quickly than the lipid peaks do. The intensity of this peak decreases very quickly as a result of the fast translational diffusion of water within the cubic phase. The intensity of the lipid peaks decreases much more slowly as the lipid translational diffusion coefficient is approximately one order of magnitude less than that of water within the cubic phase. Since all plots of the normalized natural log of the signal intensity versus the square of the gradient strength were linear, there was no influence of intermolecular NOEs on the diffusion results (Chen and Shapiro, 1999; Lucas *et al.*, 2003).

Lipid and water diffusion coefficients were determined for both peptide-containing MO:water cubic phase samples using the NMR data that were collected. It was found that the lipid and water diffusion coefficients determined for TMK and alamethicin at 20°C were very

similar. Unfortunately the TMK diffusion data were only collected at 20°C, so it was not possible to compare results for these transmembrane peptides at two different temperatures. The lipid and water diffusion coefficients obtained for the transmembrane peptide-containing cubic phases agreed well with previously published values of 5.45×10^{-6} cm^2/s for water and 1.51×10^{-7} cm^2/s for lipids at 25°C, and 6.76×10^{-6} cm^2/s for water and 2.22×10^{-7} cm^2/s for lipids at 34°C in cubic phases with a composition of 69.6 % wt/wt MO in water.

The spectral data that were collected for the measurement of peptide diffusion coefficients had very low signal intensity. This was a function of the nature of the experiment, the use of ^{15}N-editing of the spectra and of the very low peptide concentration of the samples. It is possible that acquiring more data sets, or perhaps running the experiment on a higher field spectrometer could be used to improve the quality of the data. Despite the high level of scatter in the data points, it was possible to calculate the diffusion coefficients for the transmembrane peptides within the cubic phase. The values that were determined could probably be considered accurate to an order of magnitude. It was interesting that similar values for peptide and protein diffusion coefficients were determined using fluorescence recovery after photobleaching (FRAP) studies in the MO cubic phase (Tsapis *et al.*, 2001). An 18 residue fluoroscein-tagged peptide with the sequence KKG(L)$_{12}$KKA-amide, and a fluoroscein-tagged protein were incorporated into the MO cubic phase and their diffusion coefficients were determined. The diffusion rates for the peptide and the protein were found to be 6.2×10^{-8} cm^2/s and 1.4×10^{-8} cm^2/s, respectively at 22°C. These values are within the same order of magnitude as the values that were determined for TMK and alamethicin.

4.6 Conclusions

4.6.1 Monoolein Cubic Phase Structure

The NMR data that were collected on the transmembrane peptides TMK and alamethicin in the MO cubic phase indicated that only the interfacial regions of the peptides can be studied using solution NMR techniques in the cubic phase-bound form. It appears that the lipids experience less hindered motion within the cubic phase than do the transmembrane peptides. This is shown in both the 2D heteronuclear correlation spectra and the diffusion data that were collected on the transmembrane peptides. The difference in the motion of the lipids and peptides could simply be a function of their molecular sizes, in that a lipid can rotate along its long axis with greater frequency than a peptide. It is also possible that the motion of a lipid, which only spans *half* of the bilayer, is somewhat less restricted than the motion of a transmembrane peptide which spans the *entire* bilayer and has both ends tethered in independent water channels. In order to fully

answer this question it will be necessary to study a membrane surface-binding peptide or protein in the MO cubic phase.

Chapter 5: NMR Studies of the Membrane Surface Binding Peptide Methionine-Enkephalin in Monoolein Cubic Phases

5.1 Objectives

To collect solution NMR spectra of a membrane surface-binding peptide incorporated into the MO:water cubic phase and, if sufficient NMR data is available, to also determine the structure of the lipid-bound peptide.

5.2 Introduction

Solution NMR studies of transmembrane peptides in MO cubic phases revealed that information could be gathered about the peptide residues located in the interfacial regions of the cubic phase. Based on these results, it was postulated that perhaps solution NMR data could be collected for *all* of the residues in a peptide that bound to the surface of MO:water cubic phases. Depending on the nature of this peptide-lipid interaction, it was believed that it might be possible to determine the structure of the cubic phase-bound peptide using solution NMR data.

The first step in initiating these studies was to identify a structurally well characterized peptide whose interactions with lipids and biological membranes had been thoroughly investigated. In addition to the requirement that the peptide be well characterized, it would also need to be amenable to stable isotope labelling for solution NMR studies. The peptide that was identified as an ideal peptide for study was the five residue membrane-binding peptide methionine-enkephalin. This peptide i) is fully water soluble, ii) becomes structured in the presence of lipids, iii) is small in size, iv) has been extensively characterized in many membrane mimetic environments, v) has been studied using NMR techniques, and vi) is amenable to chemical synthesis in both labelled and unlabelled forms.

5.2.1 Enkephalins

The enkephalins are endogenous opioid peptides with morphine-like activity. They bind to the

opiate receptors which are found in the central and peripheral nervous systems, as well as in the intestinal tract of mammals (Dores *et al.*, 2002). Two of the enkephalins are only five residues in length and differ in sequence by one amino acid. They are methionine-enkephalin with the sequence Tyr-Gly-Gly-Phe-Met, and leucine-enkephalin with the sequence Tyr-Gly-Gly-Phe-Leu (Hughes *et al.*, 1975). These peptides have been shown to interact with the μ, δ, and κ opiate receptors (Milon *et al.*, 1990). One of the first questions that arises when one considers the biological activity of these peptides is: how can such a small flexible peptide ligand specifically interact with its target receptor? It is believed that the membrane in which the target receptors reside play a very important role in this interaction, since it has been found that many hormones and neurotransmitters whose receptors are coupled with G-proteins, interact with phospholipid membranes (Gysin and Schwyzer, 1983; Milon *et al.*, 1990). Despite the number of studies that have been conducted, the mechanism of action of these peptides is still not well understood.

Many of the aqueous-soluble peptides known to interact with membranes do not appear to have any specific conformations in aqueous solution prior to membrane-binding (Motta *et al.*, 1987; Surewicz and Mantsch, 1988; Milon *et al.*, 1990). It has been hypothesized that these 'unstructured' amphiphilic peptides bind to membranes where they adopt conformations, and perhaps the orientation, required for binding to their target receptors (Deber and Behnam, 1984). The process in which a peptide adopts a bioactive conformation prior to receptor binding through electrostatic and hydrophobic interactions with cell membranes has been called 'membrane catalysis' (Marcotte *et al.*, 2003). It is an effective mechanism for a peptide to find its target receptor since the three-dimensional search it was undergoing in solution is reduced to a two-dimensional search upon membrane binding, *and* the peptide is now in a conformation that favours receptor binding.

5.2.1.1 Therapeutic Interest in Enkephalins

The structure and mode of action of the enkephalins has been of interest to researchers since it was first recognized that these peptides had the same receptors as morphine. It was believed that if the relationships between the structure and activity of these peptides could be understood, it would be possible to develop synthetic analgesics for therapeutic use. Initially it was assumed that the enkephalins would take on a conformation analogous to that of morphine because they bound to the same receptor (Milon *et al.*, 1990). Attempts were made to relate the structural and conformational features of the enkephalins to the geometry of morphine (Jarrell *et al.*, 1980). However, this approach was unsuccessful because morphine has a rigid structure and the enkephalins are very flexible peptides whose conformations are not very rigid or well defined.

Some of the structural elements of the enkephalins that are believed to be important for pharmacological activity are the tyrosine amide and the phenolic groups. Evidence has suggested that the orientation of the tyrosine and phenylalanine rings with respect to each other is important for determining receptor subtype selectivity. Recent studies have indicated that a folded conformation in which the tyrosine and phenylalanine aromatic rings are in close proximity to one another may be selective for binding to δ-opiate receptors, whereas a conformation in which the aromatic rings point in different directions may be selective for the binding to μ-opiate receptors (Marcotte *et al.*, 2004). Once the high resolution structure of an opiate receptor has been solved, it might be possible to establish which structural elements of these peptides are important for receptor binding and activation.

In addition to the inherent properties of the enkephalins themselves, it is believed that the interaction of these peptides with cell membranes may also be important in determining the type of opiate receptor (μ, δ, or κ) the peptides interact with *in vivo* (Schwyzer, 1986). This hypothesis is supported by evidence that the conformations adopted by the enkephalins are dependent on the medium in which the peptides are studied. It is possible that variations in the composition of opiate receptor-containing cell membranes *in vivo* may be involved in modulating the receptor selectivity of the enkephalins. For these reasons it was believed that study of the membrane-bound conformations of the enkephalins might provide some insights into the conformations of these peptides in their receptor-bound forms (Jarrell *et al.*, 1980).

As a result of the potential therapeutic applications of the enkephalins, a considerable amount of research has been conducted on these peptides and their analogues. In addition to *in vitro* and *in vivo* studies of the activity of enkephalins and their analogues, the structure and lipid interactions of these peptides have been extensively characterized using a variety of techniques including X-ray diffraction, and circular dichroism, infrared (IR), UV-visible, UV-Raman and NMR spectroscopy (Behnam and Deber, 1984; Deber and Behnam, 1984; Deber and Behnam, 1985; Surewicz and Mantsch, 1988; Milon *et al.*, 1990; Picone *et al.*, 1990; Graham *et al.*, 1992; Hicks *et al.*, 1992; D'Alagni *et al.*, 1996; Amodeo *et al.*, 1998; Marcotte *et al.*, 2003; Marcotte *et al.*, 2004). The results of some of the NMR studies that have been conducted on methionine-enkephalin (MetEnk), and the closely related leucine-enkephalin (LeuEnk), are discussed below.

5.2.2 NMR Studies of Enkephalins

MetEnk has been the subject of numerous NMR studies that have looked at many aspects of the interactions and structure of this peptide in lipidic environments. In aqueous solution, NMR studies of MetEnk have not revealed any distinguishable secondary structure (Graham *et al.*,

1992; D'Alagni *et al.*, 1996). However, studies of MetEnk in the presence of both charged and uncharged lipids in the form of membranes, vesicles and micelles/bicelles, have shown that the peptide becomes structured upon lipid binding. The structures determined for lipid-bound MetEnk have varied depending on the nature of the lipidic environment used in the structural studies. The general consensus of these studies was that both LeuEnk and MetEnk adopt a β-turn type conformation in membrane-mimetic environments that is stabilized by a hydrogen bond between the Gly^2 C=O and the Met^5 NH (Marcotte *et al.*, 2003). However, in a recent structural study of MetEnk in phospholipid bicelles it was shown that the peptide can adopt *several* folded conformations upon membrane-association (Marcotte *et al.*, 2004). The extended conformation of MetEnk and the average structure determined for MetEnk in the presence of anionic bicelles are shown in Figure 5.1 and Figure 5.2, respectively (Marcotte *et al.*, 2004).

5.2.2.1 ^{13}C NMR Studies of MetEnk and Met-enkephalinamide in Unilamellar Vesicles

^{13}C NMR was used to study the binding of MetEnk and methioinine-enkephalinamide to anionic phosphatidylserine (PS) unilamellar lipid vesicles as a model for the binding of MetEnk to its target membrane and receptors (Jarrell *et al.*, 1980). The peptides for study were labelled with ^{13}C at the α-carbon of Gly^2 or Gly^3. The effects of pH, salt and morphine on the binding of MetEnk to PS vesicles were investigated. Peptide binding was found to be dependent on pH, with maximal binding occurring under slightly acidic conditions. High ionic strength (0.5 - 1.0 M) was shown to inhibit the binding of the peptides to the PS vesicles. It was also found that morphine was an antagonist of MetEnk binding, since morphine binds more strongly to PS vesicles than does MetEnk. No aggregation of MetEnk was observed in aqueous solution at low (1 mg/mL) or at high (100 mg/mL) concentrations, and the peptide was found to have a very high aqueous solubility.

5.2.2.2 ^{1}H and ^{13}C NMR Studies of MetEnk and LeuEnk in Phospholipid Micelles

^{1}H and ^{13}C NMR was used to study the association of MetEnk and LeuEnk with zwitterionic lysophosphatidylcholine and anionic lysophosphatidylglycerol micelles (Deber and Behnam, 1984). The authors were interested in looking at the hydrophobic and electrostatic contributions to the membrane-binding of the enkephalins. Samples were prepared with peptide concentrations of 8.72 mM for ^{1}H NMR studies and 27.3 - 28.2 mM for ^{13}C NMR studies. The results of the

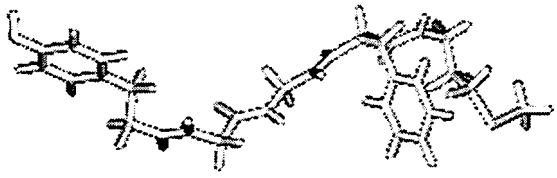
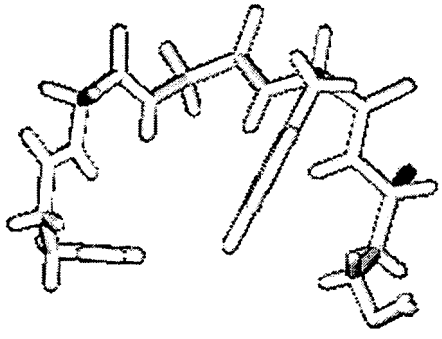


Figure 5.1. Extended Structure of Methionine-Enkephalin



PDB Accession Code 1PLW

Figure 5.2. Average Structure Determined for Methionine-Enkephalin in Anionic Bicelles Using ^{1}H NMR

NMR studies indicated that the zwitterionic peptides associated electrostatically with the negatively charged phospholipids, and hydrophobically with *both* the neutral and the negatively charged phospholipids. They were also able to identify possible sites on the peptide that were believed to be involved in major interactions with lipids.

5.2.2.3 Transferred NOE ^{1}H NMR Studies of LeuEnk and its Analogues in Liposomes

The biological activities of a number of peptide hormones and neurotransmitters have been found to correlate with their affinity for phospholipid membranes, and the conformations that these peptides adopt in the membrane-bound state. In one investigation the liposome-bound conformations of LeuEnk, its *active* analogues [D-Ala^2]Leu-enkephalin and [D-Ala^2]Leu-enkephalinamide and its *inactive* analogue [L-Ala^2]Leu-enkephalin, were studied using transferred nuclear Overhauser effect ^{1}H NMR spectrosocopy (Milon *et al.*, 1990). Previous reports in the literature had shown that if the Gly^2 of LeuEnk was replaced with a D-amino acid, the activity of the peptide was preserved or enhanced, however if Gly^2 was replaced with an L-amino acid, there was a complete loss of peptide activity. The results that were obtained from these NMR studies showed that in the lipid-bound form, the active enkephalin analogues had a common folded conformation which differed from the lipid-bound conformation of the inactive analogue.

In this study pH was found to have a significant effect on the quality of the NMR spectra that were collected. The pH at which the experiments were conducted had to be carefully selected to minimize overlap among the NH and α-proton signals of LeuEnk. It was found that there was complete overlap of the NH proton signals of Gly^3, Phe^4 and Leu^5 from pH 3.8 - 4.5 and of the α-proton signals of Leu^5 and Tyr^1 from pH 4.5 - 5.5. It was also found that the NH proton signals of Gly^2 and Gly^3 broadened and disappeared above pH 5.5.

5.2.2.4 ^{31}P and ^{2}H NMR Studies of MetEnk in Phospholipid Bicelles

A ^{31}P and ^{2}H solid-state NMR study was performed to look at the effects of phospholipid headgroups on the localization of MetEnk in membranes, since this process was known to be modulated by both hydrophobic and electrostatic interactions (Marcotte *et al.*, 2003). The study looked at the interaction of MetEnk with both zwitterionic and anionic bicelles. The bicelles were prepared with dimyristoylphosphatidylcholine (DMPC) and dicaproylphosphatidylcholine (DCPC), and depending on the experiment, up to 10 % of the DMPC was replaced with the zwitterionic phosphatidylethanolamine (DMPE) or the anionic phosphatidylglycerol (DMPG) or phosphatidylserine (DMPS), to create bicelles with different headgroup compositions. NMR samples were prepared using 20 % lipid (wt/vol) in deionized water with a pH of ~ 5.5. The peptide:lipid molar ratio used in the samples was 1:25.

The NMR data revealed that the interaction of MetEnk with the bicelles caused observable changes in the lipid dynamics of the system. It was found that the insertion depth of MetEnk into the bicelles was modulated by their composition, with a different degree of MetEnk insertion being observed for each of the four membrane systems studied. MetEnk was located at the polar/apolar interface in the bicelles. The insertion depth of the peptide into the bicelles was greatest for the DMPC bicelles, followed by the DMPE-containing, DMPS-containing and DMPG-containing bicelles at the pH studied. It appeared that the peptide was more likely to remain at the surface of anionic bicelles where electrostatic interactions seemed to play a more dominant role in membrane association than did hydrophobic interactions. These results provided some insights into the variability of peptide conformations and membrane interactions that might occur *in vivo* depending on the composition of cell membranes and the pH at different locations throughout the body where enkephalins are known to be active.

5.2.2.5 ^{1}H NMR Studies of the Conformation of MetEnk in Phospholipid Bicelles

The conformations adopted by MetEnk in the presence of isotropic bicelles were studied using multidimensional ^{1}H NMR (Marcotte *et al.*, 2004). Bicelles of two different lipid compositions were used to investigate the effects of 'membrane' composition on peptide conformation. The zwitterionic bicelles used for the studies were composed of DMPC and DCPC. Negatively charged bicelles were prepared by replacing 10 mol % DMPC with DMPG. Samples with peptide:lipid ratios of 1:25 were prepared in deionized water at pH 4.5 and 4.8. Experiments were conducted at 22°C.

This study showed that MetEnk adopts several conformations in the presence of bicelles. The structures that were determined differed slightly depending on whether the peptide was studied in the presence of zwitterionic, or anionic bicelles. These results suggest that MetEnk may be capable of forming *both* μ- and δ-selective conformers for binding to the μ- and δ-opiate receptors. The ability of MetEnk to form several conformations upon bicelle-binding is in contrast to the results obtained in previous structural studies where the enkephalins were shown to adopt well defined lipid-dependent β-turn type conformations in various membrane-mimetic environments (Behnam and Deber, 1984; Milon *et al.*, 1990). Pulsed-field gradient NMR techniques were used to determine the proportion of peptide associated with the bicelles. It was found that 65 % of the peptide was bound to PC bicelles whereas 59 % of the peptide was bound to bicelles containing 10 % PG.

5.3 Materials and Methods

5.3.1 Methionine-Enkephalin Peptide Synthesis

Peptide syntheses were performed at the University of Alberta in the Department of Chemistry by Dr. Nathaniel Martin in the laboratory of Dr. John Vederas. As previously discussed in Chapter 3, it was necessary to obtain stable isotope labelled peptide for study in MO:water cubic phases. Therefore both unlabelled and uniformly ^{13}C-, ^{15}N-labelled methionine enkephalin were synthesized using standard *O*-fluorenylmethyl-oxycarbonyl (Fmoc) chemistry.

5.3.1.1 Resin Preparation

Peptide synthesis proceeds from the C-terminal to the N-terminal residue of a peptide, so methionine was the first residue to be coupled to the resin. L-methionine-N-Fmoc (50 mg, 0.13 mmol, 1.2 eq. relative to resin) and diisopropylethylamine (Hunig's base) (92 μL, 0.53 mmol, 4.0 eq. relative to amino acid) were dissolved in 5 mL of dichloromethane (DCM) and added to 100 mg of dry 2-chloro-tritylchloride resin (estimated capacity = 1.1 mmol/g) and stirred for 2 hours. The resin was washed with 3 × 5 mL DCM:methanol:diisopropylethylamine (17:2:1), 3 × 5 mL DCM and then with 2 × 5 mL N,N-dimethylformamide (DMF). The resin was dried *in vacuo* over KOH before treatment with 10 mL of piperidine:DMF (2:8) for 3 minutes. The resin was drained and the piperidine:DMF treatment was repeated 3 times to ensure that the Fmoc group had been removed. The resin was then washed with 6 × 5 mL of DMF to remove all traces of piperidine.

5.3.1.2 Residue Coupling

Phenylalanine (Position 4)

L-phenylalanine-N-Fmoc (85 mg, 0.22 mmol, 2 eq. to resin loading), the coupling reagent PyBOP (110 mg, 0.21 mmol, 1.95 eq.) and 1-hydroxybenzotriazole hydrate (30 mg, 0.22 mmol, 2 eq.), a reagent to suppress racemization, were dissolved in 5 mL of DMF. Diisopropylethylamine (190 μL, 1.1 mmol, 10 eq. relative to resin loading) was added to the amino acid solution. This mixture was then added to the resin and stirred for 2 hours. The resin was washed with 3 × 5 mL DMF. The Fmoc protected peptidyl resin was end capped by treatment with 10 mL of acetic anhydride:DMF (2:8) for 15 minutes. The resin was drained and the acetic anhydride:DMF treatment was repeated 3 times. The resin was washed with 3 × 5 mL DMF. The Fmoc protected peptidyl resin was treated with 10 mL of piperidine:DMF (2:8) for 3 minutes. The resin was drained and the piperidine:DMF treatment was repeated 3 times. The resin was washed with 6 × 5 mL DMF.

Glycine (Position 3)

Glycine-N-Fmoc (65 mg, 0.22 mmol, 2 eq. to resin loading), PyBOP (110 mg, 0.21 mmol, 1.95 eq.) and 1-hydroxybenzotriazole hydrate (30 mg, 0.22 mmol, 2 eq.) were dissolved in 5 mL of DMF. Diisopropylethylamine (190 μL, 1.1 mmol, 10 eq. relative to resin loading) was added to the amino acid solution. This mixture was then added to the resin and stirred for 2 hours. The resin was washed with 3 × 5 mL DMF. The Fmoc protected peptidyl resin was end capped by treatment with 10 mL of acetic anhydride:DMF (2:8) for 15 minutes. The resin was drained and the acetic anhydride:DMF treatment was repeated 3 times. The resin was washed with 3 × 5 mL DMF. The Fmoc protected peptidyl resin was treated with 10 mL of piperidine:DMF (2:8) for 3 minutes. The resin was drained and the piperidine:DMF treatment was repeated 3 times. The resin was washed with 6 × 5 mL DMF.

Glycine (Position 2)

The second glycine residue in the peptide sequence was coupled using the procedure described above for the coupling of the third glycine residue in the peptide sequence.

Tyrosine (N-terminal Position)

L-tyrosine-N-Fmoc-*O*-t-butyl (100 mg, 0.22 mmol, 2 eq. to resin loading), PyBOP (110 mg, 0.21 mmol, 1.95 eq.) and 1-hydroxybenzotriazole hydrate (30 mg, 0.22 mmol, 2 eq.) were dissolved in 5 mL of DMF. Diisopropylethylamine (190 μL, 1.1 mmol, 10 eq. relative to resin loading) was added to the amino acid solution. This mixture was then added to the resin and stirred for 2 hours. The resin was washed with 3 × 5 mL DMF. The Fmoc protected peptidyl resin was treated with 10 mL of piperidine:DMF (2:8) for 3 minutes. The resin was drained and the piperidine:DMF treatment was repeated 3 times. The resin was washed with 6 × 5 mL DMF, followed by washing with 6 × 5 mL DCM.

5.3.1.3 Cleavage from Resin

The resin was placed in a dry flask and treated with ~ 5 mL of a trifluoroacetic acid solution (94 % TFA, 2.5 % H_2O, 2.5 % ethanedithiol, 1 % triisopropylsilane). The flask was stoppered and sonicated at room temperature for 1 hour. The resin was removed by filtration and washed twice with ~ 5 mL of TFA. The volume of the filtrate was reduced by evaporation and dissolved in ~ 10 mL of the initial HPLC solvent.

5.3.1.4 HPLC Purification

The peptide was purified by reversed-phase HPLC on a Vydac C_{18} prep scale column (2.2 × 25 cm, 300 Å, 10 μ) using a water (0.1 % TFA):acetonitrile (0.1 % TFA) solvent system. The flow rate was 10.0 mL/min and the absorbance was measured at 220 nm. The HPLC method involved running 20 % acetonitrile (0.1 % TFA) for 5 minutes, followed by a gradient from

20 % → 45 % acetonitrile over 17 minutes, returning from 45 % → 20 % acetonitrile in 30 seconds and then running 20 % acetonitrile for 5 minutes. MetEnk eluted as a single isolated peak with a retention time of 13 - 15 minutes using these HPLC conditions. Approximately 20 mg each of the labelled and unlabelled peptides were obtained following HPLC purification.

5.3.1.5 Purity Assessment

The purity of the peptides was assessed by reversed-phase HPLC, and the molecular weight of the peptides was confirmed using MALDI-TOF mass spectrometry. The molecular weight of unlabelled MetEnk was 573.66 g and the molecular weight of ^{15}N-,^{13}C-labelled MetEnk was 605.30 g. A 1D ^{1}H NMR spectrum and a 2D ^{1}H COSY NMR spectrum were collected on the unlabelled MetEnk to confirm sample purity and peptide sequence. A 1D ^{13}C NMR spectrum and a 2D ^{1}H HMQC NMR spectrum were collected on the ^{15}N-,^{13}C-labelled MetEnk to confirm sample purity and uniform ^{13}C labelling.

5.3.2 NMR Sample Preparation

For NMR studies it is optimal to have as high a peptide concentration as possible in a sample, in particular when working with samples which contain high concentrations of lipids. In studies of MetEnk in bicelles, the peptide was incorporated at a ratio of 1:25 peptide:lipid without compromising bicelle integrity (Marcotte *et al.*, 2003). Since the effects of MetEnk on MO cubic phase behaviour were not known, cubic phase samples containing three different concentrations of peptide were prepared. Samples were prepared in deionized water to minimize any potential influences of salts on the interaction of MetEnk with the cubic phase (Jarrell *et al.*, 1980; Milon *et al.*, 1990). The pH of the 10 % D_2O solution used for sample preparation was ~ 5.5. At this pH MetEnk is believed to exist in a zwitterionic form because the pK_a values of the amino and carboxyl groups of the peptide are 7.5 and 3.9 respectively (Jarrell *et al.*, 1980; Marcotte *et al.*, 2003).

Cubic phases with a composition of 70 % MO (wt/wt) in water containing 3.5 mg, 1.5 mg and 0.4 mg of ^{15}N,^{13}C-labelled MetEnk were prepared using a 10 % D_2O solution at ~ pH 5.5 following the procedures outlined previously in Section 4.3.1 for the preparation of transmembrane peptide-containing cubic phase NMR samples. These samples were called MetEnk:MO Sample '1', MetEnk:MO Sample '2' and MetEnk:MO Sample '3', respectively. MetEnk:MO Sample '1' had a peptide concentration of 10.5 mM and a peptide:lipid ratio of 1:180, MetEnk:MO Sample '2' had a peptide concentration of 7.5 mM and a peptide:lipid ratio of 1:254, and MetEnk:MO Sample '3' had a peptide concentration of 2.0 mM and a peptide:lipid ratio of 1:967. Two 'control' samples were also prepared for comparison of NMR spectra with those collected on the MetEnk-containing cubic phase samples. The first 'control' sample was of

70 % MO cubic phases that *did not* contain any peptide. The second 'control' sample was of 2.1 mg of ^{15}N,^{13}C-labelled MetEnk dissolved in ~ 600 μL of a 10 % D_2O solution at ~ pH 5.5. It was called MetEnk:H_2O, and this sample was stored at 4°C when not in use.

5.3.3 NMR Data Collection

5.3.3.1 Multidimensional Heteronuclear NMR Spectra

NMR experiments were conducted at the University of Alberta in the Department of Biochemistry with the assistance of Dr. Brian Sykes and Dr. Leo Spyracopoulos. NMR data were collected at 30°C on a Varian Inova 500 MHz spectrometer equipped with a 5 mm triple resonance z-pulse field gradient probe (unless otherwise noted). All ^{1}H, ^{13}C and ^{15}N chemical shifts were directly or indirectly referenced relative to an external DSS solution set to 0.0 ppm. 1D ^{1}H NMR spectra were collected on all of the NMR samples that were prepared (see Section 5.3.2), except for MetEnk:MO Sample '1' because this sample had not formed a homogeneous cubic phase. 2D ^{1}H-^{15}N HSQC spectra were collected on all of the peptide-containing samples that were successfully prepared. It was found that MetEnk:MO Sample '2' yielded heterogeneous spectra, so no further NMR experiments were conducted on this sample. In addition to the standard 2D ^{1}H-^{15}N HSQC spectra, it was possible to collect 1D ^{15}N-edited ^{1}H spectra because the peptide NH peaks did not overlap. 1D ^{15}N-edited ^{1}H spectra were collected on MetEnk:H_2O and MetEnk:MO Sample '3'. A 2D aromatic region ^{1}H-^{13}C HSQC spectrum, and a phase-sensitive 2D aliphatic region ^{1}H-^{13}C HSQC spectrum, were collected on MetEnk:MO Sample '3'.

The well resolved NH peaks in the ^{1}H-^{15}N HSQC spectra of MetEnk:MO Sample '3' made it possible to run the standard 3D ^{15}N-edited TOCSY and 3D ^{15}N-edited NOESY experiments as 2D experiments which reduced both the time required for data acquisition, and the size of the data files. A 2D ^{15}N-edited TOCSY-HSQC spectrum was acquired with the DIPSI isotropic mixing sequence (Marion *et al.*, 1989a; Cavanagh and Rance, 1992). Two 2D ^{15}N-edited gradient-enhanced nuclear Overhauser enhancement spectroscopy (NOESY)-HSQC spectra were acquired using mixing times of 150 ms and 250 ms (Zhang *et al.*, 1994). The data from these experiments were plotted (see Figure C 1 in Appendix C) in order to determine which mixing time was optimal for maximizing signal-to-noise while minimizing spin diffusion. It was found that a mixing time of 250 ms yielded good quality NOESY spectra. A 2D aromatic region ^{13}C-edited NOESY spectrum, and a 2D ^{13}C-edited HSQCNOESYHSQC spectrum were collected on MetEnk:MO Sample '3' at 30°C on a Varian Inova 600 MHz spectrometer equipped with a 5 mm triple resonance z-pulse field gradient probe.

5.3.3.2 Pulsed Field Gradient NMR Measurement of Diffusion

Pulsed-field gradient NMR experiments to measure lipid, peptide and water diffusion in MetEnk:H_2O and MetEnk:MO Sample '3' at 20°C, 30°C and 40°C were conducted at the University of British Columbia Laboratory for Molecular Biophysics by Dr. Mark Okon using a Varian Unity 500 MHz spectrometer equipped with a 5 mm triple resonance z-pulse field gradient probe. The pulse sequences used for the experiments were previously described in Section 4.3.3. Water suppression was achieved by using the gradients to dephase the water signal in the experiment conducted to measure MetEnk diffusion in pure water.

5.3.4 NMR Data Analysis

Multidimensional heteronuclear NMR spectral data were processed and analyzed using the software programs *VNMR version 5.1* (Varian), *NMRPipe* and *NMRView* (Johnson and Blevins, 1994; Delaglio *et al.*, 1995). It was possible to assign all of the peaks in the spectra collected on MetEnk:MO Sample '3', however it was not possible to make stereospecific assignments. The N-terminal amino group was not observable in any of the spectra that were collected, but this was not unexpected as this group is rarely observable in solution NMR spectra. The NOE peaks from the 2D ^{15}N-edited NOESY spectrum collected using a mixing time of 250 ms were quantified using peak volume with the software program *NMRView* using the NOE automatic calibration function.

The NMR data collected to measure lipid, peptide and water diffusion were processed and analyzed using the software program *MestReC* as previously described in Section 4.3.4 (Cobas and Sardina, 2003). Exponential multiplication with a line broadening of 10 Hz was applied to all data sets collected to measure lipid and water diffusion. For data sets collected to measure peptide diffusion, exponential multiplication using a line broadening of 8 Hz, in conjunction with a 90° shifted sine squared function were applied prior to Fourier transformation of the data. Data extracted from the processed NMR spectra were imported into *Excel* spreadsheets and plotted appropriately in order to determine the lipid, peptide and water diffusion coefficients for all of the data sets that were collected.

5.3.5 Structure Calculations

NMR experiments where NOEs were measured provided the data used for structure calculations. It is known that an NOE is observed between two hydrogen atoms if they are located less than ~ 5 Å from one another in space (Wüthrich, 1986). Since the intensity of the NOE peak is proportional to the distance between the two hydrogen atoms, it is possible to derive distance constraints from the peak intensities of the observable NOE signals. When this information is

combined with sequence specific chemical shift assignments for a polypeptide chain, conformational space can be searched for spatial arrangements of the polypeptide chain that are compatible with the NOE-derived distance constraints. NOEs can be observed between atoms of the *same* amino acid residue or between atoms from *different* amino acid residues. It is the NOEs between non-adjacent amino acid residues (non-sequential NOEs) that are the most useful for defining the 3D structure of a polypeptide. These non-sequential NOEs indicate which regions of the polypeptide chain are close together in the folded conformation of the peptide/protein.

In the case of the NOE data that was collected on MetEnk:MO Sample '3', there was a notable lack of unambiguous non-sequential NOE peaks. A number of the observed NOE peaks were ambiguous, i.e. they had more than one possible assignment. There was one peak where the possible assignments included both sequential and non-sequential assignment possibilities, so it was at least conceivable that this peak represented a non-sequential NOE. The lack of non-sequential NOEs was consistent with the peptide existing in a conformationally flexible, poorly defined structure under these conditions.

Standard protein structure calculation methods are poor at generating structure ensembles that reflect conformationally flexible peptides, so care must be taken to avoid generating a misleadingly narrow set of structures that would imply a better defined structure than is really present. To address this issue, and also to take full advantage of the ambiguous NOE data, a special structure calculation strategy was employed (Booth *et al.*, 2004). In this method, a large variety of structures are generated to try to sample a large region of conformational space and then from these structures, the subset that was consistent with the NMR data was selected.

Structures were generated by simulated annealing using the program *Crystallography and NMR System (CNS) version 1.2* (Brünger *et al.*, 1998). The interproton distance constraints used for structure calculations were obtained from the ^{15}N-edited NOESY and ^{13}C-edited aromatic NOESY spectra that were collected on MetEnk:MO Sample '3'. A total of 900 structures were calculated using the inter-proton distances derived from the NOE data. In order to maximize the variety of structures generated, several structure calculation runs were performed, with each run using different combinations of possible assignments for the ambiguous NOEs. All structure calculations were run remotely from SFU using the computing facilities at the Advanced Computation and Visualization Centre at Memorial University of Newfoundland under the guidance of Dr. Valerie Booth in the Department of Biochemistry at Memorial University of Newfoundland.

The pool of structures generated for cubic phase-bound MetEnk was evaluated using the software program *Ensemble* to define the subset of structures that were consistent with the NMR data (Choy and Forman-Kay, 2001). *Ensemble* is a program that was developed for the analysis of a collection of structures representing the unfolded state of a protein. The ultimate goal of this program is to use a weighted ensemble of rigid structures to represent the unfolded state of a protein. The relative weighting of each of these rigid structures is proportional to the time-averaged existence of each particular conformation. Since the structures calculated for cubic phase-bound MetEnk were of an extended nature, having many similarities with the conformations of unfolded peptides and proteins, it seemed appropriate to use a program such as *Ensemble* to analyze the structures. A combination of NOE and/or chemical shift constraints were used to evaluate the pool of structures calculated for cubic phase-bound MetEnk and determine which ones were most consistent with the experimental data. The software program *Molmol* was used to visualize the structures and analyze the calculated structures (Koradi *et al.*, 1996). Supplementary information relevant to the structure calculation procedure can be found in Appendix C.

5.4 Results

5.4.1 MetEnk-Containing Cubic Phase Preparation

When the MetEnk-containing cubic phase samples were prepared it was very difficult to obtain homogeneous clear samples, especially for MetEnk:MO Sample '1' which contained the highest peptide concentration. The samples had cloudy regions, even after extensive centrifugation and prolonged equilibration times under ambient conditions so additional water was added to see if this would improve sample clarity by favouring cubic phase formation over lamellar phase formation. It appeared that the addition of small quantities of water to the samples was necessary for the formation of clear homogeneous cubic phase samples.

5.4.2 Multidimensional Heteronuclear NMR Spectra

5.4.2.1 1D ^{1}H Spectra

MetEnk:H_2O Sample

The 1D ^{1}H spectra of MetEnk in water with and without decoupling showed that the peptide sample was very pure and that the peptide was well dispersed in aqueous solution.

MetEnk:MO Sample '2' and MetEnk:MO Sample '3'

1D ^{1}H spectra of the MetEnk-containing cubic phases were collected for comparison with the spectrum of cubic phases that did not contain peptide. Due to the high lipid:peptide ratios in these samples and limitations in the dynamic range of the NMR spectrometers, it was only

possible to observe the peaks arising from MO in these spectra. The spectra appeared to be identical to the ones collected previously of G-type MO:water cubic phases. It appeared that the addition of MetEnk to the MO cubic phases at the concentrations studied did not cause any observable changes to the 1D ^{1}H spectrum of the cubic phase.

5.4.2.2 1D ^{1}H-^{15}N HSQC Spectra

MetEnk:H_2O Sample

A ^{15}N-edited, 1D ^{1}H spectrum of MetEnk in water was collected for comparison with the data collected on MetEnk:MO Sample '3'.

MetEnk:MO Sample '3'

In order to observe the ^{1}H resonances originating from MetEnk that was incorporated into the MO cubic phase it was necessary to collect a ^{15}N-edited ^{1}H spectrum. In this type of experiment the only protons that are observable are those directly attached to ^{15}N. In this sample there were four observable protons, the amide backbone protons of G^2, G^3, F^4 and M^5. As one might expect, the N-terminal protons of Y^1 were not observable due to their rapid exchange with water.

5.4.2.3 2D ^{1}H-^{15}N HSQC Spectra

MetEnk:H_2O Sample

The 2D ^{1}H-^{15}N HSQC correlation spectrum of MetEnk in water was collected for comparison with the spectrum of MetEnk in the MO cubic phase. The spectrum showed four sharp well resolved peaks and is included in Appendix C as Figure C.1.

MetEnk:MO Sample '2' and MetEnk:MO Sample '3'

The spectra collected on MetEnk in the cubic phase differed between the two samples. The spectrum collected on MetEnk:MO Sample '2' had 7 distinct cross peaks, whereas the spectrum collected on MetEnk:MO Sample '3' had four distinct cross peaks. Both spectra are included in Appendix C. Figure C.2. shows the 2D ^{1}H-^{15}N HSQC spectrum collected on MetEnk:MO Sample '3' (which is also shown in Figure 5.3), and Figure C.3. the 2D ^{1}H-^{15}N HSQC spectrum collected on MetEnk:MO Sample '2'. When these two spectra were superimposed on one another it was found that the four peaks in the spectrum of MetEnk:MO Sample '3' were coincident with four of the peaks in the spectrum of MetEnk:MO Sample '2'. When the spectrum of MetEnk in water was superimposed on the spectrum of MetEnk:MO Sample '2' it was found that one of the peaks was coincident with a peak that was *also* coincident with a peak in the MetEnk Sample '3' spectrum, however the other three peaks in the water spectrum were found to be coincident with the three remaining peaks in the MetEnk Sample '2' spectrum. Figure C.4. in Appendix C shows a superimposition of the 2D ^{1}H-^{15}N HSQC spectra collected on MetEnk in water and on MetEnk:MO Sample '3', for comparison with the spectrum collected on MetEnk:MO Sample '2' which is shown in Figure C.3.

5.4.2.4 2D ^{1}H-^{13}C HSQC Spectra

MetEnk:MO Sample '3'

The 2D ^{1}H-^{13}C HSQC spectrum of MetEnk:MO Sample '3' is shown in Figure 5.5 and the 2D aromatic region ^{1}H-^{13}C HSQC spectrum is shown in Figure 5.4. It was possible to completely assign this spectrum using chemical shift information provided by the NOESY and TOCSY experiments that were run on this sample, as well as information provided by reference spectra from the literature for glycerol and oleic acid, the constituents of MO.

5.4.2.5 2D ^{15}N-edited TOCSY Spectra

A 2D ^{15}N-edited TOCSY experiment was run on MetEnk:MO Sample '3'. The spectrum is shown in Figure 5.7. It was possible to completely assign the spectrum using the information provided by the ^{5}N-edited NOESY spectrum.

5.4.2.6 2D ^{15}N-edited NOESY Spectra

2D ^{15}N-edited NOESY experiments were run on MetEnk:MO Sample '3' using 150 ms and 250 ms mixing times. The NOESY spectrum collected using a mixing time of 250 ms is shown in Figure 5.8. The 150 ms NOESY spectrum only differed from the 250 ms spectrum with respect to peak intensities (they were lower in the 150 ms spectrum), so the 150 ms NOESY spectrum is not shown. The information provided by this spectrum was used to obtain complete peak assignments for the spectra that were collected on MetEnk:MO Sample '3'. The NOE data from the 150 ms and 250 ms experiments were used to plot NOE build-up curves (shown in Appendix C in Figure C 5) to determine which mixing time would yield the highest quality NOE data. This data would be used to generate NOE-derived constraints for structure calculation purposes.

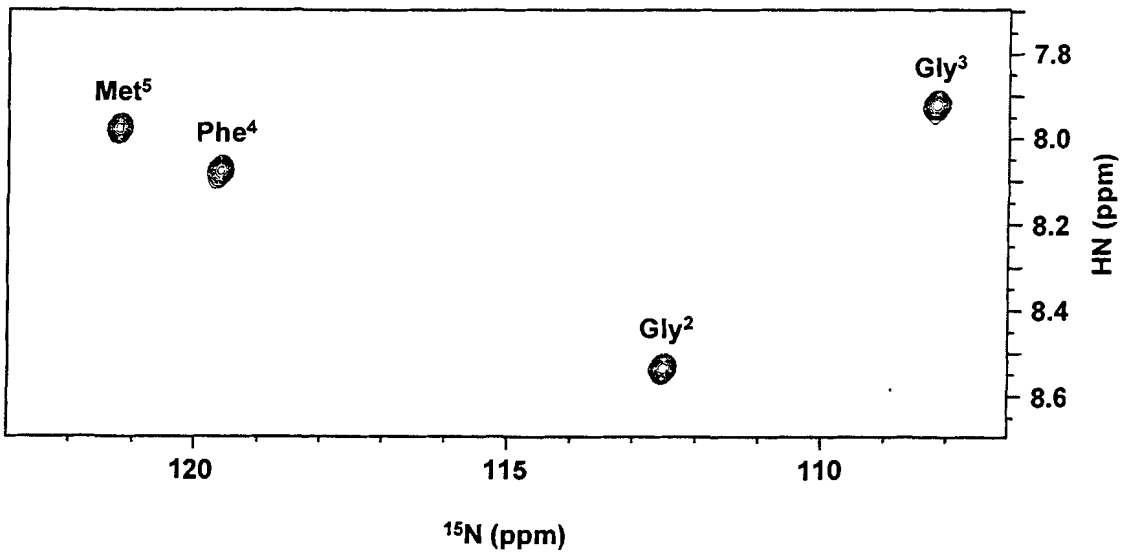


Figure 5.3. ^{1}H-^{15}N HSQC Spectrum of MetEnk:MO Sample '3' at 30°C

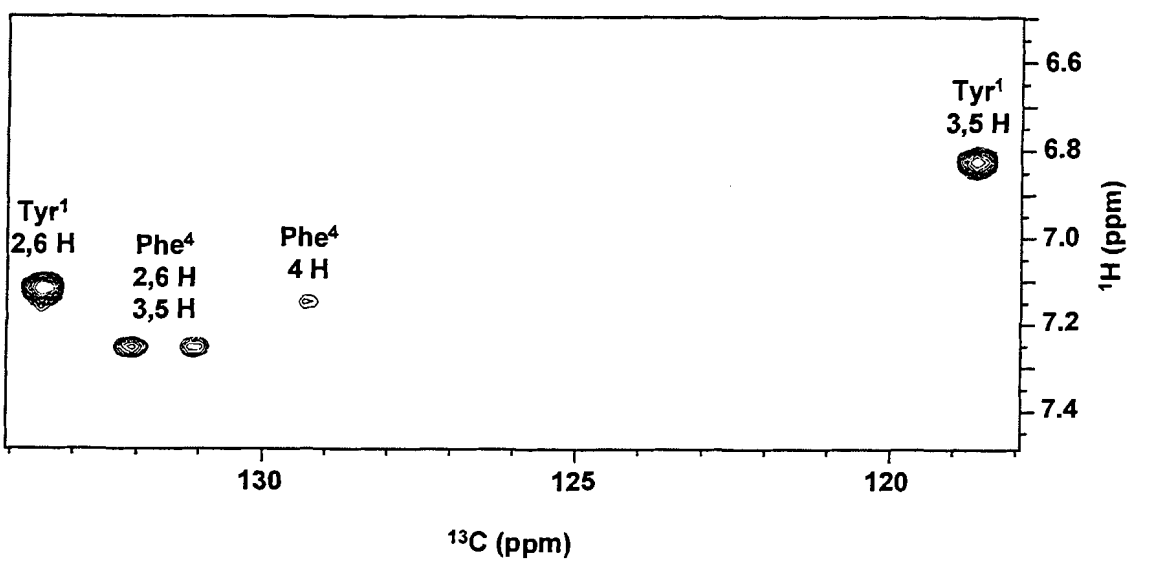


Figure 5.4. ^{1}H-^{13}C Aromatic Region HSQC Spectrum of MetEnk:MO Sample '3' at 30°C

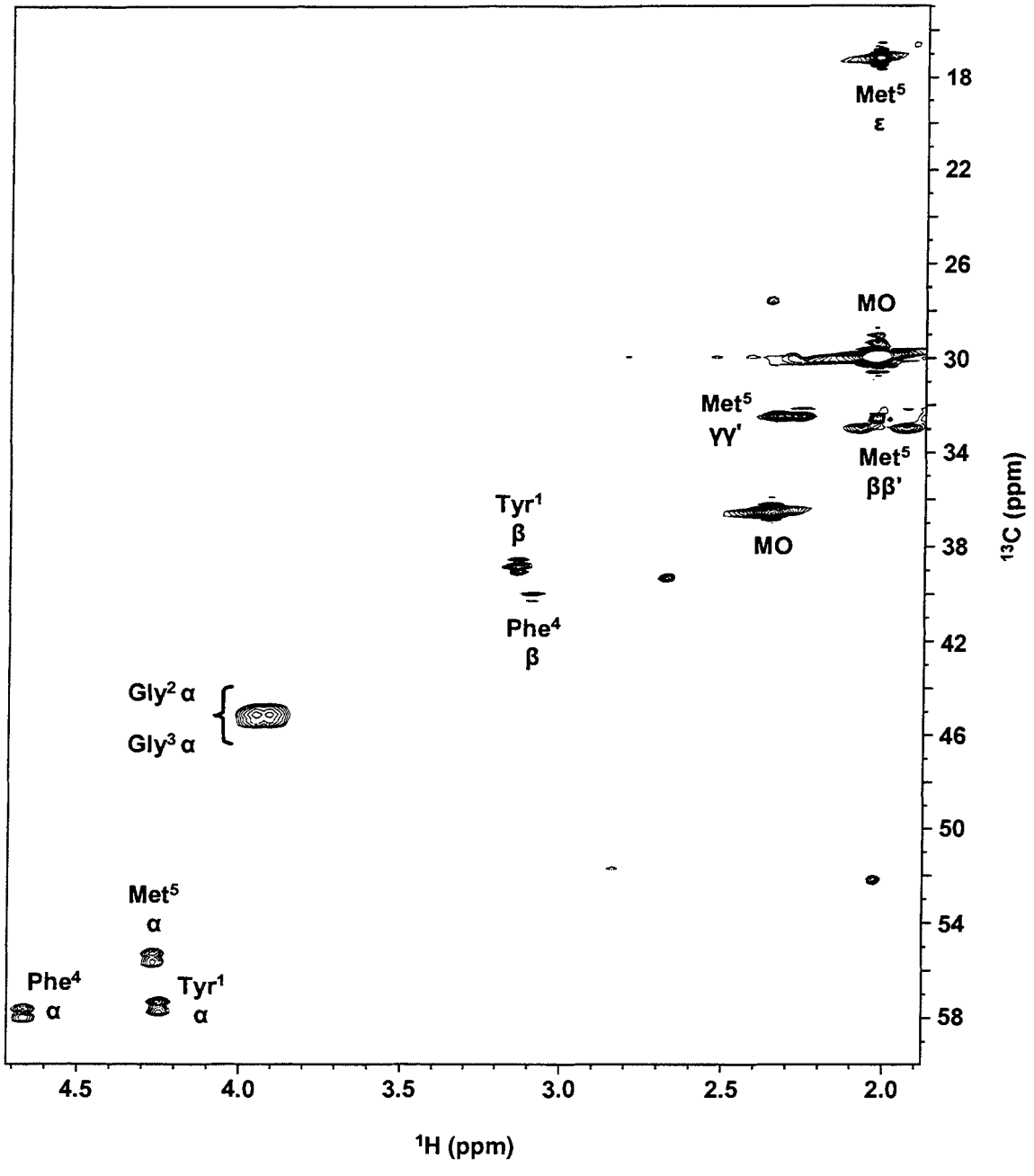


Figure 5.5. ^{1}H-^{13}C HSQC Spectrum of MetEnk:MO Sample '3' at 30°C

Table 5.1. Chemical Shifts Determined for Methionine Enkephalin in MetEnk:MO Sample '3' at 30°C

Amino Acid	Atom ID*	Chemical Shift (δ) in ppm	NMR Experiment Chemical Shift Derived From
Tyr[1]	1.N	–	–
	1.HT1	–	–
	1.CA	57.5	^{1}H-^{13}C HSQC
	1.HA	4.25	^{1}H-^{13}C HSQC
	1.CB	38.8	^{1}H-^{13}C HSQC
	1.HB#2	3.14	^{1}H-^{13}C HSQC
	1.HB#1	3.14	^{1}H-^{13}C HSQC
	1.CG	–	–
	1.CD#1	133.5	^{1}H-^{13}C Aromatic HSQC
	1.HD#1	7.12	^{1}H-^{13}C Aromatic HSQC
	1.CE#1	118.6	^{1}H-^{13}C Aromatic HSQC
	1.HE#1	6.82	^{1}H-^{13}C Aromatic HSQC
	1.CZ	–	–
	1.HH	–	–
	1.CE#2	118.6	^{1}H-^{13}C Aromatic HSQC
	1.HE#2	6.82	^{1}H-^{13}C Aromatic HSQC
	1.CD#2	133.5	^{1}H-^{13}C Aromatic HSQC
	1.HD#2	7.12	^{1}H-^{13}C Aromatic HSQC
	1.C	–	–
Gly[2]	2.N	112.5	^{1}H-^{15}N HSQC
	2.HN	8.54	^{1}H-^{15}N HSQC
	2.CA	45.2	^{1}H-^{13}C HSQC
	2.HA#2	3.95	^{1}H-^{13}C HSQC
	2.HA#1	3.91	^{1}H-^{13}C HSQC
	2.C	–	–
Gly[3]	3.N	108.2	^{1}H-^{15}N HSQC
	3.HN	7.92	^{1}H-^{15}N HSQC
	3.CA	45.2	^{1}H-^{13}C HSQC
	3.HA#2	3.95	^{1}H-^{13}C HSQC
	3.HA#1	3.91	^{1}H-^{13}C HSQC
	3.C	–	–

Cont'd...

Table 5.1. Continued

Amino Acid	Atom ID*	Chemical Shift (δ) in ppm	NMR Experiment Chemical Shift Derived From
Phe[4]	4.N	119.6	^{1}H-^{15}N HSQC
	4.HN	8.08	^{1}H-^{15}N HSQC
	4.CA	57.8	^{1}H-^{13}C HSQC
	4.HA	4.67	^{1}H-^{13}C HSQC
	4.CB	40.0	^{1}H-^{13}C HSQC
	4.HB#2	3.10	^{1}H-^{13}C HSQC
	4.HB#1	3.10	^{1}H-^{13}C HSQC
	4.CG	–	–
	4.CD#1	132.1	^{1}H-^{13}C Aromatic HSQC
	4.HD#1	7.25	^{1}H-^{13}C Aromatic HSQC
	4.CE#1	131.1	^{1}H-^{13}C Aromatic HSQC
	4.HE#1	7.25	^{1}H-^{13}C Aromatic HSQC
	4.CZ	129.3	^{1}H-^{13}C Aromatic HSQC
	4.HZ	7.15	^{1}H-^{13}C Aromatic HSQC
	4.CE#2	131.1	^{1}H-^{13}C Aromatic HSQC
	4.HE#2	7.25	^{1}H-^{13}C Aromatic HSQC
	4.CD#2	132.1	^{1}H-^{13}C Aromatic HSQC
	4.HD#2	7.25	^{1}H-^{13}C Aromatic HSQC
	4.C	–	–
Met[5]	5.N	121.2	^{1}H-^{15}N HSQC
	5.HN	7.98	^{1}H-^{15}N HSQC
	5.CA	55.5	^{1}H-^{13}C HSQC
	5.HA	4.27	^{1}H-^{13}C HSQC
	5.CB	33.0	^{1}H-^{13}C HSQC
	5.HB#2	2.08	^{1}H-^{13}C HSQC
	5.HB#1	1.94	^{1}H-^{13}C HSQC
	5.CG	32.5	^{1}H-^{13}C HSQC
	5.HG#2	2.33	^{1}H-^{13}C HSQC
	5.HG#1	2.27	^{1}H-^{13}C HSQC
	5.CE	17.2	^{1}H-^{13}C HSQC
	5.HE1	2.03	^{1}H-^{13}C HSQC
	5.C	–	–

* In the column for *Atom ID,* the number indicates the residue number starting from the amino terminus of the peptide. The first letter indicates the atom type, and if it is followed by another letter, this indicates the position of the atom in the amino acid, where *A* refers to atoms closest to the amide bond and those farther away are indicated sequentially using the Greek alphabet (α, β, γ, δ, ε and ζ), ending with *Z*. A number sign (*#*) indicates degeneracy between atoms. If the first letter is not followed by another letter, then it is either the amide nitrogen, *N* or the carbonyl carbon, *C*. The amino terminal protons are always indicated as *HT*.

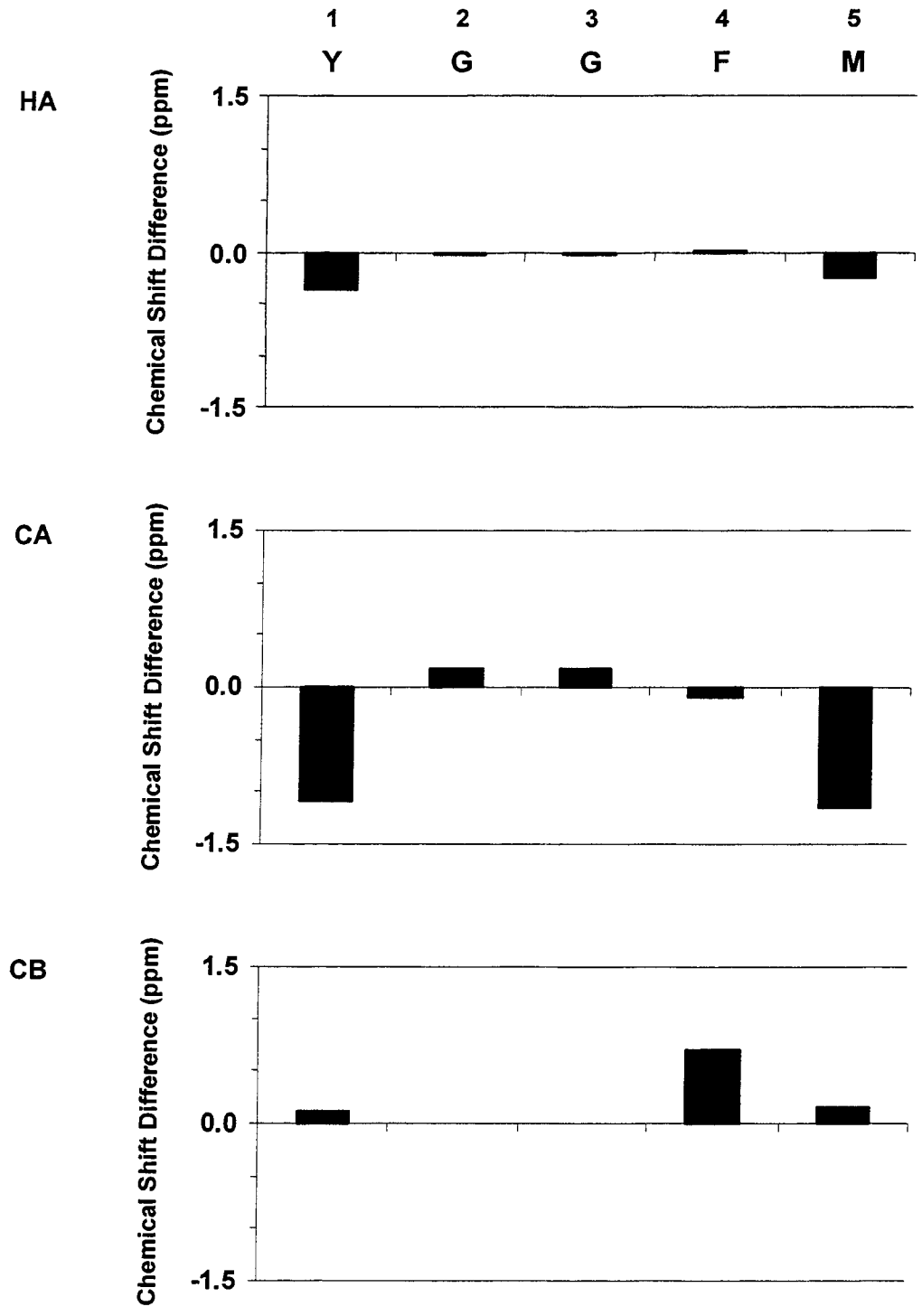


Figure 5.6. Chemical Shift Differences for the HA, CA & CB Peaks in MetEnk:MO Sample '3'

The differences between the observed chemical shifts of the HA, CA and CB peaks in MetEnk Sample '3', and the random coil chemical shifts supplied with CSI v.2.0 SGI 6.0 ($\Delta\delta$) were calculated by subtracting $\delta_{random\ coil}$ from $\delta_{observed}$ (Wishart and Sykes, 1994).

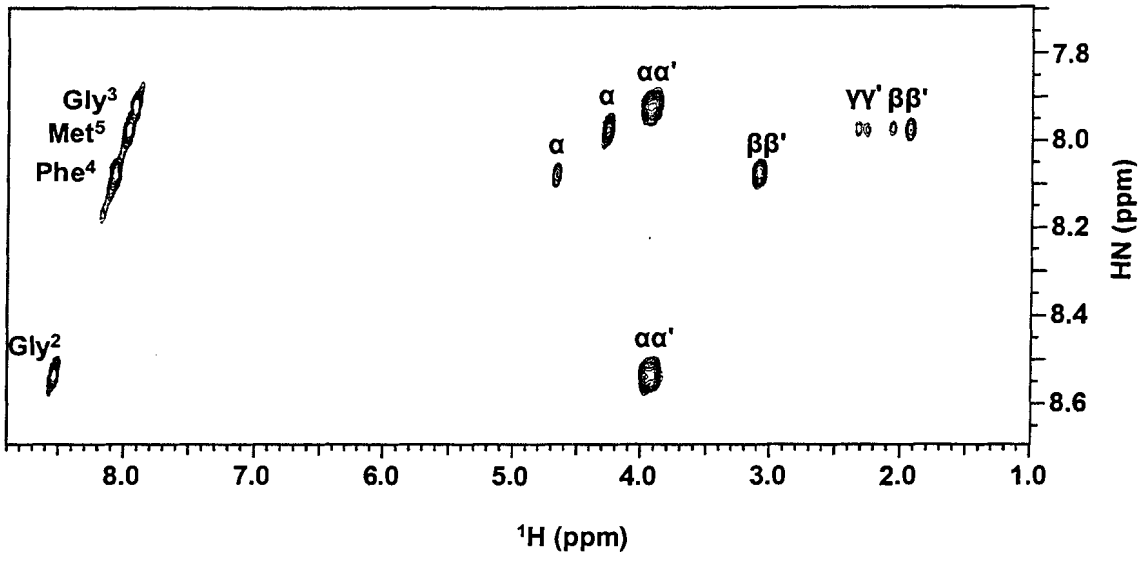


Figure 5.7. ^{1}H-^{15}N TOCSY Spectrum of MetEnk:MO Sample '3' at 30°C

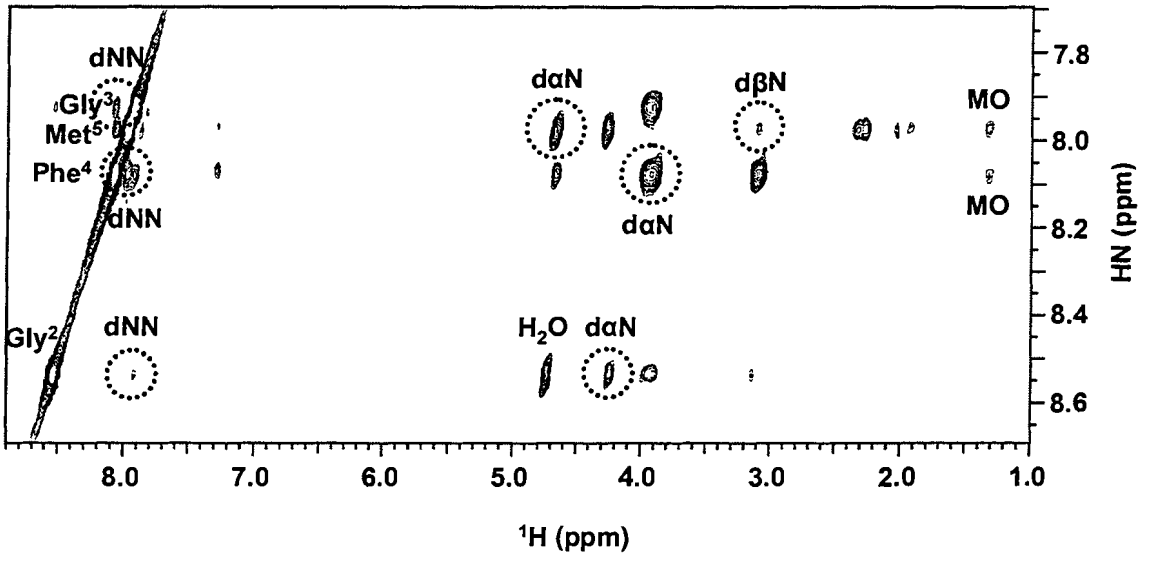


Figure 5.8. ^{1}H-^{15}N NOESY Spectrum of MetEnk:MO Sample '3' at 30°C

Dotted circles indicate interresidue NOE peaks which are labelled according to the naming convention shown in Figure 5.11. The NOE peaks observed between MetEnk and water, and MetEnk and MO are also labelled.

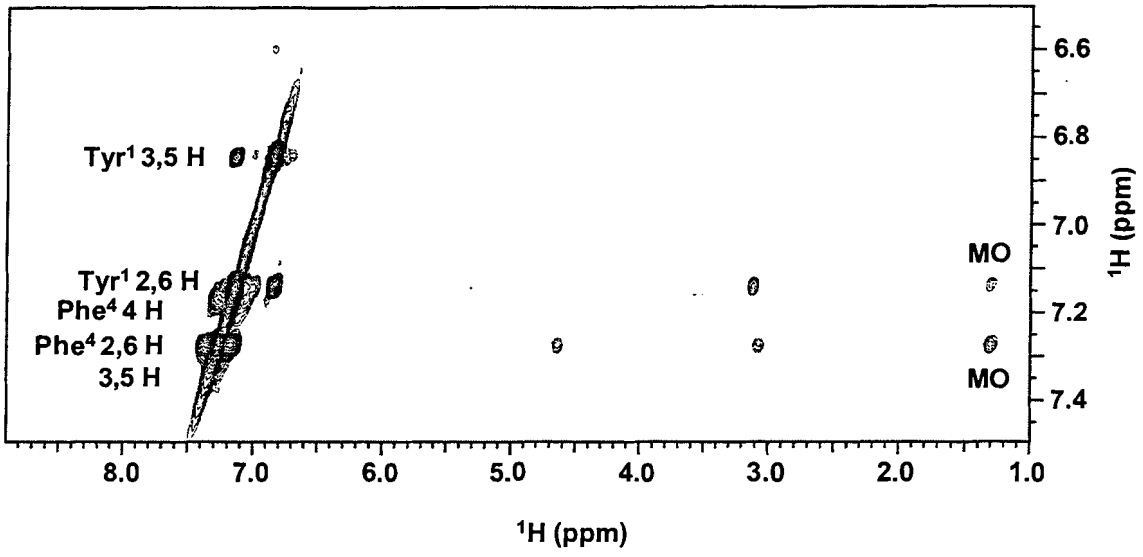


Figure 5.9. **^{1}H-^{13}C Aromatic NOESY Spectrum of MetEnk:MO Sample '3' at 30°C**

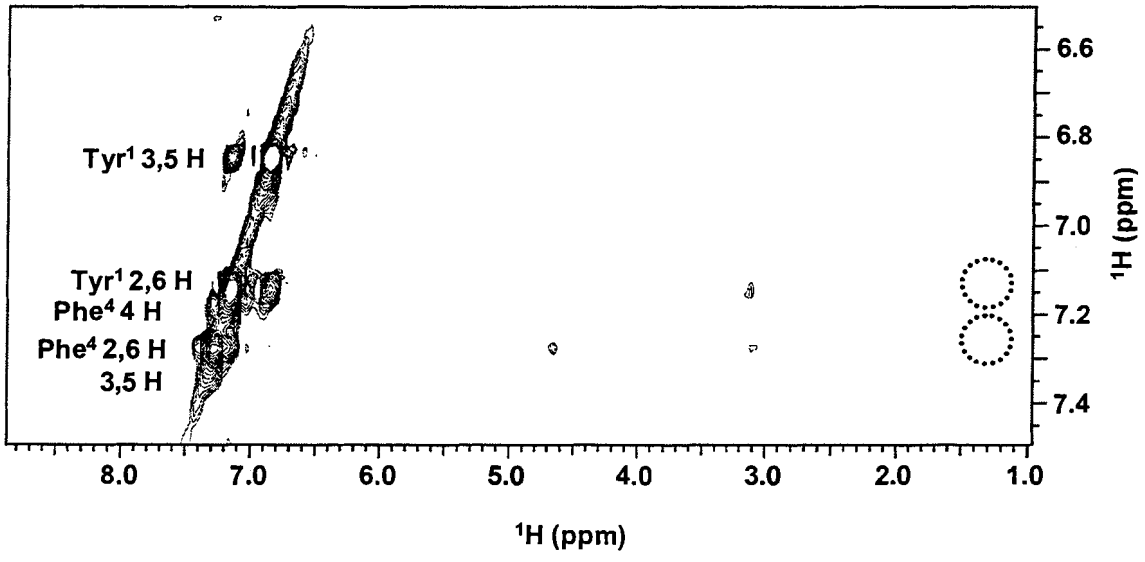


Figure 5.10. **^{1}H-^{13}C HSQCNOESYHSQC Spectra of MetEnk:MO Sample '3' at 30°C**

Dotted circles indicate the absence of NOE peaks that were observed in the ^{1}H-^{13}C aromatic NOESY spectrum shown above in Figure 5.9.

5.4.2.7 ^{1}H-^{13}C Aromatic NOESY and HSQCNOESYHSQC Spectra

Spectra were collected on MetEnk:MO Sample '3' to verify which NOE peaks arose from peptide-peptide interactions, and which ones arose from lipid-peptide interactions. The two experiments that were conducted for these purposes were the 2D aromatic region ^{13}C-edited NOESY spectrum and the 2D ^{13}C-edited HSQCNOESYHSQC which are shown in Figure 5.9 and Figure 5.10.

5.4.3 Chemical Shift Data

The NMR spectra collected on MetEnk:MO Sample '3' were analyzed with the aid of the software program *NMRView*. A list of all of the atoms in MetEnk, along with their experimentally determined chemical shifts are shown in Table 5.1. In this table (also known as a peaklist), the NMR experiment from which the chemical shift was derived is also listed. The notation used to denote the various atoms in MetEnk is the notation used by the software program *CNS*. It was possible to obtain non-stereospecific assignments for all of the peaks in the spectra collected on MetEnk:MO Sample'3'.

A comparison of the chemical shifts for cubic phase-bound MetEnk with those for amino acids in a random coil conformation was made. The random coil HA, CA and CB chemical shifts used for the analysis were the ones supplied with CSI_v2.0 (Wishart and Sykes, 1994). The difference between the chemical shifts for cubic phase-bound MetEnk and random coil chemical shifts is plotted in Figure 5.6.

Table 5.2. Unambiguous NOEs Observed for MetEnk:MO Sample '3' at 30°C

NOE Type	'From' Atom and Residue	'To' Atom and Residue	Strength (W, M or S)
Tyr^1			
intraresidue	Tyr^1 HA	Tyr^1 HD1 & HD2	W
intraresidue	Tyr^1 HB1 & HB2	Tyr^1 HD1 & HD2	M
intraresidue	Tyr^1 HD1 & HD2	Tyr^1 HE1 & HE2	M
dαN	Tyr^1 HA	Gly^2 HN	M
dβN	Tyr^1 HB1 & HB2	Gly^2 HN	W
Gly^2			
intraresidue	Gly^2 HN	Gly^2 HA	M
dNN	Gly^2 HN	Gly^3 HN	W
Gly^3			
dNN	Gly^3 HN	Phe^4 HN	M
Phe^4			
intraresidue	Phe^4 HN	Phe^4 HA	M
intraresidue	Phe^4 HN	Phe^4 HB1 & HB2	S
intraresidue	Phe^4 HA	Phe^4 HD1 & HD2	W
intraresidue	Phe^4 HB1 & HB2	Phe^4 HD1 & HD2	W
intraresidue	Phe^4 HE1 & HE2	Phe^4 HZ	W
dNN	Phe^4 HN	Met^5 HN	M
dαN	Phe^4 HA	Met^5 HN	S
dβN	Phe^4 HB1 & HB2	Met^5 HN	W
Met^5			
intraresidue	Met^5 HN	Met^5 HA	S
intraresidue	Met^5 HN	Met^5 HB1 & HB2	W
intraresidue	Met^5 HN	Met^5 HG1 & HG2	M

Table 5.3. Ambiguous NOEs Observed for MetEnk:MO Sample '3' at 30°C

Peak ID* and NOE Type	'From' Atom and Residue	'To' Atom and Residue	Strength (W, M or S)
Peak 7			
dαN	Gly^3 HA	Phe^4 HN	S
dαN (i, i+2)	Gly^2 HA	Phe^4 HN	S
Peak 9			
intraresidue	Phe^4 HN	Phe^4 HD1 & HD2	M
intraresidue	Phe^4 HN	Phe^4 HD1 & HD2	M
Peak 12			
intraresidue	Met^5 HN	Met^5 HB1 & HB2	W
intraresidue	Met^5 HN	Met^5 HE1	W
Peak 13			
intraresidue	Met^5 HN	Met^5 HB1 & HB2	W
intraresidue	Met^5 HN	Met^5 HE1	W
Peak 18			
dδN	Phe^4 HD1 & HD2	Met^5 HN	W
dεN	Phe^4 HE1 & HE2	Met^5 HN	W
Peak 23			
intraresidue	Gly^3 HN	Gly^3 HA	S
dαN	Gly^2 HA	Gly^3 HN	S

* The peak numbers listed in the *Peak ID* column correspond to those found in the peak list file *new_noe_list.xpk* that was generated using the software program *NMRView*.

Figure 7.1 from Kurt Wüthrich (1986) *NMR of Proteins and Nucleic Acids* John Wiley & Sons Inc.

Figure 5.11. Diagram Showing Sequential and Non-sequential NOEs in a Polypeptide Chain

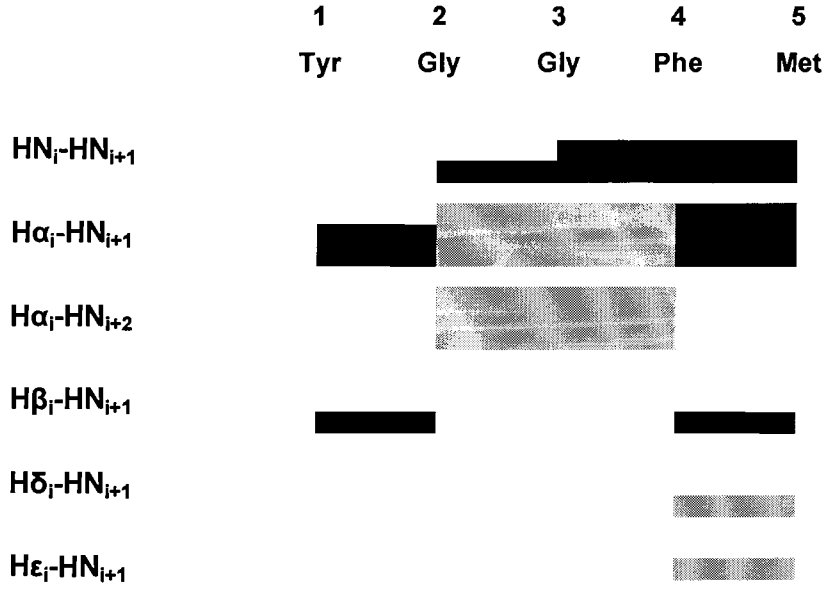


Figure 5.12. Schematic Diagram Showing the Observed NOEs for MetEnk:MO Sample '3'

The solid bars indicate unambiguous NOE peaks and the grey bars indicate ambiguous NOE peaks. The thickness of the bars indicates the relative intensity of the NOE peaks as strong, medium or weak, where the thickest bar represents the most intense NOE peak.

5.4.4 Structure Calculations

5.4.4.1 NOE Constraints

In addition to the 12 intraresidue NOE peaks that were observed in spectra collected on MetEnk:MO Sample '3', there were 7 sequential and 5 ambiguous NOE peaks. Table 5.2 contains a listing of the unambiguous NOEs, and Table 5.3 contains a list of the ambiguous NOEs that were observed for cubic phase-bound MetEnk. NOE-derived distance constraints were used to perform a series of structure calculation runs in which *all* of the unambiguous NOE constraints, and various combinations of the ambiguous NOE constraints were employed. This strategy was used to ensure that the pool of calculated structures would accurately represent all of the possible conformations for cubic phase-bound MetEnk that were consistent with the experimental data.

5.4.4.2 Calculated Structures

A total of 900 structures were calculated for cubic phase-bound MetEnk using various combinations of the NOE data provided from NMR experiments. These structures needed to be evaluated and interpreted to give an accurate picture of the structure of cubic phase-bound MetEnk. The first step in evaluating the structures generated by *CNS* was to check the structures for NOE violations. This was done by inspecting the output files generated by *CNS* for each structure calculation run to see if any of the structures were not compatible with the inputted NOE constraints. Only four of the calculated structures were found to have NOE violations. These structures were removed from the pool of calculated structures and were not used in any further analyses.

5.4.4.3 Analysis of MetEnk Structures Using Ensemble

The program *Ensemble* was used to determine which of the calculated structures were most likely to represent the cubic phase-bound conformations of the MetEnk peptide. The pool of structures that was analyzed contained the 896 structures that were calculated using NOE-derived distance constraints *and* 100 structures that were calculated without the use of distance constraints. *Ensemble* evaluated the pool of structures to determine which ones were most consistent with the observed NOE peaks and the experimentally determined chemical shifts. There were 7 high scoring structures (with weights > 0.003) in the *Ensemble* analysis conducted using NOE and chemical shift data. A backbone alignment of these structures is shown in Figure 5.13. There were many structures in the pool which were low scoring (with weights < 0.0001) according to the *Ensemble* analysis. A group of 7 structures was selected from this pool to represent the conformational diversity of the structures. A backbone alignment of these structures is shown in Figure 5.14 for comparison with the high scoring structures.

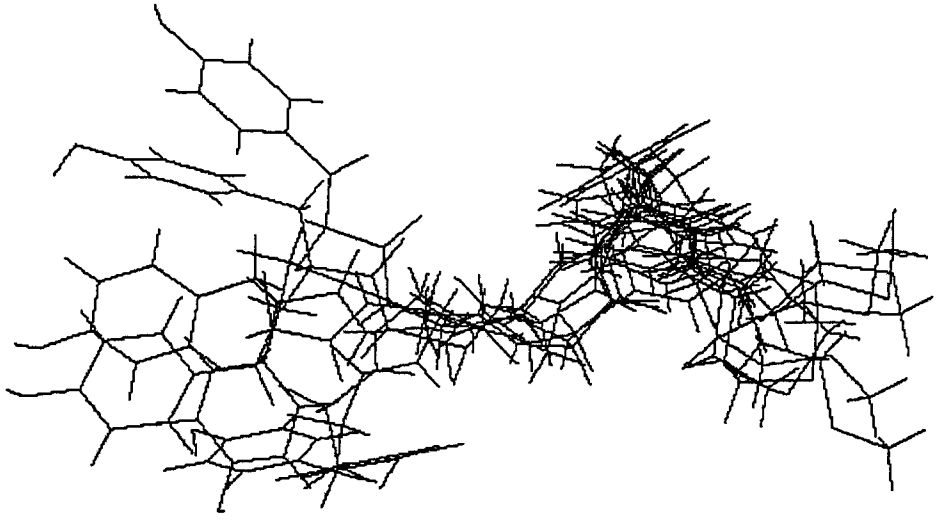


Figure 5.13. Backbone Alignment of the 7 Structures With Weights > 0.003

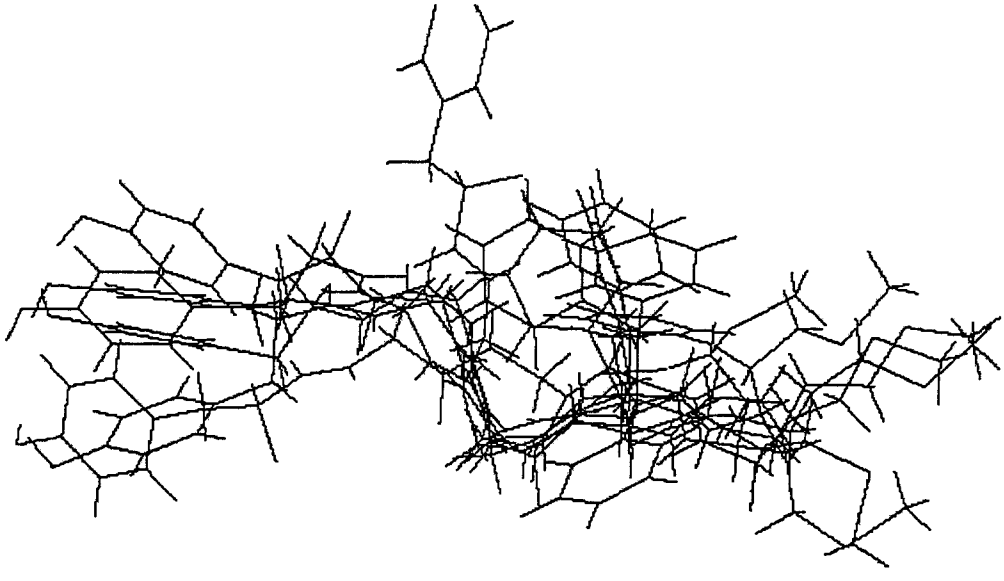


Figure 5.14. Backbone Alignment of 7 Representative Structures With Weights < 0.0001

5.4.5 NMR Measurement of Diffusion

The NMR data collected on MetEnk in water at 30°C and MetEnk:MO Sample '3' at 20°C, 30°C and 40°C were analyzed in order to measure the lipid, water and peptide diffusion rates in these samples. The experimentally determined lipid, water and peptide diffusion coefficients are summarized in Table 5.4. Data that had been collected in 2003 to measure lipid and water diffusion at 30°C on the same MetEnk:MO sample were analyzed for comparison with the data collected in 2004. The NMR spectra collected to measure lipid and water, and peptide diffusion in MetEnk:MO Sample 3' at 30°C are shown in Figure 5.15 and Figure 5.16 respectively.

For MetEnk:MO Sample '3' the highest quality peptide diffusion data were obtained at 30°C, however there was a significant amount of scatter in the data points due to the low signal-to-noise of the spectra. The peptide signal was significantly weaker than for MetEnk in water, and the peaks were much broader and less well resolved. Only the diffusion coefficient calculated for MetEnk in the MO cubic phase at 30°C will be reported because the data was the most reliable. A plot of the peak intensities as a function of the square of the gradient strength for MetEnk in the cubic phase at 30°C is shown in Figure 5.19. A plot of the normalized natural Log of the peak intensities as a function of the square of the gradient strength is shown in Figure 5.20. The slope of this line is 1.6×10^{-7} cm^2/s, the peptide diffusion coefficient for MetEnk:MO Sample '3' at 30°C. As a 'negative control' a data set was taken from a region of the spectrum without any visible peaks (~ 10 ppm). When these data were plotted it was very clear that these were random data points that were not undergoing exponential decay.

For MetEnk:MO Sample '3' it was possible to calculate the lipid and water diffusion coefficients at all temperatures. Only the first 5 or 6 data points were used to calculate the water diffusion coefficients because the signal had decayed to zero. All of the data points were used to calculate the lipid diffusion coefficients for the first three peaks in the lipid spectrum (0.8, 1.2 and 2.0 ppm) because these data were the most reliable. Comparison of the data collected on MetEnk:MO Sample '3' in 2003 at 30°C, with the data collected in 2004 at 30°C revealed changes in the peaks at 3.8 and 5.2 ppm. These changes were likely due to degradation of some of the MO molecules over time resulting in the release of 'free' oleic acid (the resonances at 5.2 ppm) and glycerol (the resonances from 3.5 - 4.2 ppm). Spectral differences were also observed in the peaks at 5.2 ppm in the spectra of MetEnk:MO Sample '3' collected at 20°C and 30°C versus the spectra collected at 40°C. Only the diffusion coefficient calculated for MO in the cubic phase at 30°C (1.7×10^{-7} cm^2/s) will be reported because the data was the most reliable. The water diffusion coefficient determined in the MO cubic phase was 3.3×10^{-6} cm^2/s, 3.9×10^{-6} cm^2/s and 4.4×10^{-6} cm^2/s at 20°C, 30°C and 40°C, respectively.

For MetEnk in water it was possible to calculate decay rates for each of the four observable peptide NH peaks. The diffusion coefficient of MetEnk in aqueous solution was calculated as the average of these four decay rates. The NMR spectra collected to measure peptide diffusion are shown in Figure 5.17 and the spectra collected to measure water diffusion are shown in Figure 5.18. The peaks undergo exponential decay and by the sixth spectrum/gradient the signal has decayed to zero, so only the first 6 or 7 data points were used to calculate the peptide diffusion coefficient. A plot of the peak intensities as a function of the square of the gradient strength for MetEnk in water at 30°C is shown in Figure 5.21. A plot of the normalized natural Log of the peak intensities as a function of the square of the gradient strength is shown in Figure 5.22. Plots of the Data points 2 - 15 were used to calculate the water diffusion coefficient in this sample. A plot of the water peak intensities as a function of the square of the gradient strength at 30°C is shown in Figure 5.23. A plot of the normalized natural Log of the water peak intensities as a function of the square of the gradient strength is shown in Figure 5.24. The peptide diffusion coefficient was found to be 6.9×10^{-6} cm^2/s and the water diffusion coefficient was found to be 2.2×10^{-5} cm^2/s at 30°C.

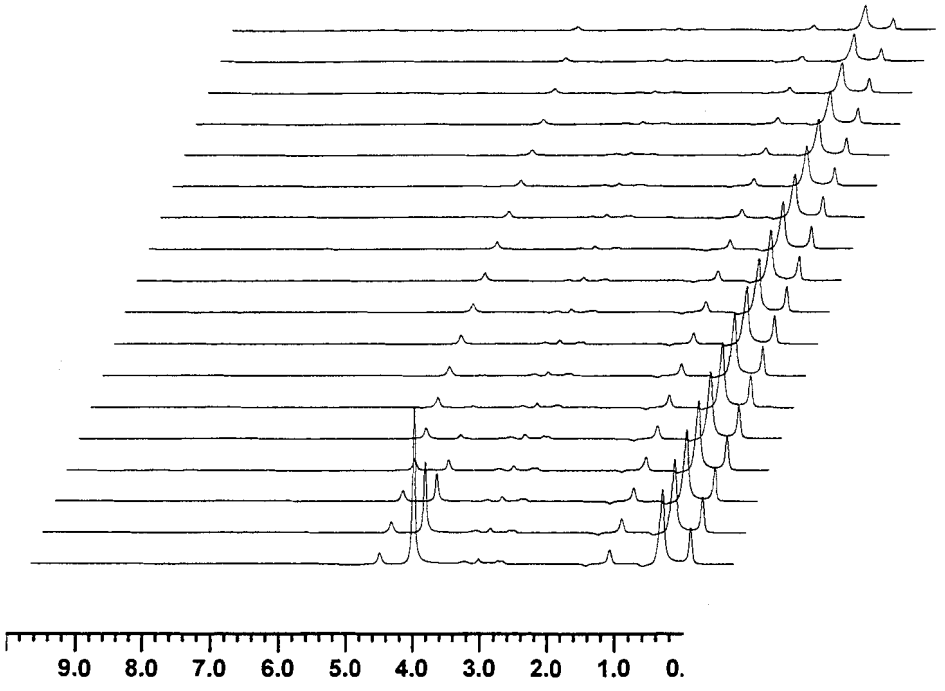


Figure 5.15. **^{1}H NMR Spectra of MetEnk:MO Sample '3' Collected to Measure Lipid and Water Diffusion at 30°C**

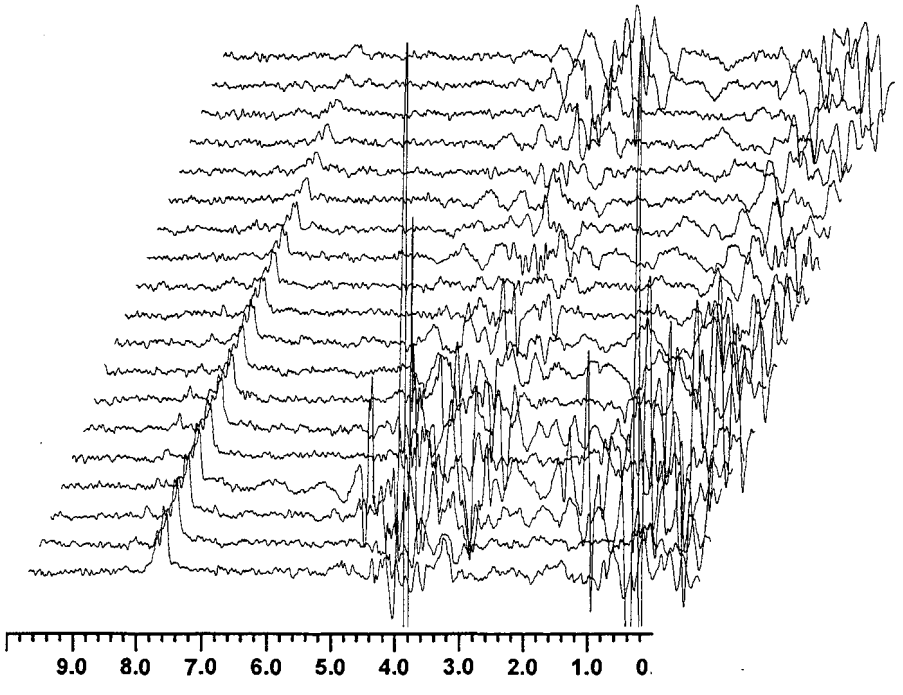


Figure 5.16. **^{15}N-edited ^{1}H NMR Spectra of MetEnk:MO Sample '3' Collected to Measure Peptide Diffusion at 30°C**

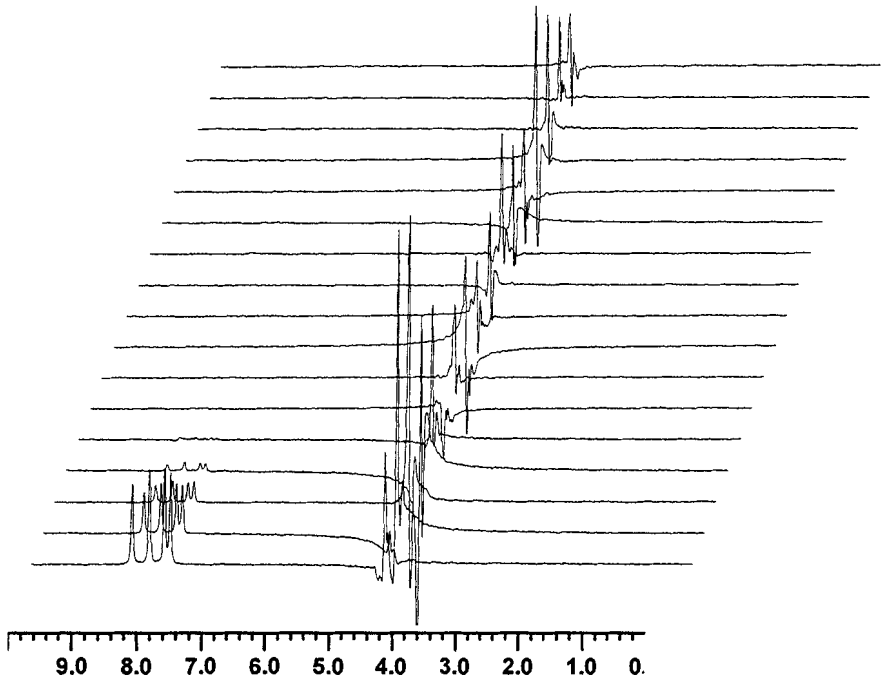


Figure 5.17. **^{15}N-edited ^{1}H NMR Spectra of the MetEnk:H_2O Sample Collected to Measure Peptide Diffusion at 30°C**

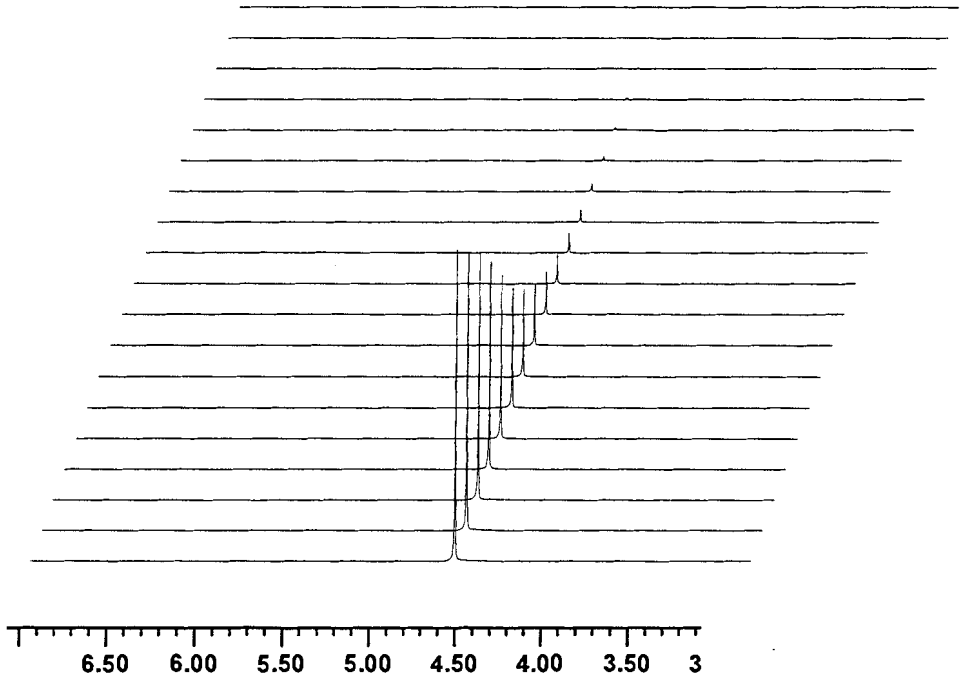


Figure 5.18. **^{1}H NMR Spectra of the MetEnk:H_2O Sample Collected to Measure Water Diffusion at 30°C**

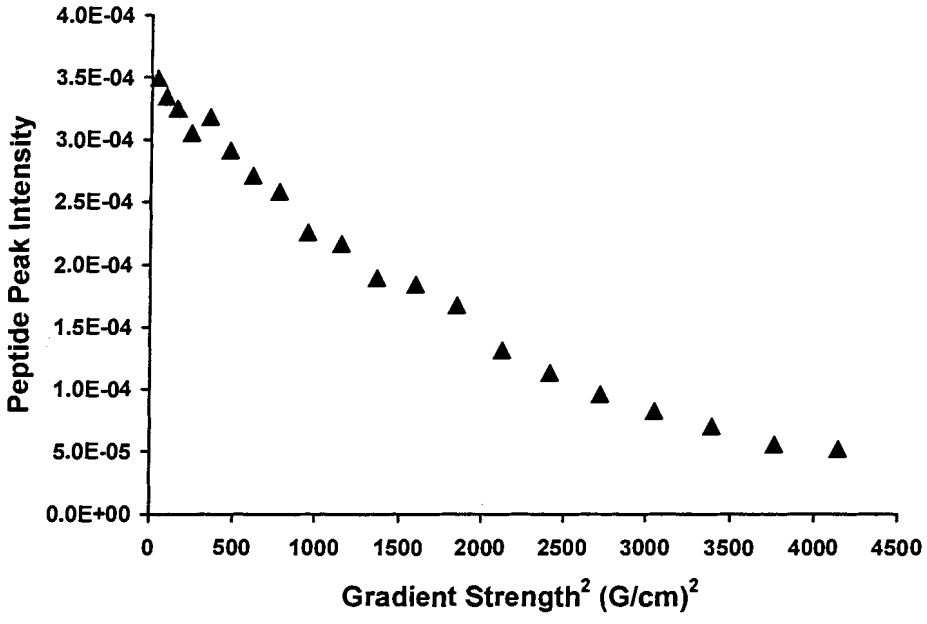


Figure 5.19. Plot of Peptide Peak Intensity versus Gradient Strength for MetEnk:MO Sample '3' at 30°C

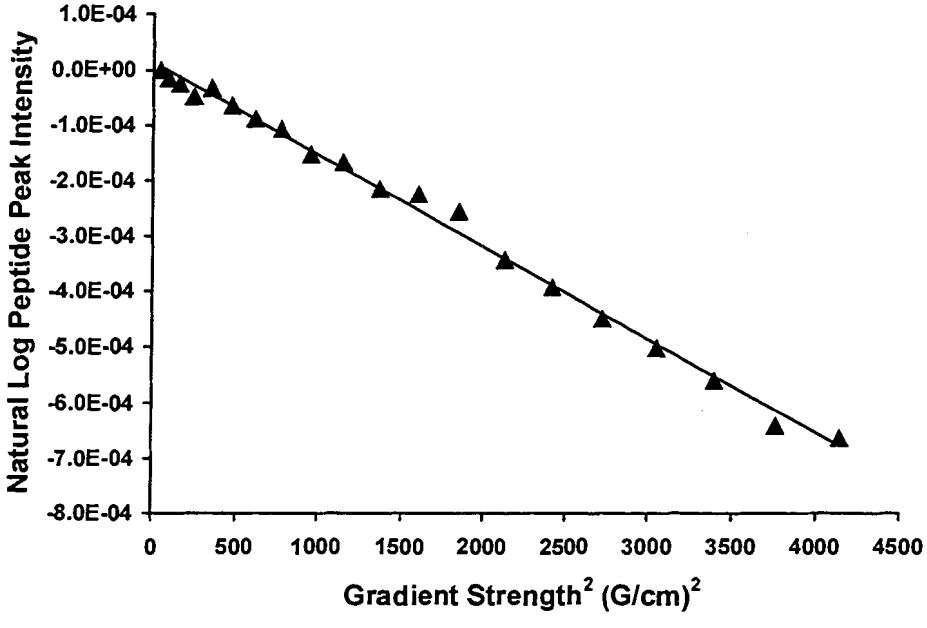


Figure 5.20. Natural Log Plot of Peptide Peak Intensity versus Gradient Strength for MetEnk:MO Sample '3' at 30°C

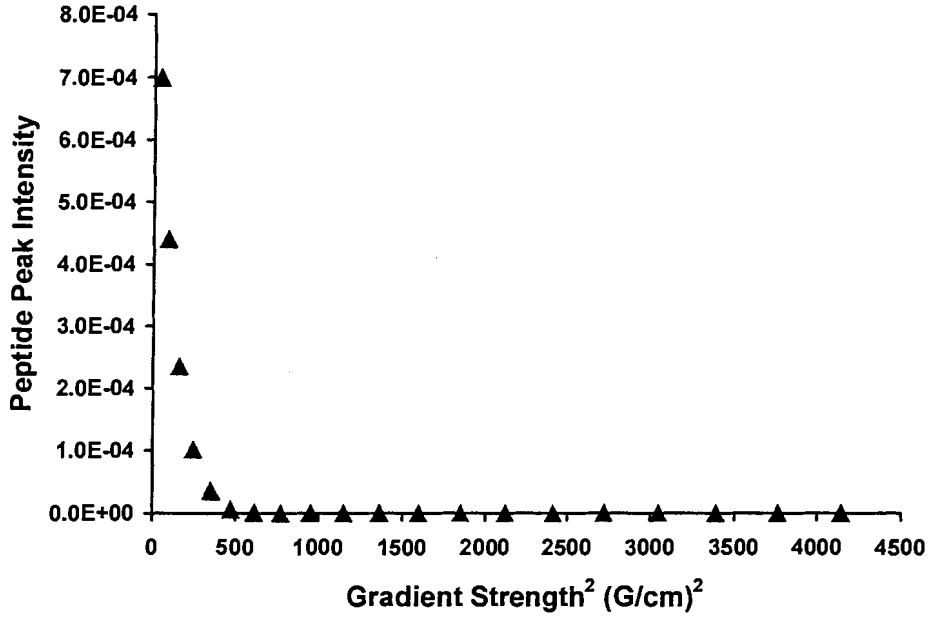


Figure 5.21. **Plot of Peptide Peak Intensity versus Gradient Strength for the MetEnk:H_2O Sample at 30°C**

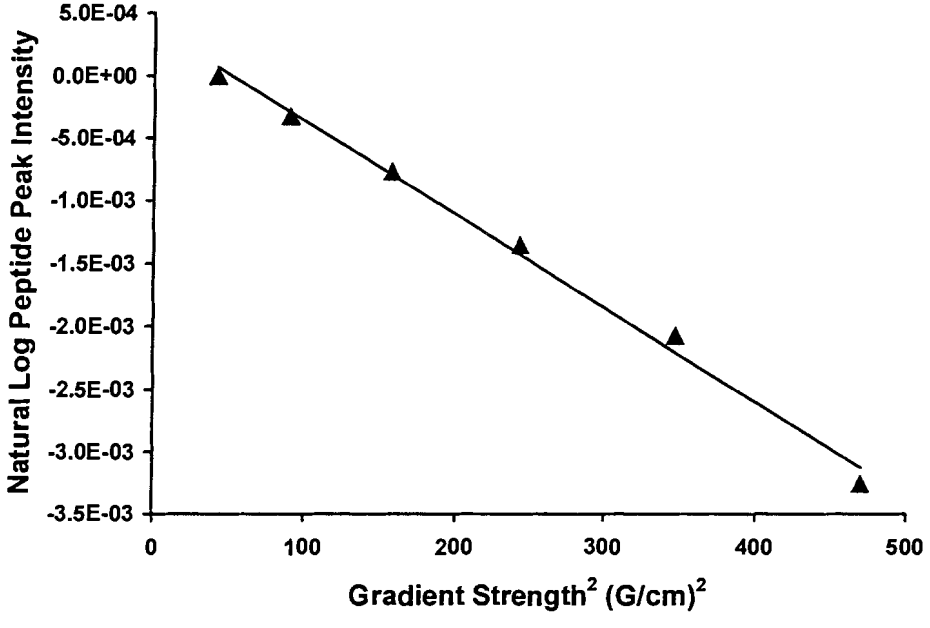


Figure 5.22. **Natural Log Plot of Peptide Peak Intensity versus Gradient Strength for the MetEnk:H_2O Sample at 30°C**

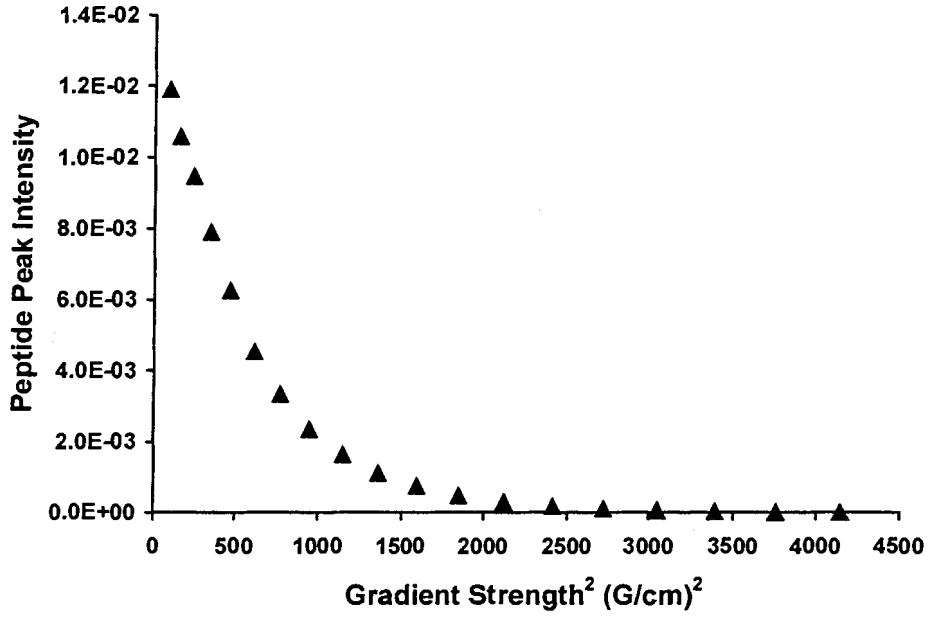


Figure 5.23. **Plot of Water Peak Intensity versus Gradient Strength for the MetEnk:H_2O Sample at 30°C**

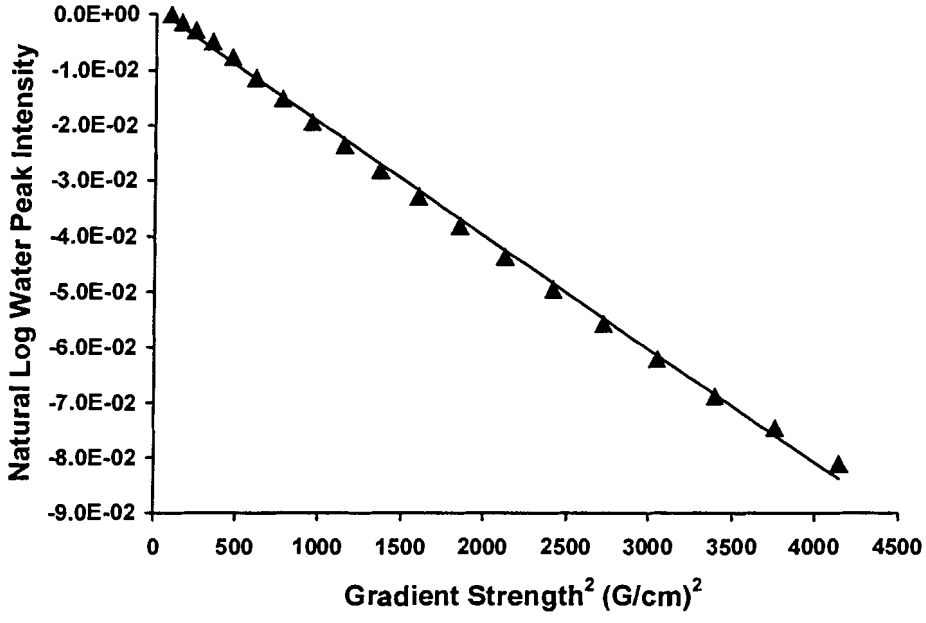


Figure 5.24. **Natural Log Plot of Peptide Peak Intensity versus Gradient Strength for the MetEnk:H_2O Sample at 30°C**

Table 5.4. Summary of Diffusion Data Collected on MetEnk in Water and in Cubic Phases

Experiment Description	Diffusion Coefficient (cm^2/s)	Temperature (°C)	δ (s)	Δ (s)
MetEnk in Water H_2O Diffusion at 30°C	$D_{WATER} = 2.2 \times 10^{-5}$	30	0.002	0.03
MetEnk in Water Peptide Diffusion at 30°C	$D_{PEPTIDE} = 6.9 \times 10^{-6}$	30	0.002	0.5
MetEnk:MO Sample '3' Lipid & H_2O Diffusion at 20°C	$D_{LIPID} = 1.1 \times 10^{-7}$ $D_{WATER} = 3.3 \times 10^{-6}$	20	0.002	1
MetEnk:MO Sample '3' Peptide Diffusion at 20°C	$D_{PEPTIDE} = 1.0 \times 10^{-7}$	20	0.002	1
MetEnk:MO Sample '3' Lipid & H_2O Diffusion at 30°C	$D_{LIPID} = 1.7 \times 10^{-7}$ $D_{WATER} = 3.9 \times 10^{-6}$	30	0.002	1
MetEnk:MO Sample '3' Peptide Diffusion at 30°C	$D_{PEPTIDE} = 1.6 \times 10^{-7}$	30	0.002	1
MetEnk:MO Sample '3' Lipid & H_2O Diffusion at 40°C	$D_{LIPID} = 2.2 \times 10^{-7}$ $D_{WATER} = 4.4 \times 10^{-6}$	40	0.002	1
MetEnk:MO Sample '3' Peptide Diffusion at 40°C	$D_{PEPTIDE} = 2.3 \times 10^{-7}$	40	0.002	1
MetEnk:MO Sample '3' Lipid & H_2O Diffusion at 30°C **one year earlier**	$D_{LIPID} = 1.7 \times 10^{-7}$ $D_{WATER} = 3.2 \times 10^{-6}$	30	0.002	1

5.5 Discussion

5.5.1 Effects of Methionine-Enkephalin on Monoolein:Water Phase Behaviour

When cubic phase samples containing MetEnk were prepared, they behaved as though they had a higher lipid content than they actually did. Since MetEnk is known to interact with the surfaces of membranes, perhaps it acts in a similar manner in the cubic phase. If MetEnk is located at the polar/apolar interface of the MO 'bilayer', as it is in bicelles, it would behave as a lipid with respect to its affects on the phase behaviour of the system (Marcotte *et al.*, 2003). In the cubic phase the highest areas of strain are at the water/lipid interface, so any molecule that interacted with the interface would have more of an effect on the phase behaviour of the system than a molecule that was for example sequestered within the cubic phase. It seemed possible to counter act this effect by increasing the hydration level of the system. A similar observation was made in studies of lysozyme in the MO cubic phase. In the phase diagram of MO:lysozyme:water the phase boundaries for the cubic phase are shifted towards higher water content than in the phase diagram of MO:water at a relatively low lysozyme concentration (Ericsson *et al.*, 1983).

5.5.2 NMR Spectra of Cubic Phase-Bound Methionine-Enkephalin

5.5.2.1 2D ^{1}H-^{15}N Correlation Spectra

Good quality NMR spectra were obtained for MetEnk in the MO cubic phase at a peptide:lipid ratio of ~ 1:1000. Sharp well resolved peaks were observed in all of the NMR spectra that were collected indicating that the peptide was experiencing rapid isotropic reorientation in the cubic phase-bound state. Since it was not necessary to employ water suppression in the NMR experiments that were conducted, no problems were encountered with variability in the intensities of the amide proton peaks. In previous NMR studies of MetEnk in lipid micelles or bicelles, it had been observed that the intensity of the Gly^{2} amide proton resonance was diminished compared to the intensity of the signal from the Gly^{3}, Phe^{4} and Met^{5} resonances when experiments involved pre-saturation of the water signal (Deber and Behnam, 1985; Marcotte *et al.*, 2004).

These data were consistent with high resolution ^{13}C solid state NMR spectra collected on LeuEnk, the sister peptide to MetEnk, bound to dihexanoylphosphatidylcholine: dimyristoylphosphatidylcholine bicelles (Sanders II and Landis, 1995). The authors of the paper were surprised at the high resolution of the spectra that were obtained as this was unprecedented for oriented membrane proteins at the time that the study was conducted. It was believed that the resolution of the spectra were likely derived from a combination of the conformational dynamics

of the bicelle-bound peptide and/or to motions of the whole peptide, including on-surface wobbling and diffusion, and rapid dissociation of the protein into isotropic solution, followed by tumbling and reassociation. Evidence supporting the presence of extensive motional averaging of the bicelle bound peptide were provided by the small magnitudes of the chemical shift anisotropies and the dipolar couplings observed for bicelle-associated LeuEnk.

5.5.2.2 2D ^{1}H-^{13}C Correlation Spectra

In the aromatic region ^{1}H-^{13}C HSQC spectrum of cubic phase-bound MetEnk it was observed that the peak intensities of the Tyr^{1} ring protons were much more intense than those of the Phe^{4} ring protons. This was likely attributable to differences in the mobility of these two side chains in the membrane-bound state which would result in differential line broadening of the peaks (Milon *et al.*, 1990). These data suggest that the Tyr residue is likely located closer to the surface of the cubic phase 'bilayer' than the Phe residue, and that the motion of the Tyr side chain is not restricted. In contrast, the peak intensities of the Phe ring protons decrease towards the end of the side chain, suggesting that this residue is likely buried well within the cubic phase 'bilayer' where its motion is restricted. Additional evidence supporting a location for Tyr^{1} near the surface of the cubic phase 'bilayer' was the presence of a stong NOE between Tyr^{1} and water. The proposed 'bilayer' surface location of Tyr^{1} is in contrast to some of the findings reported in the literature which indicated that the Tyr^{1} residue of both MetEnk and LeuEnk is located at the headgroup/acyl chain interface in phospholipid micelles and liposomes (Behnam and Deber, 1984; Milon *et al.*, 1990)

It was possible to assign all of the peptide peaks in the phase sensitive ^{1}H-^{13}C HSQC spectrum that was collected on cubic phase-bound MetEnk. Even though the peptide was fully ^{13}C-labelled, the peptide peaks were still of a much weaker intensity than the lipid peaks. This is because even at natural abundance, the amount of ^{13}C signal from MO is still about 100 × the intensity of the signals derived from the peptide. The difference in signal intensities from the lipids and the peptide were useful for initially identifying the peaks originating from the peptide. Luckily, there wasn't any overlap of any of the peptide peaks with the lipid peaks. It was therefore possible to observe and assign all of the peptide peaks that were expected to be observable in the ^{1}H-^{13}C HSQC spectrum. The phase-sensitive nature of the experiment made it possible to unambiguously assign the γ, β, and ε protons of Met^{5}, it was not however possible to obtain stereospecific assignments for the γ and β protons.

5.5.2.3 2D ^{15}N-edited TOCSY and NOESY Spectra

The information in the TOCSY and NOESY experiments made it possible to make sequence specific assignments for all for of the peaks in the spectra that were collected. Intraresidue and

sequential NOE peaks were observed in the NOESY spectrum. Unfortunately no unambiguous non-sequential NOE peaks were observed in either of the NOESY spectra that were collected (with 150 ms and 250 ms mixing times). This result was not unexpected given the inherent flexibility of a peptide only five residues in length. The absence of strong non-sequential NOE peaks in the NOESY spectra of cubic phase-bound MetEnk indicated that the peptide was probably not folded up into a compact conformation upon lipid-binding. This result was unexpected given the number of non-sequential NOEs that had been observed in previous studies (Milon *et al.*, 1990; Marcotte *et al.*, 2004). There was however a strong NOE peak between Tyr^1 and water indicating that this residue is likely in an interfacial location when bound to MO cubic phases.

5.5.2.4 2D ^{13}C-edited NOESY and HSQCNOESYHSQC Spectra

In addition to the strong NOE peak between water and Tyr^1, NOE cross peaks were seen at ~ 1.5 ppm. These crosspeaks were believed to be with the lipids, but they could also have been from the Met^5 side chain. To rule out this possibility a 2D aromatic region ^{13}C-edited NOESY spectrum, and a 2D ^{13}C-edited HSQCNOESYHSQC were collected. Any NOE peaks that were visible in the ^{13}C-edited NOESY spectrum that were not visible in the ^{13}C-edited HSQCNOESYHSQC did not originate from peptide-peptide interactions, but rather from peptide-lipid interactions. The NOE peaks at ~ 1.5 ppm were not visible in the ^{13}C-edited HSQCNOESYHSQC spectrum confirming that these NOE peaks originated from the lipids. These NOE peaks between Phe^4 and Met^5 and the oleic acid side chain of MO provide evidence for the location of these residues at the headgroup/acyl chain interface of the cubic phase.

5.5.3 Chemical Shifts

Using the information provided by the various heteronuclear NMR spectra that were collected it was possible to assign all of the observable backbone and side chain atom ^{1}H, ^{13}C and ^{15}N chemical shifts. Only small differences were observed in the chemical shifts of the free versus lipid bound peptide. These results are consistent with those published in the literature where small changes were found in the chemical shifts upon binding of the peptide to phospholipid bicelles (Marcotte *et al.*, 2004). In Figure 5.6 the chemical shifts of cubic phase-bound MetEnk were compared with the chemical shifts for polypeptides in random coil conformations. It was found that there were only minor differences between the chemical shifts for cubic phase-bound MetEnk and the chemical shifts for random coil polypeptides. This indicated that the peptide did not contain significant amounts of α and β secondary structure (Wishart, 1991).

5.5.4 Structure Calculations for Cubic Phase-Bound MetEnk

The results of the *Ensemble* analyses that were conducted on the pool of structures calculated for cubic phase-bound MetEnk are shown graphically in Figure 5.13 and Figure 5.14. The structures shown in Figure 5.13 are the 7 most highly weighted conformers (those with weights > 0.003) from the pool of calculated structures. The 7 structures shown in Figure 5.14 are representative of those conformers with weights < 0.0001. It does not appear that MetEnk adopts a well defined structure upon binding to the MO cubic phase, although there may be some conformational preference in the structures that were highly weighted in the *Ensemble* analyses. This result was not surprising given the lack of long-range NOEs.

In the recent paper in which the structures of bicelle-bound MetEnk were calculated, approximately 60 NOE derived distance constraints were used for the structure calculations (Marcotte *et al.*, 2004). Whereas only 19 unambiguous, and 6 ambiguous sequential or intraresidue NOE constraints were available for use in structure calculations for cubic phase-bound MetEnk. Given the paucity of distance constraints that were available for structure calculation purposes, it is not overly surprising that MO-bound MetEnk was not found to have a well defined structure. The observed chemical shifts for cubic phase-bound MetEnk also supported an extended conformation for the cubic phase-bound peptide. These results are interesting since MetEnk has been shown to be structured in many other lipid environments. Recently the conformations of MetEnk in zwitterionic and anionic bicelles were determined using ^{1}H NMR techniques (Marcotte *et al.*, 2004). It was found that MetEnk existed in slightly different conformations depending on the nature of the lipids it was interacting with. This result was consistent with the NOE data that was collected on these peptides in the two different types of bicelles, since many similar NOEs were observed. The differences in the NOEs observed in the zwitterionic versus anionic lipid environments were between the side chains of Tyr^1 and Met^5, Phe^4 and Met^5, and to additional NOEs that were observed between Tyr^1 and Phe^4 in the anionic bicelles. These results demonstrated that MetEnk can adopt many different conformations in membrane-mimetic environments which is consistent with the natural high flexibility of this peptide (and its close relatives). It was postulated that the structural differences observed between the conformations determined for MetEnk in anionic versus zwitterionic bicelles could result from differences in the mobility of the N-terminus (Marcotte *et al.*, 2004). This is related to natural variations in the phi and psi angles of glycine which is a much less constrained amino acid. In the case of cubic phase-bound MetEnk, the absence of detectable structure formation upon lipid-binding could be due to differences in the interactions of MetEnk with MO, a polar lipid, than with PC, a zwitterionic lipid. Previous studies have indicated that salt interferes with

the interaction of MetEnk with both charged and zwitterionic lipids (Jarrell *et al.*, 1980; Milon *et al.*, 1990). This implies that there is an electrostatic component to the interaction of MetEnk with lipids, even if they have a neutral net charge. Another factor to consider is that although the most stable conformation for the enkephalins may be a folded type of structure, this does not account for the contributions of membrane association to the overall stability of the peptide structure.

Evidence from this, and other studies conducted on MetEnk in membrane-mimetic environments indicate that MetEnk is located at the lipid/water interface. A number of intermolecular NOEs have been observed between the lipid headgroups and the peptide residues in studies conducted in phospholipid bicelles (Marcotte *et al.*, 2004). It has also been shown in NMR studies in both micellar and bicellar systems that the Gly^2 NH is affected by water suppression (Deber and Behnam, 1985; Marcotte *et al.*, 2004). This observation indicates that this residue is located closer to the membrane surface than the other residues of the peptide. Interactions of the peptide with lipid head groups would restrict the motion of the peptide in its lipid-bound form. The NMR data collected on MetEnk bound to MO cubic phases did not show an interaction of Tyr^1 with the lipid headgroups, but instead showed that this residue was quite mobile. A strong NOE peak with water provided evidence that Tyr^1 was probably located in the interfacial region of the cubic phase 'bilayer'. This difference in the mode of interaction of MetEnk with MO cubic phases versus other membrane-mimetic systems, could account for the structural differences that were observed for cubic phase-bound MetEnk. If electrostatics play an important role in the interaction of the amino terminus of the peptide with zwitterionic lipid headgroups, it is likely that this interaction also has an effect on the structure of the peptide. This influence of electrostatic interactions on the structures adopted by MetEnk in lipid environments is consistent with the role that electrostatic interactions may play in modulating the structure of MetEnk *in vivo* since the membranes of the central nervous system can be composed of up to 24 % anionic phospholipids (Marcotte *et al.*, 2004).

5.5.5 Methionine-Enkephalin Diffusion in the Monoolein Cubic Phase

NMR was used to measure the diffusion of the lipid, peptide and water components of MetEnk-containing MO cubic phases. Since all plots of the normalized natural log of the signal intensity versus the square of the gradient strength were linear, there was no influence of intermolecular NOEs on the diffusion results (Chen and Shapiro, 1999; Lucas *et al.*, 2003). The diffusion coefficient that was observed for cubic phase-associated MetEnk ($D_{observed}$) represented an average of the diffusion coefficients for the free (D_{free}) and the bound (D_{bound}) peptide because the peptide was in rapid exchange between the free and bound states on the NMR time scale (Whitehead *et al.*, 2001). In this two site model, D_{bound} corresponds to the lipid diffusion

coefficient and D_{free} corresponds to the diffusion coefficient of the peptide in water (Stilbs, 1982). The following equation has been used in studies of micelle-associated neuropeptides to describe the relationship between $D_{observed}$ and the proportions of free and lipid-bound peptide (Gao and Wong, 1998; Whitehead *et al.*, 2001):

$$D_{observed} = X_{free} D_{free} + X_{bound} D_{bound} \quad [5.1]$$

In this equation, X_{free} represents the mole fraction of free peptide and X_{bound} the mole fraction of bound peptide. The proportion of peptide that is lipid-associated can be determined using the values measured for $D_{observed}$, D_{free} and D_{bound}. In solutions containing micelles or bicelles, the diffusion of 'free' unbound peptide will be affected by the presence of the micelles or bicelles which will impede the diffusion of the peptide in the solution. To compensate for this impediment of free diffusion, an 'obstruction' factor must be incorporated into calculations of the proportion of lipid-associated peptide in micellar and bicellar solutions (Gaemers and Bax, 2001).

The diffusion rate of free MetEnk in aqueous solution is ~ 31.8 % of that of water molecules in the same solution. This value was used to calculate the 'expected' diffusion coefficient for unbound MetEnk in the MO:water cubic phase to be $\sim 1.2 \times 10^{-6}$ cm^2/s at 30°C. The experimentally determined values for $D_{observed}$, D_{free} and D_{bound} (shown in Table 5.4) were used to calculate the proportion of MetEnk that was bound to MO cubic phases at 30°C. It was found that ~ 100 % of the MetEnk in the sample was lipid bound. Since the peptide:lipid ratio in these samples was very high, (~ 1:1000) it was not surprising that such a high proportion of the peptide was lipid-associated. It is interesting to note that the high level of lipid association is not concomitant with the folding of the peptide into a 'structured' conformation. The lipid and water diffusion coefficients determined for MetEnk-containing MO cubic phases were consistent with those measured for the transmembrane peptide-containing MO cubic phases discussed in Section 4.4.4. The relatively high amount of lipid-associated peptide is consistent with results published in the literature for MetEnk and LeuEnk. In NMR studies of LeuEnk in mixtures of 1,2-dihexanoyl-*sn*-3-glycerophosphocholine (DHPC) and DMPC (1:2.0) at 40°C, it was found that the peptide was greater than 95 % associated with the lipid interface (Sanders II and Landis, 1994). The molar ratio of peptide:DMPC in these samples was roughly 1:10, and the total lipid content of the samples was ~ 25 % (wt/vol). The samples were prepared in 70 mM sodium phosphate, 30 mM KCl in D_2O at a pD of 7.0. In another NMR study conducted on MetEnk in the presence of zwitterionic DMPC, or anionic DMPG-doped DMPC bicelles at a peptide:lipid molar ratio of 1:25, it was found that 65 % and 59 % of the peptide was lipid bound, respectively. (Marcotte *et al.*, 2004).

Chapter 6: Discussion, Conclusions and Future Work

6.1 Discussion

My thesis research has shown that both the transmembrane peptides TMK and alamethicin, and the membrane surface-associating peptide MetEnk, can be successfully incorporated into the MO cubic phase and studied using solution NMR techniques. The results that were obtained provide valuable insights into the interactions and behaviour of membrane peptides, and by inference membrane proteins, incorporated into the bicontinuous cubic phases formed by mixtures of MO and water. This knowledge will further our understanding of lipid-peptide/protein interactions, lipid polymorphism and the mechanisms by which cubic phases can mediate the crystallization of membrane proteins.

6.1.1 NMR Studies of Transmembrane Peptides in MO Cubic Phases

In solution NMR studies conducted on transmembrane peptides incorporated into the MO cubic phase, it was found that only the signals originating from the amino acid residues located at the N- and/or C-termini of the peptides could be observed. These residues would be located in the interfacial regions of the cubic phase 'bilayer' if the peptides were incorporated into the cubic phase in the predicted transbilayer orientation. The NMR results that were obtained were somewhat unexpected because the lipid spectra collected on MO cubic phases yielded narrow well resolved peaks. By analogy, it was believed that NMR spectra of hydrophobic or amphiphilic peptides incorporated into the MO cubic phase would have similarly narrow well resolved peaks. Although the spectra collected on cubic phase-bound transmembrane peptides did contain *some* narrow peaks, most of the expected peaks were too broad to be observable using solution NMR techniques. The peaks that could not be observed were the ones originating from residues located in the membrane-spanning regions of the peptides. Broad peaks are known to be

observed in NMR spectra when molecular motions are slow, or in cases where a protein exists in an anisotropic state (Sanders II and Landis, 1995). There were a number of scenarios where one, or both, of these conditions could be met.

6.1.1.1 Peptide Aggregation

The peptides would experience anisotropic motions if they had formed molecular aggregates that were too large to 'tumble' rapidly within the milieu of the cubic phase. Although the high lipid-to-peptide ratios in the samples (1800:1 and 270:1 for TMK and alamethicin, respectively), made it unlikely that this would occur, the *possibility* that aggregation could have occurred could not be ruled out since both of the peptides were known to self-associate under certain conditions. The TMX-1 peptide, which the TMK peptide sequence was based on, formed soluble aggregates in aqueous solution (Wimley and White, 2000). Alamethicin was also known to self-associate as it is a channel forming peptide whose self-association is dependent on the presence of a transmembrane potential (the primary driving force behind peptide self-association), as well as on pH and peptide concentration (Franklin *et al.*, 1994).

Since both NMR samples were prepared by dissolving the peptides in molten lipid *prior* to the addition of water, it was unlikely that the peptides would have formed aggregates in aqueous solution. The alacoil motif (Ala at i, $i + 7$ positions along the helix) which was present in the 'parent' TMX-1, and could have potentially favoured the formation of anti-parallel α-helical dimers of the peptide, was intentionally disrupted in the design of the TMK peptide, so at the high lipid-to-peptide ratio in the sample (1800:1), peptide self-association would not be favoured (Wimley and White, 2000). In previous studies of alamethicin in detergent micelles at similar concentrations to the one used in the cubic phase sample, the peptide was monomeric and it was found to be associated with the interior of the detergent micelles (Franklin *et al.*, 1994). For these reasons it was believed that monomeric alamethicin would be incorporated into the cubic phase 'bilayer' in a transmembrane orientation.

Hydrophobic Mismatch

There are other mechanisms that could drive peptide aggregation within the cubic phase. One of these is hydrophobic mismatch. If the hydrophobic thickness of the MO 'bilayer' was not equal to the hydrophobic length of the peptide, this could favour peptide self-association (de Planque *et al.*, 2002; de Planque and Killian, 2003; Killian, 2003; Orzaez *et al.*, 2005). Although the hydrophobic lengths of the peptides were calculated to be comparable to the hydrophobic thickness of the MO 'bilayer' based on data provided from X-ray diffraction studies, it was still possible for mismatch to occur depending on the conformations of the peptides in the cubic phase environment. This was one of the reasons why alamethicin was chosen as the second peptide to

study, because it had a hydrophobic length that was shorter than that of the TMK peptide. If hydrophobic mismatch had driven the aggregation of the TMK peptide, a shorter peptide would be less likely self-association via the same mechanism. Another advantage of studying alamethicin was that it was an uncharged peptide, so it would be less likely to cause perturbation of the structure of the cubic phase, since the presence of charged moieties has been shown to increase the lattice dimensions of the cubic phase (Giorgione *et al.*, 1998; Aota-Nakano *et al.*, 1999). However, the NMR spectra that were obtained for both peptides were similar, which indicated that the results were not 'artefacts' caused by peptide aggregation or perturbation of the cubic phase, but were rather due to the inherent spectral properties of transmembrane peptides incorporated into MO cubic phases.

6.1.1.2 Lipid and Peptide Lateral Diffusion

The lateral diffusion of a peptide in the cubic phase is more restricted than that of a lipid (Cherry, 1979; Lindblom and Rilfors, 1989). In addition to the differences in size between a lipid and a peptide molecule, there are differences in the localization of these molecules within the cubic phase and this will affect their diffusion. A lipid molecule 'sits' in one leaflet of the cubic phase 'bilayer' and is free to diffuse independently of the other leaflet of the 'bilayer'. However, the diffusion of a transmembrane peptide *is* dependent on both leaflets of the bilayer since its N- and C-termini are 'tethered' at the aqueous interface of separate water channels, so diffusion of the peptide depends on the coordinated movement of both termini of the peptide through the cubic phase structure. Unlike TMK, alamethicin does not contain any charged residues so its diffusion would be slightly less restricted due to fewer interactions in the interfacial region of the cubic phase. This was another reason why alamethicin was selected as the second transmembrane peptide for study. However, no significant differences could be detected between the NMR spectra collected on the alamethicin versus the TMK peptide, or in their diffusion rates within the MO cubic phase. It would be interesting to compare the lipid diffusion rates of bipolar lipids in the bicontinuous cubic phase with those of the transmembrane peptides in the bicontinuous cubic phase, since the diffusion of both of these types of molecules would be similarily restricted (Luzzati, 1997).

Another factor that could influence peptide diffusion within the cubic phase would be irregularities or 'defects' in the cubic phase structure, such as transient alterations in the hydrophobic thickness of the 'bilayers' or variations in the dimensions of the water channels. Depending on the nature of such 'defects', there could be an energetic penalty to either partition a hydrophilic residue into the hydrophobic interior of the 'bilayer' or to partition a hydrophobic residue into the interfacial or aqueous regions of the cubic phase. The presence of 'defects' could

act to reduce the areas that would be accessible to a diffusing peptide compared to a diffusing lipid. In the case of alamethicin, peptide diffusion would be less likely to be affected by such 'defects' because the termini of the peptide are not anchored at the aqueous interface, unlike the TMK peptide which has a tryptophan anchored N-terminus and a charged C-terminus. However, it is unlikely that there are any irregularities in the structure of the cubic phase since no anisotropic signal is present in the NMR spectra of cubic phases and the 'liquid-like' properties of the cubic phase requires that all molecules have equivalent packing conditions (Hyde *et al.*, 1984).

6.1.1.3 Lipid and Peptide Conformational Flexibility and Axial Rotation

Broad lines are observed in NMR spectra when molecular motions are slowed. In the case of peptides inserted into bilayers in a transmembrane orientation, both rotational and conformational motions of the peptide will be decreased resulting in increases in the observed peptide linewidths (Sanders II and Landis, 1995). It is also likely that the rotational and conformational motions of the lipids in the immediate vicinity of the bound peptides will be dampened. Both of these restrictions of molecular motion will act to increase the time required for sample relaxation to occur, a process that plays an important role in determining whether or not the signals from lipid-associated peptide residues will be observable in NMR spectra (Sanders II and Landis, 1995). It is likely that restriction of the conformational and rotational motions of residues located within the cubic phase 'bilayer' is the reason why only the signals originating from the most flexible residues in the peptides (those residing at the N- and C-termini of the peptide), are observable using solution NMR techniques.

The effect of lipid association on the NMR spectra of membrane-binding peptides and proteins cannot be predicted. In solution and solid state NMR studies of micelle and bicelle-associated peptides it has been found that some peptides yield high quality NMR spectra with sharp well resolved peaks upon lipid binding, whereas other peptides do not (Sanders II and Landis, 1995; Andersson and Mäler, 2002; Biverståhl *et al.*, 2004). It appears that spectral quality is dependent on the nature of the particular peptide or protein that is studied since this phenomenon has been observed for both transmembrane and surface-binding peptides. It is believed that this effect is modulated, at least in part, by a local ordering of the lipids in the immediate vicinity of the lipid-bound peptide (Andersson and Mäler, 2002). In one solid state NMR study where spectra were acquired on a number of different membrane peptides and proteins that had been reconstituted into lipid bicelles, a considerable amount of variation was observed in the quality of the NMR spectra that could be collected (Sanders II and Landis, 1995). Alamethicin was one of the peptides that was studied. A ^{13}C NMR spectrum of alamethicin was

recorded at pH 5.5 on a sample with a peptide:DMPC molar ratio of 1:36 at 40°C. It was found that the bicelles oriented as expected in the magnetic field, however the ^{13}C NMR spectrum was somewhat noisy and had broad aromatic resonances. In contrast, the ^{13}C NMR spectrum of LeuEnk (the sister peptide to MetEnk), was recorded at pHs between 6 - 7.3 on samples with a peptide:DMPC molar ratio of < 1:8 at temperatures ranging from 33 - 50°C. The bicelles were found to orient normally and good quality ^{13}C NMR spectra could be recorded. These conditions were very similar to the ones used for the collection of NMR data on alamethicin in the MO cubic phase. It was interesting to note that although the spectra of alamethicin that were obtained in bicelles were of better quality than the spectra observed in the cubic phase, in both systems the spectra observed for the enkephalins were of similar high quality (Sanders II and Landis, 1995; Marcotte *et al.*, 2004). This result implied that despite the obvious differences in the mode of membrane-association between these two peptides (transbilayer versus membrane surface-associated), there are likely inherent differences in the interactions of these two peptides with lipids in the membrane-bound state.

6.1.2 Methionine-Enkephalin Interactions With MO Cubic Phases

6.1.2.1 Effects on Monoolein:Water Phase Behaviour

The incorporation of MetEnk into the MO:water cubic phase altered the phase behaviour of the system as judged by comparison with the phase behaviour of 'pure' non-peptide containing mixtures of MO and water (see Figure 2.1 for the MO:water phase diagram). It was found that even at low MetEnk concentrations where the lipid:peptide molar ratio was 1000:1, homogeneous cubic phases were only formed after additional water had been added to the sample. As the peptide concentrations of the samples were increased, larger amounts of water needed to be added in order to facilitate cubic phase formation. The sample that was prepared with the highest peptide concentration (containing a lipid:peptide molar ratio of 180:1), did not form a uniform cubic phase, even after additional water had been added. It appeared that the sample contained predominantly cubic phase lipid, with the coexistence of regions of lamellar phase lipid. This effect of MetEnk on the formation of MO:water cubic phases can be explained in terms of lipid molecular shape theory.

The addition of a membrane-surface binding peptide to the MO cubic phase will increase the lateral packing pressure in the interfacial region of the cubic phase 'bilayer'. This overall increase in the molecular volume occupying the interfacial region of the bilayer will not be compensated for by a concomitant increase in the molecular volume *within* the bilayer since MetEnk only interacts with the surfaces of membranes. As more and more peptide becomes bound to the bilayer, it will begin to affect the properties of the cubic phase as the pressure profile in the

bilayer becomes significantly altered. A point will be reached where the bilayer cannot incorporate any more peptide while maintaining its current morphology. At this point 'saturation' is reached and excess peptide remains in the aqueous phase until the peptide concentration becomes high enough to induce the formation of lamellar phases. This was seen in MetEnk:MO Sample '1' which had a lipid:peptide ratio of 180:1 where the coexistence of the lamellar and cubic phases were observed. It is conceivable that the effect of MetEnk on lipid packing within the cubic phase would be similar to the effect of magainin on PC membranes, however due to differences in the molecular shapes of the lipids the effect on phase behaviour would be opposite.

The magainin peptides have been shown to induce nonbilayer structures in PC membranes at concentrations where the peptides are biologically active, which is consistent with the membrane lytic activity of these peptides (Bechinger, 1997). These peptides adopt α-helical structures upon membrane binding and insert into bilayers in an orientation where the helix axis is parallel to the bilayer surface. Molecular modelling has shown that the diameter of the helix (10 - 12 Å) is insufficient to extend to the midplane of the bilayer in which it is inserted (Bechinger, 1997). Consequently this creates a distortion of the bilayer that extends over a diameter of up to 100 Å and causes a decreased average bilayer thickness (Ludtke *et al.*, 1995). Above certain peptide concentrations a bilayer structure can no longer be maintained and a phase transition occurs. In the case of the MO cubic phases, the lipids can be described as cone shaped, i.e. they have a relatively small headgroup in comparison to the cross-sectional area of their acyl chain. The effect of MetEnk binding to the cubic phase would be to increase the effective cross-sectional area in the headgroup region of the 'bilayer' favouring the formation of lamellar phases.

It is interesting to note that in studies where DOPC (a lamellar phase forming lipid with a cylindrical averaged shape) was incorporated into the MO cubic phase, the lattice parameter of the cubic phase increased (Cherezov *et al.*, 2002). It was believed that this was due to the effect of DOPC in the cubic phase 'bilayer' where the presence of a cylindrical shaped lipid would act to decrease the degree of curvature at the bilayer/water interface. In order for a cubic structure to form, the cubic lattice must expand. This can only occur if excess water is present. Perhaps the binding of MetEnk alters the phase behaviour of the MO cubic phase in a similar manner. If this were the case, then in the presence of excess water the MetEnk-containing cubic phases that were studied would be able to swell and form uniform cubic phases in the presence of higher concentrations of peptide.

The effect of the peptide WLFLLKKK on monoolein phase behaviour was studied using X-ray diffraction techniques (Masum *et al.*, 2003). It was found that this peptide could induce the formation of lamellar phases in monoolein:water mixtures that would form D-type cubic phases

in the absence of peptide. The authors believed that this effect was dependent on the electrostatic interactions of the peptide in the cubic phase, as well as on the effects of the bound peptide on the spontaneous curvature of the leaflets of the monoolein 'bilayer'.

In addition to the ability of MetEnk to alter the phase behaviour of mixtures of MO and water, saturation in the amount of MetEnk that could become associated with the MO cubic phase was observed in MetEnk:MO Sample '2', a sample with a lipid:peptide molar ratio of 254:1. The spectrum of this sample (shown in Figure C 3 in Appendix C) had two populations of peaks, one set of peaks were similar to those observed for MetEnk in aqueous solution and the other set of peaks were similar to those observed in the spectrum of MetEnk:MO Sample '3' where ~ 100 % of the peptide was lipid-associated. A similar observation had been made in studies of membrane-associating polypeptides in bicellar solutions where it was found that the binding of peptides to bicelles could be saturatable (Sanders II and Landis, 1995). In the bicellar system the peptides and lipids had been co-dissolved in solvent prior to hydration to ensure proper mixing of the lipids with peptides. In spite of this, a certain quantity of peptide was present in the aqueous milieu in an aggregated form when it was expected that *all* of the peptide would have been bicelle-associated.

6.1.3 Modelling of Peptide Association With Monoolein Cubic Phases

Molecular dynamics simulations of the TMK, alamethicin and MetEnk peptides were conducted in MO bilayers in collaboration with Peter Tieleman in the Department of Biological Sciences at the University of Calgary. The simulations of the peptides in MO bilayers were used as models for the interactions of the peptides with the 'bilayers' in MO cubic phases since the thickness of MO bilayers does not vary between the lamellar and cubic phases (Marrink and Tieleman, 2001; Cherezov *et al.*, 2002). These simulations were conducted in hydrated MO bilayers for 1 ns. The results of the simulations are shown in Figure 6.1. The positions of the peptides within the bilayers are consistent with the NMR data that were collected on these peptides in MO cubic phases. The peptide residues that were NMR observable are located in the interfacial regions of the MO bilayers, whereas the unobservable residues are sequestered within the bilayers. Given the position of alamethicin within the MO bilayer, it is actually quite remarkable that NMR signals could be observed for any of the residues in this peptide.

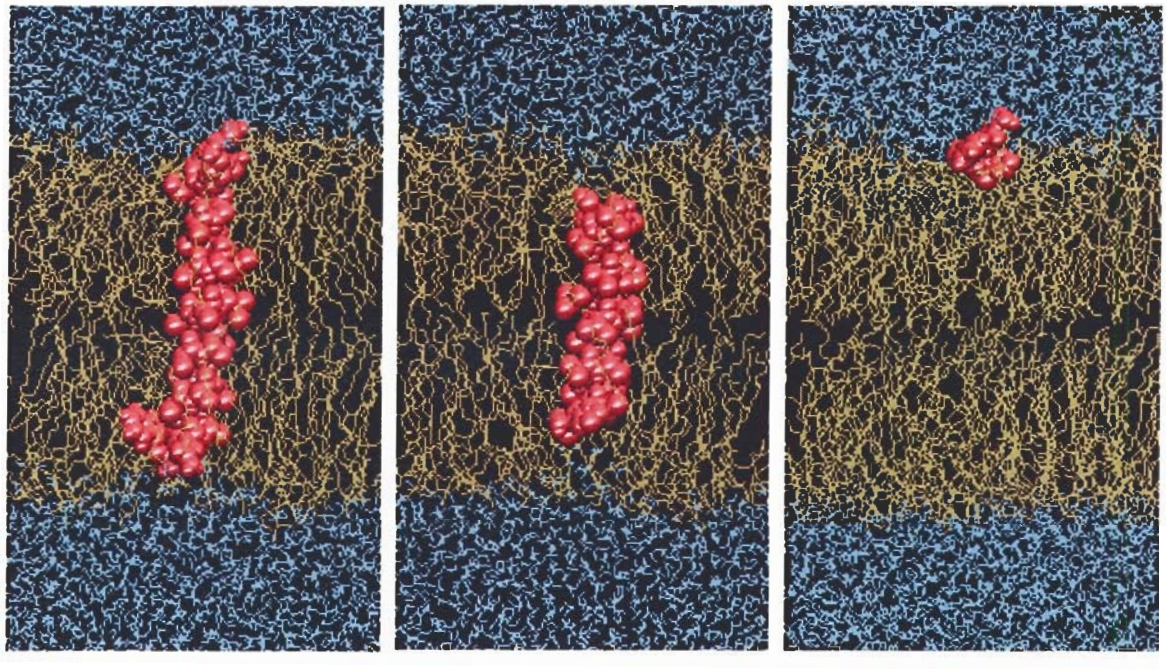


Figure 6.1. **1 ns Snapshots From Molecular Dynamics Simulations of the TMK, Alamethicin and MetEnk Peptides, Respectively, in Monoolein Bilayers**

6.1.3.1 Generation of Membrane Protein Crystals in Lipid Cubic Phases

There has been, and still is, a considerable amount of interest in the application of cubic phases as membrane-mimetic environments for the growth of membrane protein crystals. This is evident from the number of papers discussing the mechanisms and application of lipid cubic phases to membrane protein crystal generation that have been published in recent years (Rummel *et al.*, 1998; Nollert *et al.*, 1999; Ai and Caffrey, 2000; Pebay-Peyroula *et al.*, 2000; Rouhani *et al.*, 2002; Chupin *et al.*, 2003; Gohon and Popot, 2003; Misquitta and Caffrey, 2003; Sennoga *et al.*, 2003; Qutub *et al.*, 2004). Despite the high level of activity in this field, only a few membrane proteins have been successfully crystallized using lipid cubic phases. These proteins include: bacteriorhodopsin, the acetylcholine receptor, the photosynthetic reaction center from *Rhodobacter sphaeroides* and brome mosaic plant virus (Belrhali *et al.*, 1999; Casselyn *et al.*, 2001; Katona *et al.*, 2003; Paas *et al.*, 2003; Schenkl *et al.*, 2003). To better understand cubic phases and their applications in crystallography, studies of the effects of detergents, proteins, various hydrophobic and hydrophilic molecules and hydration, on the phase behaviour of MO:water systems have been conducted (Jeong *et al.*, 2002; Navarro *et al.*, 2002; Barauskas, 2003; Lendermann and Winter, 2003; Sparr *et al.*, 2004). The application of MO cubic phases for the crystallization of soluble proteins has also been investigated (Tanaka *et al.*, 2004).

Many researchers have developed models to explain the mechanism of protein crystal nucleation and growth in lipid cubic phases, as this process is not well understood. In the case of the crystallization of bacteriorhodopsin, it has been proposed that the major crystallization-inducing factor is the change in membrane structure that occurs in the cubic phase upon dehydration modulated through the addition of salt (Grabe *et al.*, 2003). Bacteriorhodopsin crystals were generated by solubilization of the protein in MO cubic phases, followed by dehydration of the cubic phase by exposure to salt (which caused shrinking of the dimensions of the unit cell of the cubic phase) and resulted in clustering of the proteins in locally flattened regions of the cubic phase 'bilayer'. The role of salt in the crystallization process is two-fold, it is necessary to dehydrate the cubic phases, and it also plays a crucial role in protein crystal formation by masking the electrostatic interactions between the protein molecules in the closely packed crystal (Nollert *et al.*, 2001).

Attempts have been made to extend the applicability to this technique to the crystallization of a wider range of membrane proteins by altering the properties of the MO cubic phase through the addition of other lipid components. Since MO itself is not a common membrane component, it was thought that perhaps this was limiting the range of membrane proteins that could be crystallized from the MO cubic phase. Monoolein cubic phases containing DOPC, DOPE,

DOPS, cardiolipin, lysoPC, a polyethylene glycol-lipid, 2-monoolein, oleamide and cholesterol were studied using X-ray diffraction techniques (Cherezov *et al.*, 2002). It was found that all of these lipids could be incorporated into the MO cubic phase to some extent without altering the cubic phase. All of the lipids, except for the anionic lipids and lysoPC, could be incorporated to 20 - 25 mol % without altering the phase behaviour of the system. Above this 'concentration limit', the incorporated lipids would induce a phase transition, or they would saturate the cubic phase and become separated out as crystals. The data provided by studies such as these could be used to rationally design cubic phases with different lipid profiles to match the specific lipid requirements of the target membrane protein to be studied.

6.2 Conclusions

The bicontinuous cubic phases formed by mixtures of MO and water are suitable membrane-mimetic environments for solution NMR studies of membrane surface-associating peptides, and potentially proteins. They may also be suitable environments for studies of the interfacial regions of cubic phase bound-transmembrane peptides and proteins depending on the effects of the incorporated polypeptide on the phase behaviour of the system. The primary advantages of employing cubic phases as membrane-mimetic environments for studies of membrane peptides and proteins are the exceptional stability of the samples (> 1 year), and the high lipid-to-peptide/protein ratios that can be maintained in the samples.

6.3 Future Work

It would be interesting to follow up on a number of aspects related to the application of MO cubic phases in NMR studies of membrane peptides and proteins.

6.3.1 Solution NMR

Although the transmembrane regions of membrane proteins cannot be observed by solution NMR in the cubic phase, the interfacial and aqueous regions are NMR observable. For this reason it would be interesting to study the structure of the extramembranous domains of monotopic membrane proteins. In this type of study the structure of a domain could be investigated *with* its membrane-binding domain intact. Studies of this type are important because the structures of many extramembranous membrane protein domains have been determined in the *absence* of the membrane spanning region, i.e. the membrane spanning domain had been removed to improve protein solubility and amenability to crystallization. Membrane proteins with two membrane spanning domains could also be studied in the cubic phase. It would be of particular interest to characterize the structure of an extramembranous domain that was anchored in the membrane by

two membrane spanning helices. Since cubic phase samples containing membrane peptides were prepared by simply dissolving the peptide in molten lipid prior to hydration, a modified sample preparation protocol would need to be developed for studies of whole intact proteins. It is likely that a protocol similar to the ones used for the incorporation of membrane proteins into cubic phases for the purposes of crystallization could be employed.

It would also be possible to study receptor-ligand interactions, protein-protein, or protein-lipid interactions within the cubic phase. For example, an opiate receptor could be incorporated into the cubic phase, and the structure and interactions of the *ligand* (small molecules or water soluble peptides) for the receptor could be studied using solution NMR techniques. It would be possible to study the protein-protein interactions of a water soluble protein with an integral membrane protein that had been reconstituted into the cubic phase. Specific lipids that are known to act as receptors for certain proteins could be incorporated into the cubic phase. NMR could then be used to study the interactions of these proteins with their target lipids (Rosen and Liao, 2003).

6.3.1.1 Surface-Associating Membrane Peptides and Proteins

Membrane proteins which associate with bilayers through lipid anchored, or membrane-binding domains such as amphipathic helices could be studied in the cubic phase. For these types of studies it would be recommended that the phase behaviour of the system be characterized in the presence of varying amounts of protein to determine how much protein can be incorporated before alterations in the phase behaviour of the system are observed. Since the cubic phase can incorporate other lipid constituents up to ~ 20 mol %, it would be possible to study the effects of different lipids on membrane association and on protein structure. Functional studies of proteins can also be conducted in the cubic phase due to its optical transparency and isotropic nature (Giorgione *et al.*, 1998; Navarro *et al.*, 2002; Liu and Caffrey, 2005).

Methionine-Enkephalin

It would be interesting to study the structure of MetEnk in cubic phases with varying lipid compositions since it appears that the interaction of the N-terminus of the peptide with charged moities in the headgroups of zwitterionic and anionic lipids plays an important role in modulating the structure of this peptide in membrane environments. It has been shown that the zwitterionic lipid POPC, a major component of cell membranes, can be incorporated into MO cubic phases to ~ 25 mol % without perturbing the phase behaviour of the system (Lindblom and Rilfors, 1989; Cullis *et al.*, 1996; Cherezov *et al.*, 2002). However, since POPC and MetEnk would have similar effects on the properties of the cubic phase (as previously discussed), it would be advisable to incorporate less than 25 mol % of POPC into MetEnk-containing cubic phases.

MetEnk is known to interact with nerve cell membranes which contain the lipid PE, so it would also be interesting to study the structure of this peptide in PE-containing cubic phases (Cullis *et al.*, 1996). These studies could be conducted using cubic phases prepared from mixtures of MO, PE and water, or mixtures of PE, alamethicin and water which are also known to form bicontinuous cubic phases (Keller *et al.*, 1996; Giorgione *et al.*, 1998; Cherezov *et al.*, 2002). Since the dynamically averaged shape of PE is a cone, similar to that of MO, and significantly different from the cylindrical shape of PC, it is conceivable that a greater amount of MetEnk could be incorporated into PE-containing cubic phases than PC-containing cubic phases due to the complementary effects of these two molecules on the pressure profiles of the cubic phase 'bilayer'. Whereas MetEnk will increase the packing pressure in the headgroup regions, PE will increase the packing pressure *within* the 'bilayer', as opposed to PC which increases the packing pressure in *both* the headgroup and acyl chain regions of the 'bilayer'. The electrostatic contributions to MetEnk lipid-binding and structure could be investigated using cubic phases containing the anionic lipid PS, a natural constituent of eukaryotic cell membranes (Marcotte *et al.*, 2003).

6.3.2 Solid State NMR

It would also be interesting to collect solid state NMR data on transmembrane peptides in the MO cubic phase for comparison with the data collected on this system using solution NMR techniques. It is unclear whether or not the transmembrane residues of cubic phase incorporated peptides would be observable in solid-state NMR spectra. These types of studies could be conducted on alamethicin in the cubic phase since this transmembrane peptide can be incorporated to relatively high concentrations into the MO cubic phase without causing perturbation of the cubic phase. It is however unlikely that these types of studies could be conducted on other transmembrane peptides in the cubic phase because, at the high peptide concentrations that are required for solid-state NMR studies, transmembrane peptides perturb the phase behaviour of cubic phase forming systems (Farés *et al.*, 2002; Chupin *et al.*, 2003).

6.3.3 Cubic Phase Tailoring

A considerable amount of work has been conducted with respect to the incorporation of different lipids into MO cubic phases to potentially expand their applications for protein structure determination, and as drug delivery systems. These cubic phases have been characterized using X-ray diffraction, to determine the lattice dimensions of the cubic phases, and by solid state NMR, to study the lipid packing properties in these phases. For the most part, these cubic phases have not been characterized using ^{1}H NMR techniques. It would be interesting to determine

whether or not the incorporation of different types of lipids would have an effect on the ^{1}H NMR spectra of lipids or incorporated peptides, or on the observed lipid, peptide and water diffusion rates characteristic of 'pure' MO cubic phases since the lattice-dimensions and lateral packing pressure profiles of cubic phases change upon incorporation of charged and uncharged lipid and protein molecules (Giorgione *et al.*, 1998; Aota-Nakano *et al.*, 1999; Cherezov *et al.*, 2002; Chupin *et al.*, 2003; Misquitta *et al.*, 2004). Cubic phases with different physical characteristics could be used in studies of lipid-protein interactions to investigate the effects of different lipids or membrane properties on a protein's structure or its interactions with lipid membranes.

References

Aggeli, A., Bannister, M. L., Bell, M., Boden, N., Findlay, J. B. C., Hunter, M., Knowles, P. F. and Yang, J.-C. (1998). "Conformation and Ion-Channeling Activity of a 27-Residue Peptide Modeled on the Single-Transmembrane Segment of the IsK (minK) Protein." *Biochemistry* **37**:8121-8131.

Ai, X. and Caffrey, M. (2000). "Membrane Protein Crystallization in Lipidic Mesophases: Detergent Effects." *Biophysical Journal* **79**:394-405.

Altieri, A. S., Hinton, D. P. and Byrd, R. A. (1995). "Association of Biomolecular Systems Via Pulsed-Field Gradient NMR Self-Diffusion Measurements." *Journal of the American Chemical Society* **117**:7566-7567.

Amodeo, P., Naider, F., Picone, D., Tancredi, T. and Temussi, P. A. (1998). "Conformational sampling of bioactive conformers: a low-temperature NMR study of 15N-Leu-enkephalin." *Journal of Peptide Science* **4**:253-265.

Andersson, A. and Mäler, L. (2002). "NMR solution structure and dynamics of motilin in isotropic phospholipid bicellar solution." *Journal of Biomolecular NMR* **24**:103-12.

Aota-Nakano, Y., Li, S. J. and Yamazaki, M. (1999). "Effects of electrostatic interaction on the phase stability and structures of cubic phases of monoolein/oleic acid mixture membranes." *Biochimica et Biophysica Acta* **1461**:96-102.

Aranda, F. J., Killian, J. A. and de Kruijff, B. (1987). "Importance of the tryptophans of gramicidin for its lipid structure modulating activity in lysophosphatidylcholine and phosphatidylethanolamine model membranes. A comparative study employing gramicidin analogs and a synthetic alpha-helical hydrophobic polypeptide." *Biochimica et Biophysica Acta* **901**:217-228.

Aurora, R. and Rose, G. D. (1998). "Helix capping." *Protein Science* **7**:21-38.

Barauskas, J., Razumas, V., Talaikyte, Z., Bulovas, A., Nylander, T., Tauraite, D., Butkus, E. (2003). "Towards redox active liquid crystalline phases of lipids: a monoolein/water system with entrapped derivatives of ferrocene." *Chemistry and Physics of Lipids* **123**:87-97.

Basáñez, G., Ruiz-Argüello, M. B., Alonso, A., Goñi, F. M., Karlsson, G. and Edwards, K. (1997). "Morphological Changes Induced by Phospholipase C and by Sphingomyelinase on Large Unilamellar Vesicles: A Cryo-Transmission Electron Microscopy Study of Liposome Fusion." *Biophysical Journal* **72**:2630-2637.

Bazzo, R., Tappin, M. J., Pastore, A., Harvey, T. S., Carver, J. A. and Campbell, I. D. (1988). "The structure of melittin: A ^{1}H-NMR study in methanol." *European Journal of Biochemistry* **173**:139-146.

Bechinger, B. (1997). "Structure and Functions of Channel-Forming Peptides: Magainins, Cecropins, Melittin and Alamethicin." *The Journal of Membrane Biology* **156:**197-211.

Bechinger, B., Aisenbrey, C. and Bertani, P. (2004). "The alignment, structure and dynamics of membrane-associated polypeptides by solid-state NMR spectroscopy." *Biochimica et Biophysica Acta* **1666:**190-204.

Bechinger, B., Skladnev, D. A., Ogrel, A., Li, X., Rogozhkina, E. V., Ovchinnikova, T. V., O'Neil, J. D. J. and Raap, J. (2001). "^{15}N and ^{31}P Solid-State NMR Investigations on the Orientation of Zervamicin II and Alamethicin in Phosphatidylcholine Membranes." *Biochemistry* **40:**9428-9437.

Behnam, B. A. and Deber, C. M. (1984). "Evidence for a folded conformation of methionine- and leucine-enkephalin in a membrane environment." *The Journal of Biological Chemistry* **259:**14935-40.

Belrhali, H., Nollert, P., Royant, A., Menzel, C., Rosenbusch, J. P., Landau, E. M. and Pebay-Peyroula, E. (1999). "Protein, lipid and water organization in bacteriorhodopsin crystals: a molecular view of the purple membrane at 1.9 A resolution." *Structure* **7:**909-917.

BioRad (2000). Sub-Cell GT Agarose Gel Electrophoresis Systems Instruction Manual: p. 9.

Biverståhl, H., Andersson, A., Gräslund, A. and Mäler, L. (2004). "NMR solution structure and membrane interaction of the N-terminal sequence (1-30) of the bovine prion protein." *Biochemistry* **43:**14940-7.

Bloom, M. (1995). "Trans-Membrane Peptide and Protein Structure in Fluid Membranes via NMR." *Biophysical Journal* **69:**1631-1632.

Booth, V., Waring, A. J., Walther, F. J. and Keough, K. M. (2004). "NMR structures of the C-terminal segment of surfactant protein B in detergent micelles and hexafluoro-2-propanol." *Biochemistry* **43:**15187-94.

Bouchard, M., Davis, J. H. and Auger, M. (1995). "High-Speed Magic Angle Spinning Solid-State 1H Nuclear Magnetic Resonance Study of the Conformation of Gramicidin A in Lipid Bilayers." *Biophysical Journal* **69:**1933-1938.

Boyd, B. J. (2003). "Characterisation of drug release from cubosomes using the pressure ultrafiltration method." *International Journal of Pharmaceutics* **260:**239-247.

Briggs, J., Chung, H. and Caffrey, M. (1996). "The Temperature-Composition Phase Diagram and Mesophase Structure Characterization of the Monoolein/Water System." *Journal de Physique II France* **6:**723-751.

Brown, L. R., Braun, W., Kumar, A. and Wüthrich, K. (1982). "High resolution nuclear magnetic resonance studies of the conformation and orientation of melittin bound to a lipid-water interface." *Biophysical Journal* **37:**319-328.

Brown, L. R., Lauterwein, J. and Wüthrich, K. (1980). "High-resolution ^{1}H-NMR studies of self-aggregation of melittin in aqueous solution." *Biochimica et Biophysica Acta* **622:**231-244.

Brown, L. R. and Wüthrich, K. (1981). "Melittin bound to dodecylphosphocholine micelles: ^{1}H-NMR assignments and global conformational features." *Biochimica et Biophysica Acta* **647:**95-111.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. and Warren, G. L. (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." *Acta Crystallographica Section D - Biological Crystallography* **54:**905-921.

Cai, M., Huang, Y., Sakaguchi, K., Clore, G. M., Gronenborn, A. M. and Craigie, R. (1998). "An efficient and cost-effective isotope labelling protocol for proteins expressed in *Escherichia coli*." *Journal of Biomolecular NMR* **11:**97-102.

Casselyn, M., Perez, J., Tardieu, A., Vachette, P., Witz, J. and Delacroix, H. (2001). "Spherical plant viruses: interactions in solution, phase diagrams and crystallization of brome mosaic virus." *Acta Crystallographica Section D - Biological Crystallography* **57:**1799-812.

Cavanagh, J. and Rance, M. (1992). "Suppression of Cross-Relaxation Effects in Tocsy Spectra Via a Modified Dipsi-2 Mixing Sequence." *Journal of Magnetic Resonance* **96:**670-678.

Chekmenev, E. Y., Hu, J., Gor'kov, P. L., Brey, W. W., Cross, T. A., Ruuge, A. and Smirnov, A. I. (2005). "(15)N and (31)P solid-state NMR study of transmembrane domain alignment of M2 protein of influenza A virus in hydrated cylindrical lipid bilayers confined to anodic aluminum oxide nanopores." *Journal of Magnetic Resonance* **173:**322-7.

Chen, A. and Shapiro, M. (1999). "Nuclear overhauser effect on diffusion measurements." *Journal of the American Chemical Society* **121:**5338-5339.

Cherezov, V., Clogston, J., Misquitta, Y., Abdel-Gawad, W. and Caffrey, M. (2002). "Membrane protein crystallization in meso: lipid type-tailoring of the cubic phase." *Biophysical Journal* **83:**3393-407.

Cherry, R. J. (1979). "Rotational and lateral diffusion of membrane proteins." *Biochimica et Biophysica Acta* **559:**289-327.

Chou, J. J., Kaufman, J.D., Stahl, S.J., Wingfield, P.T., Bax, A. (2002). "Micelle-Induced Curvature in a Water-Insoluble HIV-1 Env Peptide Revealed by NMR Dipolar Coupling Measurement in Stretched Polyacrylamide Gel." *Journal of the American Chemical Society* **124:**2450-2451.

Choy, W. Y. and Forman-Kay, J. D. (2001). "Calculation of ensembles of structures representing the unfolded state of an SH3 domain." *Journal of Molecular Biology* **308:**1011-32.

Chung, H. and Caffrey, M. (1994). "The Neutral Area Surface of the Cubic Mesophase: Location and Properties." *Biophysical Journal* **66:**377-381.

Chupin, V., Killian, J. A., Breg, J., de Jongh, H. H. J., Boelens, R., Kaptein, R. and de Kruijff, B. (1995). "PhoE Signal Peptide Inserts into Micelles as a Dynamic Helix-Break-Helix Structure, Which Is Modulated by the Environment. A Two-Dimensional ^{1}H NMR Study." *Biochemistry* **34:**11617-11624.

Chupin, V., Killian, J. A. and de Kruijff, B. (2003). "Effect of phospholipids and a transmembrane Peptide on the stability of the cubic phase of monoolein: implication for protein crystalization from a cubic phase." *Biophysical Journal* **84:**2373-81.

Chupin, V., Leenhouts, J. M., de Kroon, A. I. P. M. and de Kruijff, B. (1996). "Secondary Structure and Topology of a Mitochondrial Presequence Peptide Associated with Negatively Charged Micelles. A 2D ^{1}H-NMR Study." *Biochemistry* **35:**3140-3146.

Cobas, J. C. and Sardina, F. J. (2003). "Nuclear magnetic resonance data processing. MestRe-C: A software package for desktop computers." *Concepts in Magnetic Resonance Part A* **19A:**80-96.

Cullis, P. R., Fenske, D. B. and Hope, M. J. (1996). Physical properties and functional roles of lipids in membranes. Biochemistry of Lipids, Lipoproteins and Membranes. Vance, D. E. and Vance, J. E. Amsterdam, Elsevier.

D'Alagni, M., Delfini, M., Di Nola, A., Eisenberg, M., Paci, M., Roda, L. G. and Veglia, G. (1996). "Conformational study of [Met5]enkephalin-Arg-Phe in the presence of phosphatidylserine vesicles." *European Journal of Biochemistry* **240:**540-549.

Dan, N. and Safran, S. A. (1998). "Effect of Lipid Characteristics on the Structure of Transmembrane Proteins." *Biophysical Journal* **75:**1410-1414.

Davis, J. H., Auger, M. and Hodges, R. S. (1995). "High Resolution ^{1}H Nuclear Magnetic Resonance of a Transmembrane Peptide." *Biophysical Journal* **69:**1917-1932.

Davis, J. H., Clare, D. M., Hodges, R. S. and Bloom, M. (1983). "Interaction of a Synthetic Amphiphilic Polypeptide and Lipids in a Bilayer Structure." *Biochemistry* **22:**5298-5305.

De Angelis, A. A., Nevzorov, A. A., Park, S. H., Howell, S. C., Mrse, A. A. and Opella, S. J. (2004). "High-resolution NMR spectroscopy of membrane proteins in aligned bicelles." *Journal of the American Chemical Society* **126:**15340-1.

de Boer, H. A. and Kastelein, R. A. (1986). Biased Codon Usage: An Exploration of Its Role in Optimization of Translation. Maximizing Gene Expression. Reznikoff, W., Gold, L. Boston: 225-285.

de Groot, B. L. and Grubmüller, H. (2001). "Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF." *Science* **294:**2353-7.

de Planque, M. R., Bonev, B. B., Demmers, J. A., Greathouse, D. V., Koeppe, R. E., 2nd, Separovic, F., Watts, A. and Killian, J. A. (2003). "Interfacial anchor properties of tryptophan residues in transmembrane peptides can dominate over hydrophobic matching effects in peptide-lipid interactions." *Biochemistry* **42:**5341-8.

de Planque, M. R., Boots, J. W., Rijkers, D. T., Liskamp, R. M., Greathouse, D. V. and Killian, J. A. (2002). "The effects of hydrophobic mismatch between phosphatidylcholine bilayers and transmembrane alpha-helical peptides depend on the nature of interfacially exposed aromatic and charged residues." *Biochemistry* **41:**8396-8404.

de Planque, M. R. and Killian, J. A. (2003). "Protein-lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring." *Molecular Membrane Biology* **20:**271-84.

de Planque, M. R. R., Greathouse, D. V., Koeppe II, R. E., Schafer, H., Marsh, D. and Killian, J. A. (1998). "Influence of Lipid/Peptide Hydrophobic Mismatch on the Thickness of Diacylphosphatidylcholine Bilayers. A ^{2}H NMR and ESR Study Using Designed Transmembrane a-Helical Peptides and Gramicidin A." *Biochemistry* **37:**9333-9345.

Deber, C. M. and Behnam, B. A. (1984). "Role of membrane lipids in peptide hormone function: Binding of enkephalins to micelles." *Proceedings of the National Academy of Sciences of the United States of America* **81:**61-65.

Deber, C. M. and Behnam, B. A. (1985). "Transfer of Peptide-Hormones from Aqueous to Membrane Phases." *Biopolymers* **24:**105-116.

Deber, C. M., Khan, A. R., Li, Z., Joensson, C., Glibowicka, M. and Wang, J. (1993). "Val to Ala mutations selectively alter helix-helix packing in the transmembrane segment of phage M13 coat protein." *Proceedings of the National Academy of Sciences of the United States of America* **90:**11648-11652.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995). "NMRPipe: A multidimensional spectral processing system based on UNIX pipes." *Journal of Biomolecular NMR* **6:**277-293.

Dempsey, C. E. (1990). "The actions of melittin on membranes." *Biochimica et Biophysica Acta* **1031:**143-161.

Dempsey, C. E. and Handcock, L. J. (1996). "Hydrogen Bond Stabilities in Membrane-Reconstituted Alamethicin from Amide-Resolved Hydrogen-Exchange Measurements." *Biophysical Journal* **70:**1777-1788.

Deng, Y., Marko, M., Buttle, K. F., Leith, A., Mieczkowski, M. and Mannella, C. A. (1999). "Cubic Membrane Structure in Ameoba (*Chaos carolinensis*) Mitochondria Determined by Electron Microscopic Tomography." *Journal of Structural Biology* **127:**231-239.

Dores, R. M., Lecaude, S., Bauer, D. and Danielson, P. B. (2002). "Analyzing the evolution of the opioid/orphanin gene family." *Mass Spectrometry Reviews* **21:**220-43.

Drevet, P., Lemaire, C., Gasparini, S., Zinn-Justin, S., Lajeunesse, E., Ducancel, F., Pinkasfeld, S., Courcon, M., Tremeau, O., Boulain, J. C. and Menez, A. (1997). "High-Level Production and Isotope Labeling of Snake Neurotoxins, Disulfide-rich Proteins." *Protein Expression and Purification* **10:**293-300.

Dumas, F., Lebrun, M. C. and Tocanne, J. F. (1999). "Is the protein/lipid hydrophobic matching principle relevant to membrane organization and functions?" *FEBS Letters* **458:**271-7.

Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T. and MacKinnon, R. (2002). "X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity." *Nature* **415:**287-94.

Ellena, J. F., Moulthrop, J., Wu, J., Rauch, M., Jaysinghne, S., Castle, J. D. and Cafiso, D. S. (2004). "Membrane position of a basic aromatic peptide that sequesters phosphatidylinositol 4,5 bisphosphate determined by site-directed spin labeling and high-resolution NMR." *Biophysical Journal* **87:**3221-33.

Emerson, S. D., Waugh, D. S., Scheffler, J. E., Tsao, K.-L., Prinzo, K. M. and Fry, D. C. (1994). "Chemical Shift Assignments and Folding Topology of the RAS-Binding Domain of Human RAF-1 As Determined by Heteronuclear Three-Dimensional NMR Spectroscopy." *Biochemistry* **33:**7745-7752.

Engström, S., Norden, T. P. and Hyquist, H. (1999). "Cubic phases for studies of drug partition into lipid bilayers." *European Journal of Pharmaceutical Sciences* **8:**243-254.

Epand, R. M. (1998). "Lipid polymorphism and protein-lipid interactions." *Biochimica et Biophysica Acta* **1376:**353-368.

Ericsson, B., Larsson, K. and Fontell, K. (1983). "A Cubic Protein-Monoolein-Water Phase." *Biochimica et Biophysica Acta* **729:**23-27.

Eriksson, P. O. and Lindblom, G. (1993). "Lipid and water diffusion in bicontinuous cubic phases measured by NMR." *Biophysical Journal* **64:**129-36.

Esposito, G., Carver, J. A., Boyd, J. and Campbell, I. D. (1987). "High-Resolution ^{1}H NMR Study of the Solution Structure of Alamethicin." *Biochemistry* **26:**1043-1050.

Farés, C., Sharom, F. J. and Davis, J. H. (2002). "$^{15}N,^{1}H$ Heteronuclear correlation NMR of gramicidin A in DMPC-d_{67}." *Journal of the American Chemical Society* **124:**11232-3.

Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D. and Kay, L. E. (1994). "Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation." *Biochemistry* **33:**5984-6003.

Fernández, C. and Wüthrich, K. (2003). "NMR solution structure determination of membrane proteins reconstituted in detergent micelles." *FEBS Letters* **555:**144-50.

Franklin, J. C., Ellena, J. F., Jayasinghe, S., Kelsh, L. P. and Cafiso, D. S. (1994). "Structure of Micelle-Associated Alamethicin from ^{1}H NMR. Evidence for Conformational Heterogeneity in a Voltage-Gated Peptide." *Biochemistry* **33:**4036-4045.

Fregeau Gallagher, N. L., Sailer, M., Niemczura, W. P., Nakashima, T. T., Stiles, M. E. and Vederas, J. C. (1997). "Three-Dimensional Structure of Leucocin A in Trifluoroethanol and dodecylphosphocholine Micelles: Spatial Location of Residues Critical for Biological Activity in Type IIa Bacteriocins from Lactic Acid Bacteria." *Biochemistry* **36:**15062-15072.

Gaemers, S. and Bax, A. (2001). "Morphology of Three Lyotropic Liquid Crystalline Biological NMR Media Studied by Translational Diffusion Anisotropy." *Journal of the American Chemical Society* **123:**12343-12352.

Gao, X. and Wong, T. C. (1998). "Studies of the binding and structure of adrenocorticotropin peptides in membrane mimics by NMR spectroscopy and pulsed-field gradient diffusion." *Biophysical Journal* **74:**1871-88.

Gavit, P. and Better, M. (2000). "Production of antifungal recombinant peptides in *Escherichia coli*." *Journal of Biotechnology* **79:**127-136.

Giorgione, J. R., Huang, Z. and Epand, R. M. (1998). "Increased Activation of Protein Kinase C with Cubic Phase Lipid Compared with Liposomes." *Biochemistry* **37:**2384-2392.

Glover, K. J., Whiles, J.A., Vold, R.R., Melacini, G. (2002). "Position of residues in transmembrane peptides with respect to the lipid bilayer: A combined lipid NOEs and water chemical exchange approach in phospholipid bicelles." *Journal of Biomolecular NMR* **22:**57-64.

Gohon, Y. and Popot, J. L. (2003). "Membrane protein-surfactant complexes." *Current Opinion in Colloid and Interface Science* **8:**15-22.

Grabe, M., Neu, J., Oster, G. and Nollert, P. (2003). "Protein interactions and membrane geometry." *Biophysical Journal* **84:**854-68.

Graham, W. H., Carter, E. S., 2nd and Hicks, R. P. (1992). "Conformational analysis of Met-enkephalin in both aqueous solution and in the presence of sodium dodecyl sulfate micelles using multidimensional NMR and molecular modeling." *Biopolymers* **32:**1755-64.

Griffin, R. G. (1998). "Dipolar recoupling in MAS spectra of biological solids." *Nature Structural Biology*:508-512.

Guex, N. and Peitsch, M. C. (1997). "SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling." *Electrophoresis* **18:**2714-2723.

Gulik, A., Luzzati, V., De Rosa, M. and Gambacorta, A. (1985). "Structure and polymorphism of bipolar isopranyl ether lipids from archaebacteria." *Journal of Molecular Biology* **182:**131-49.

Gysin, B. and Schwyzer, R. (1983). "Head group and structure specific interactions of enkephalins and dynorphin with liposomes: investigation by hydrophobic photolabeling." *Archives of Biochemistry and Biophysics* **225:**467-74.

Harris, R. K., Becker, E. D., De Menezes, S. M. C., Goodfellow, R. and Granger, P. (2001). "NMR nomenclature. Nuclear spin properties and conventions for chemical shifts - (IUPAC recommendations 2001)." *Pure and Applied Chemistry* **73:**1795-1818.

Harzer, U. and Bechinger, B. (2000). "Alignment of Lysine-Anchored Membrane Peptides under Conditions of Hydrophobic Mismatch: A CD, ^{15}N and ^{31}P Solid-State NMR Spectroscopy Investigation." *Biochemistry* **39:**13106-13114.

Hasler, L., Heymann, J. B., Engel, A., Kistler, J. and Walz, T. (1998). "2D Crystallization of Membrane Proteins: Rationales and Examples." *Journal of Structural Biology* **121:**162-171.

Haught, C., Davis, G. D., Subramanian, R., Jackson, K. W. and Harrison, R. G. (1998). "Recombinant Production and Purification of Novel Antisense Antimicrobial Peptide in *Escherichia coli*." *Biotechnology and Bioengineering* **57:**55-61.

He, K., Ludtke, J., Heller, W. T. and Huang, H. W. (1996). "Mechanism of Alamethicin Insertion into Lipid Bilayers." *Biophysical Journal* **71:**2669-2679.

Henry, G. D. and Sykes, B. D. (1990). "Hydrogen Exchange Kinetics in a Membrane Protein Determined by ^{15}N NMR Spectroscopy: Use of the INEPT Experiment To Follow Individual Amides in Detergent-Solubilized M13 Coat Protein." *Biochemistry* **29:**6303-6313.

Henry, G. D., Weiner, J. H. and Sykes, B. D. (1986). "Backbone Dynamics of a Model Membrane Protein: ^{13}C NMR Spectroscopy of Alanine Methyl Groups in Detergent-Solubilized M13 Coat Protein." *Biochemistry* **25:**590-598.

Hicks, R. P., Beard, D. J. and Young, J. K. (1992). "The Interactions of Neuropeptides with Membrane Model Systems: A Case-Study." *Biopolymers* **32:**85-96.

Hirsh, D. J., Hammer, J., Maloy, W. L., Blazyk, J. and Schaefer, J. (1996). "Secondary Structure and Location of a Magainin Analogue in Synthetic Phospholipid Bilayers." *Biochemistry* **35:**12733-12741.

Hochkoeppler, A., Landau, E. M., Venturoli, G., Zannoni, D., Feick, R. and Luigi Luisi, P. (1995). "Photochemistry of a Photosynthetic Reaction Center Immobilized in Lipidic Cubic Phases." *Biotechnology and Bioengineering* **46:**93-98.

Hope, M. J. and Cullis, P. R. (1981). "The role of nonbilayer lipid structures in the fusion of human erythrocytes induced by lipid fusogens." *Biochimica et Biophysica Acta* **640:**82-90.

Howell, S. C., Mesleh, M. F. and Opella, S. J. (2005). "NMR Structure Determination of a Membrane Protein with Two Transmembrane Helices in Micelles: MerF of the Bacterial Mercury Detoxification System." *Biochemistry* **44:**5196-206.

Hughes, J., Smith, T. W., Kosterlitz, H. W., Fothergill, L. A., Morgan, B. A. and Morris, H. R. (1975). "Identification of two related pentapeptides from the brain with potent opiate agonist activity." *Nature* **258:**577-80.

Huschilt, J. C., Millman, B. M. and Davis, J. H. (1989). "Orientation of α-helical peptides in a lipid bilayer." *Biochimica et Biophysica Acta* **979:**139-141.

Hyde, S. T., Andersson, S., Ericsson, B. and Larsson, K. (1984). "A cubic structure consisting of a lipid bilayer forming an infinite periodic minimum surface of the gyroid type in the glycerolmonoleat-water system." *Zeitschrift für Kristallographie* **168:**213-219.

Jansson, M., Li, Y.-C., Jendeberg, L., Anderson, S., Montelione, G. T. and Nilsson, B. (1996). "High-level production of uniformly ^{15}N- and ^{13}C-enriched fusion proteins in *Escherichia coli*." *Journal of Biomolecular NMR* **7:**131-141.

Jarrell, H. C., Deslauriers, R., McGregor, W. H. and Smith, I. C. (1980). "Interaction of opioid peptides with model membranes. A carbon-13 nuclear magnetic study of enkephalin binding to phosphatidylserine." *Biochemistry* **19:**385-390.

Jarvis, J. A., Ryan, M. T., Hoogenraad, N. J., Craik, D. J. and Høj, P. B. (1995). "Solution Structure of the Acetylated and Noncleavable Mitochondrial Targeting Signal of Rat Chaperonin 10." *The Journal of Biological Chemistry* **270:**1323-1331.

Jastimi, R. E. and Lafleur, M. (1999). "Nisin promotes the formation of non-lamellar inverted phases in unsaturated phophatidylethanolamines." *Biochimica et Biophysica Acta* **1418:**97-105.

Jensen, M. O. and Mouritsen, O. G. (2004). "Lipids do influence protein function-the hydrophobic matching hypothesis revisited." *Biochimica et Biophysica Acta* **1666:**205-26.

Jeong, S. W., O'Brien, D. F., Oradd, G. and Lindblom, G. (2002). "Encapsulation and diffusion of water-soluble dendrimers in a bicontinuous cubic phase." *Langmuir* **18:**1073-1076.

Jiang, Y., Lee, A., Chen, J., Ruta, V., Cadene, M., Chait, B. T. and MacKinnon, R. (2003). "X-ray structure of a voltage-dependent K+ channel." *Nature* **423:**33-41.

Johnson, B. A. and Blevins, R. A. (1994). "NMR View - a Computer-Program for the Visualization and Analysis of NMR Data." *Journal of Biomolecular NMR* **4:**603-614.

Jones, D. H., Ball, E. H., Sharpe, S., Barber, K. R. and Grant, C. W. M. (2000). "Expression and Membrane Assembly of a Transmembrane Region from Neu." *Biochemistry* **39:**1870-1878.

Kainosho, M. (1997). "Isotope labelling of macromolecules for structural determinations." *Nature Structural Biology*:858-861.

Katona, G., Andreasson, U., Landau, E. M., Andreasson, L. E. and Neutze, R. (2003). "Lipidic cubic phase crystal structure of the photosynthetic reaction centre from Rhodobacter sphaeroides at 2.35A resolution." *Journal of Molecular Biology* **331:**681-92.

Kay, L. E., Keifer, P. and Saarinen, T. (1992). *Journal of the American Chemical Society* **114:**10663-10665.

Keller, S. L., Gruner, S. M. and Gawrisch, K. (1996). "Small concentrations of alamethicin induce a cubic phase in bulk phosphatidylethanolamine mixtures." *Biochimica et Biophysica Acta* **1278:**241-246.

Killian, J. A. (2003). "Synthetic peptides as models for intrinsic membrane proteins." *FEBS Letters* **555:**134-8.

Killian, J. A., de Jong, A. M. P., Bijvelt, J., Verkleij, A. J. and de Kruijff, B. (1990). "Induction of non-bilayer lipid structures by functional signal peptides." *The EMBO Journal* **9:**815-819.

Killian, J. A., Salemink, I., de Planque, M.R.R., Lindblom, G., Koeppe II, R.E., Greathouse, D.V. (1996). "Induction of Nonbilayer Structures in Diacylphosphatidylcholine Model Membranes by Transmembrane a-Helical Peptides: Important of Hydrophobic Mismatch and Proposed Role of Tryptophans." *Biochemistry* **35:**1037-1045.

Kim, S., Baum, J. and Anderson, S. (1997). "Production of correctly folded recombinant [^{13}C, ^{15}N]-enriched guinea pig [Val90]-alpha-lactalbumin." *Protein Engineering* **10:**455-462.

Kohno, T., Kusunoki, H., Sato, K. and Wakamatsu, K. (1998). "A new general method for the biosynthesis of stable isotope-enriched peptides using a decahistidine-tagged ubiquitin fusion system: an application to the production of mastoparan-X uniformly enriched with 15N and 15N/13C." *Journal of Biomolecular NMR* **12:**109-21.

Koradi, R., Billeter, M. and Wüthrich, K. (1996). "MOLMOL: a program for display and analysis of macromolecular structures." *Journal of Molecular Graphics* **14:**51-5, 29-32.

Kovacs, F. A., Denny, J. K., Song, Z., Quine, J. R. and Cross, T. A. (2000). "Helix Tilt of the M2 Transmembrane Peptide from Influenza A Virus: An Intrinsic Property." *Journal of Molecular Biology* **295:**117-125.

Kuliopulos, A. and Walsh, C. T. (1994). "Production, Purification, and Cleavage of Tandem Repeats of Recombinant Peptides." *Journal of the American Chemical Society* **116:**4599-4607.

LaBean, T. H., Kauffman, S. A. and Butt, T. R. (1995). "Libraries of random-sequence polypeptides produced with high yield as carboxy-terminal fusions with ubiquitin." *Molecular Diversity* **1:**29-38.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." *Nature* **227:**680-5.

Lajmi, A. R., Wallace, T. R. and Shin, J. A. (2000). "Short, Hydrophobic, Alanine-Based Proteins Based on the Basic Region/Leucine Zipper Protein Motif: Overcoming Inclusion Body Formation and Protein Aggregation during Overexpression, Purification, and Renaturation." *Protein Expression and Purification* **18:**394-403.

Landau, E. M. and Luigi Luisi, P. (1993). "Lipidic Cubic Phases as Transparent, Rigid Matrices for the Direct Spectroscopic Study of Immobilized Membrane Proteins." *Journal of the American Chemical Society* **115:**2102-2106.

Landau, E. M. and Rosenbusch, J. P. (1996). "Lipidic cubic phases: A novel concept for the crystallization of membrane proteins." *Proceedings of the National Academy of Sciences of the United States of America* **93:**14532-14535.

Landh, T. (1995). "From entangled membranes to eclectic morphologies: cubic membranes as subcellular space organizers." *FEBS Letters* **369:**13-17.

Larsson, K., Fontell, K. and Krog, N. (1980). "Structural relationships between lamellar, cubic and hexagonal phases in monoglyceride-water systems. Possibility of cubic structures in biological systems." *Chemistry and Physics of Lipids* **27:**321-328.

Larsson, K. and Lindblom, G. (1982). "Molecular amphiphile bilayers forming a cubic phase in amphiphile-water systems." *Journal of Dispersion Science and Technology* **3:**61-66.

Lauterwein, J., Brown, L. R. and Wüthrich, K. (1980). "High-Resolution ^{1}H-NMR Studies of Monomeric Melittin in Aqueous Solution." *Biochimica et Biophysica Acta* **622:**219-230.

Lee, A. G. (2004). "How lipids affect the activities of integral membrane proteins." *Biochimica et Biophysica Acta* **1666:**62-87.

Lee, J. H., Kim, J. H., Hwang, S. W., Lee, W. J., Yoon, H. K., Lee, H. S. and Hong, S. S. (2000). "High-Level Expression of Antimicrobial Peptide Mediated by a Fusion Partner Reinforcing Formation of Inclusion Bodies." *Biochemical and Biophysical Research Communications* **277:**575-580.

Lee, J. H., Minn, I., Park, C. B. and Kim, S. C. (1998). "Acidic Peptide-Mediated Expression of the Antimicrobial Peptide Buforin II as Tandem Repeats in *Escherichia coli*." *Protein Expression and Purification* **12:**53-60.

Lemaitre, V., de Planque, M. R., Howes, A. P., Smith, M. E., Dupree, R. and Watts, A. (2004). "Solid-state 17O NMR as a probe for structural studies of proteins in biomembranes." *Journal of the American Chemical Society* **126:**15320-1.

Lendermann, J. and Winter, R. (2003). "Interaction of cytochrome c with cubic monoolein mesophases at limited hydration conditions: The effects of concentration, temperature and pressure." *Physical Chemistry Chemical Physics* **5:**1440-1450.

Lew, S., Ren, J. and London, E. (2000). "The Effects of Polar and/or Ionizable Residues in the Core and Flanking Regions of Hydrophobic Helices on Transmembrane Conformation and Oligomerization." *Biochemistry* **39:**9632-9640.

Lindblom, G., Larsson, K., Johansson, L., Fontell, K. and Forsen, S. (1979). "The Cubic Phase of Monoglyceride-Water Systems. Arguments for a Structure Based upon Lamellar Bilayer Units." *Journal of the American Chemical Society* **101:**5465-5470.

Lindblom, G. and Rilfors, L. (1989). "Cubic phases and isotropic structures formed by membrane lipids-possible biological relevance." *Biochimica et Biophysica Acta* **988:**221-256.

Liu, F., Lewis, R. N. A. H., Hodges, R. S. and McElhaney, R. N. (2001). "A Differential Scanning Calorimetric and ^{31}P NMR Spectroscopic Study of the Transmembrane α-Helical Peptides on the Lamellar-Reversed Hexagonal Phase Transition of Phosphatidylethanolamine Model Membranes." *Biochemistry* **40:**760-768.

Liu, F., Lewis, R. N. A. H., Hodges, R. S. and McElhaney, R. N. (2002). "Effect of Variations in the Structure of a Polyleucine-Based a-Helical Transmembrane Peptide on Its Interaction with Phosphatidylcholine Bilayers." *Biochemistry* **41:**9197-9207.

Liu, W. and Caffrey, M. (2005). "Gramicidin structure and disposition in highly curved membranes." *Journal of Structural Biology* **150:**23-40.

Losonczi, J. A., Tian, F. and Prestegard, J. H. (2000). "Nuclear Magnetic Resonance Studies of the N-Terminal Fragment of Adenosine Diphosphate Ribosylation Factor 1 in Micelles and Bicelles: Influence of N-Myristoylation." *Biochemistry* **39:**3804-3816.

Lucas, L. H., Yan, J., Larive, C. K., Zartler, E. R. and Shapiro, M. J. (2003). "Transferred nuclear overhauser effect in nuclear magnetic resonance diffusion measurements of ligand-protein binding." *Analytical Chemistry* **75:**627-34.

Ludtke, S., He, K. and Huang, H. (1995). "Membrane thinning caused by magainin 2." *Biochemistry* **34:**16764-9.

Lutton, E. S. (1965). "Phase Behaviour of Aqueous Systems of Monoglycerides." *Journal of the American Oil Chemists' Society* **42:**1068-1070.

Luzzati, V. (1997). "Biological significance of lipid polymorphism: the cubic phases." *Current Opinion in Structural Biology* **7:**661-668.

Luzzati, V. and Husson, F. (1962). "The structure of the liquid-crystalline phases of lipid-water systems." *The Journal of Cell Biology* **12:**207-219.

Majerle, A., Kidric, J. and Jerala, R. (2000). "Production of stable isotope enriched antimicrobial peptides in *Escherichia coli*: An application to the production of a ^{15}N-enriched fragment of lactoferrin." *Journal of Biomolecular NMR* **18:**145-151.

Marcotte, I., Dufourc, E. J., Ouellet, M. and Auger, M. (2003). "Interaction of the neuropeptide met-enkephalin with zwitterionic and negatively charged bicelles as viewed by 31P and 2H solid-state NMR." *Biophysical Journal* **85:**328-39.

Marcotte, I., Separovic, F., Auger, M. and Gagne, S. M. (2004). "A multidimensional 1H NMR investigation of the conformation of methionine-enkephalin in fast-tumbling bicelles." *Biophysical Journal* **86:**1587-600.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M. and Clore, G. M. (1989a). "Overcoming the overlap problem in the assignment of 1H NMR spectra of larger proteins by use of three-dimensional heteronuclear 1H-15N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta." *Biochemistry* **28:**6150-6.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989b). "Rapid Recording of 2d NMR-Spectra without Phase Cycling - Application to the Study of Hydrogen-Exchange in Proteins." *Journal of Magnetic Resonance* **85:**393-399.

Marrink, S.-J. and Tieleman, D. P. (2001). "Molecular Dynamics Simulation of a Lipid Diamond Cubic Phase." *Journal of the American Chemical Society* **123:**12383-12391.

Martinez, E., Jimenez, M. A., Seguí-Real, B., Vandekerckhove, J. and Sandoval, I. V. (1997). "Folding of the Presequence of Yeast pAPI into an Amphipathic Helix Determines Transport of the Protein from the Cytosol to the Vacuole." *Journal of Molecular Biology* **267:**1124-1138.

Masum, S. M., Li, S. J., Tamba, Y., Yamashita, Y., Tanaka, T. and Yamazaki, M. (2003). "Effect of de novo designed peptides interacting with the lipid-membrane interface on the stability of the cubic phases of the monoolein membrane." *Langmuir* **19:**4745-4753.

McIntosh, L. P. and Dahlquist, F. W. (1990). "Biosynthetic Incorporation of ^{15}N and ^{13}C for Assignment and Interpretation of Nuclear Magnetic Resonance Spectra of Proteins." *Quarterly Reviews of Biophysics* **23:**1-38.

McIntosh, T. J., Ting-Beall, H. P. and Zampighi, G. (1982). "Alamethicin-Induced Changes in Lipid Bilayer Morphology." *Biochimica et Biophysica Acta* **685:**51-60.

Milon, A., Miyazawa, T. and Higashijima, T. (1990). "Transferred Nuclear Overhauser Effect Analyses of Membrane-Bound Enkephalin Analogues by ^{1}H Nuclear Magnetic Resonance: Correlation between Activities and Membrane-Bound Conformations." *Biochemistry* **29:**65-75.

Miskolzie, M. and Kotovych, G. (2003). "The NMR-derived conformation of neuropeptide AF, an orphan G-protein coupled receptor peptide." *Biopolymers* **69:**201-215.

Misquitta, Y. and Caffrey, M. (2003). "Detergents Destabilize the Cubic Phase of Monoolein: Implications for Membrane Protein Crystallization." *Biophysical Journal* **85:**3084-96.

Misquitta, Y., Cherezov, V., Havas, F., Patterson, S., Mohan, J. M., Wells, A. J., Hart, D. J. and Caffrey, M. (2004). "Rational design of lipid for membrane protein crystallization." *Journal of Structural Biology* **148:**169-75.

Monette, M. and Lafleur, M. (1996). "Influence of Lipid Chain Unsaturation on Melittin-Induced Micellization." *Biophysical Journal* **70:**2195-2202.

Morein, S., Koeppe II, R. E., Lindblom, G., de Kruijff, B. and Killian, J. A. (2000). "The Effect of Peptide/Lipid Hydrophobic Mismatch on the Phase Behavior of Model Membranes Mimicking the Lipid Composition in *Escherichia coli* Membranes." *Biophysical Journal* **78:**2475-2485.

Morein, S., Standberg, E., Killian, J. A., Persson, S., Arvidson, G., Koeppe II, R. E. and Lindblom, G. (1997). "Influence of Membrane-Spanning a-Helical Peptides on the Phase Behavior of the Dioleoylphosphatidylcholine/Water System." *Biophysical Journal* **73:**3078-3088.

Morein, S., Trouard, T. P., Hauksson, J. B., Rilfors, L., Arvidson, G. and Lindblom, G. (1996). "Two-dimensional ^{1}H-NMR of transmembrane peptides from *Escherichia coli* phosphatidylglycerophosphate synthase in micelles." *European Journal of Biochemistry* **241:**489-497.

Morrow, M. R., Huschilt, J. C. and Davis, J. H. (1985). "Simultaneous Modeling of Phase and Calorimetric Behavior in an Amphiphilic Peptide/Phospholipid Model Membrane." *Biochemistry* **24:**5396-5406.

Motta, A., Tancredi, T. and Temussi, P. A. (1987). "Nuclear Overhauser effects in linear peptides. A low-temperature 500 MHz study of Met-enkephalin." *FEBS Letters* **215:**215-218.

Mouritsen, O. G. and Bloom, M. (1984). "Mattress model of lipid-protein interactions in membranes." *Biophysical Journal* **46:**141-53.

Navarro, J., Landau, E. M. and Fahmy, K. (2002). "Receptor-dependent G-protein activation in lipidic cubic phase." *Biopolymers* **67:**167-77.

Nieva, J. L., Alonso, A., Basáñez, G., Goñi, F. M., Gulik, A., Vargas, R. and Luzzati, V. (1995). "Topological properties of two cubic phases of a phospholipid:cholesterol:diacylglycerol aqueous system and their possible implications in the phospholipase C-induced liposome fusion." *FEBS Letters* **368:**143-7.

Nollert, P., Qiu, H., Caffrey, M., Rosenbusch, J. P. and Landau, E. M. (2001). "Molecular mechanism for the crystallization of bacteriorhodopsin in lipidic cubic phases." *FEBS Letters* **504:**179-86.

Nollert, P., Royant, A., Pebay-Peyroula, E. and Landau, E. M. (1999). "Detergent-free membrane protein crystallization." *FEBS Letters* **457:**205-208.

Okon, M., Frank, P. G., Marcel, Y. L. and Cushley, R. J. (2001). "Secondary structure of human apolipoprotein A-I(1-186) in lipid-mimetic solution." *FEBS Letters* **487:**390-6.

Okon, M., Frank, P. G., Marcel, Y. L. and Cushley, R. J. (2002). "Heteronuclear NMR studies of human serum apolipoprotein A-I. Part I. Secondary structure in lipid-mimetic solution." *FEBS Letters* **517:**139-43.

O'Neil, J. D. J. and Sykes, B. D. (1988). "Structure and Dynamics of a Detergent-Solublized Membrane Protein: Measurement of Amide Hydrogen Exchange Rates in M13 Coat Protein by ^{1}H NMR Spectroscopy." *Biochemistry* **27:**2753-2762.

Orzaez, M., Lukovic, D., Abad, C., Perez-Paya, E. and Mingarro, I. (2005). "Influence of hydrophobic matching on association of model transmembrane fragments containing a minimised glycophorin A dimerisation motif." *FEBS Letters* **579:**1633-8.

Ottavi, A., Tiennault, E., Maftah, A., Berjeaud, J. M., Cenatiempo, Y. and Julien, R. (1998). "An improved method to obtain a single recombinant vasoactive intestinal peptide (VIP) analog." *Biochimie* **80:**289-293.

Ottleben, H., Haasemann, M., Ramachandran, R., Gorlach, M., Muller-Esterl, W. and Brown, L. R. (1997). "An NMR study of the interaction of ^{15}N-labelled bradykinin with an antibody mimic of the bradykinin B2 receptor." *European Journal of Biochemistry* **244:**471-478.

Paas, Y., Cartaud, J., Recouvreur, M., Grailhe, R., Dufresne, V., Pebay-Peyroula, E., Landau, E. M. and Changeux, J. P. (2003). "Electron microscopic evidence for nucleation and growth of 3D acetylcholine receptor microcrystals in structured lipid-detergent matrices." *Proceedings of the National Academy of Sciences of the United States of America* **100:**11309-11314.

Papavoine, C. H. M., Christiaans, B. E. C., Folmer, R. H. A., Konings, R. N. H. and Hilbers, C. W. (1998). "Solution Structure of the M13 Major Coat Protein in Detergent Micelles: A Basis for a Model of Phage Assembly Involving Specific Residues." *Journal of Molecular Biology* **282:**401-419.

Papavoine, C. H. M., Konings, R. N. H., Hilbers, C. W. and van de Ven, F. J. M. (1994). "Location of M13 Coat Protein in Sodium Dodecyl Sulfate Micelles As Determined by NMR." *Biochemistry* **33:**12990-12997.

Park, S.-H., Shalongo, W. and Stellwagen, E. (1998). "Analysis of N-Terminal Capping Using Carbonyl-Carbon Chemical Shift Measurements." *Proteins: Structure, Function and Genetics* **33:**167-176.

Patton, J. S. and Carey, M. C. (1979). "Watching fat digestion." *Science* **204:**145-8.

Pebay-Peyroula, E., Neutze, R. and Landau, E. M. (2000). "Lipidic cubic phase crystallization of bacteriorhodopsin and cryotrapping of intermediates: towards resolving a revolving photocycle." *Biochimica et Biophysica Acta* **1460:**119-132.

Pebay-Peyroula, E., Rummel, G., Rosenbusch, J. P. and Landau, E. M. (1997). "X-ray Structure of Bacteriorhodopsin at 2.5 Angstroms from Microcrystals Grown in Lipidic Cubic Phases." *Science* **277:**1676-1681.

Penel, S., Hughes, E. and Doig, A. J. (1999a). "Side-chain Structures in the First Turn of the a-Helix." *Journal of Molecular Biology* **287:**127-143.

Penel, S., Morrison, R. G., Mortishire-Smith, R. J. and Doig, A. J. (1999b). "Periodicity in a-Helix Lengths and C-Capping Preferences." *Journal of Molecular Biology* **293:**1211-1219.

Picone, D., D'Ursi, A., Motta, A., Tancredi, T. and Temussi, P. A. (1990). "Conformational preferences of [Leu5]enkephalin in biomimetic media. Investigation by 1H NMR." *European Journal of Biochemistry* **192:**433-9.

Pilon, A. L., Yost, P., Chase, T. E., Lohnas, G. L. and Bentley, W. E. (1996). "High-Level Expression and Efficient Recovery of Ubiquitin Fusion proteins from *Escherichia coli*." *Biotechnology Progress* **12:**331-337.

Piotto, S. (2004). "Lipid aggregates inducing symmetry breaking in prebiotic polymerisations." *Origins of Life and Evolution of the Biosphere* **34:**123-32.

Portmann, M., Landau, E. M. and Luigi Luisi, P. (1991). "Spectroscopic and Rheological Studies of Enzymes in Rigid Lipidic Matrices: The Case of a-Chymotrypsin in a Lysolecithin/Water Cubic Phase." *The Journal of Physical Chemistry* **95:**8437-8440.

Presta, L. G. and Rose, G. D. (1988). "Helix Signals in Proteins." *Science* **240:**1632-1641.

Prosser, F. S., Daleman, S. I. and Davis, J. H. (1994). "The Structure of an Integral Membrane Peptide: A Deuterium NMR Study of Gramicidin." *Biophysical Journal* **66:**1415-1428.

Qutub, Y., Reviakine, I., Maxwell, C., Navarro, J., Landau, E. M. and Vekilov, P. G. (2004). "Crystallization of transmembrane proteins in cubo: mechanisms of crystal growth and defect formation." *Journal of Molecular Biology* **343:**1243-54.

Raussens, V., Slupsky, C. M., Sykes, B. D. and Ryan, R. O. (2003). "Lipid-bound structure of an apolipoprotein E-derived peptide." *The Journal of Biological Chemistry* **278:**25998-6006.

Richardson, J. S. and Richardson, D. C. (1988). "Amino Acid Preferences for Specific Locations at the Ends of α Helices." *Science* **240:**1648-1652.

Rigby, A. C., Grant, C. W. M. and Shaw, G. S. (1998). "Solution and solid state conformation of the human EGF receptor transmembrane region." *Biochimica et Biophysica Acta* **1371:**241-253.

Riley, L. G., Junius, F. K., Swanton, M. K., Vesper, N. A., Williams, N. K., King, G. F. and Weiss, A. S. (1994). "Cloning, expression, and spectroscopic studies of the Jun leucine zipper domain." *European Journal of Biochemistry* **219:**877-886.

Rilfors, L., Eriksson, P.-O., Arvidson, G. and Lindblom, G. (1986). "Relationship between Three-Dimensional Arrays of "Lipidic Particles" and Bicontinuous Cubic Lipid Phases." *Biochemistry* **25:**7702-7711.

Rinia, H. A., Kik, R. A., Demel, R. A., Snel, M. M. E., Killian, J. A., van der Eerden, J. P. J. M. and de Kruijff, B. (2000). "Visualization of Highly Ordered Striated Domains Induced by Transmembrane Peptides in Supported Phosphatidylcholine Bilayers." *Biochemistry* **39:**5852-5858.

Roosild, T. P., Greenwald, J., Vega, M., Castronovo, S., Riek, R. and Choe, S. (2005). "NMR structure of Mistic, a membrane-integrating protein for membrane protein expression." *Science* **307:**1317-21.

Rosen, H. and Liao, J. (2003). "Sphingosine 1-phosphate pathway therapeutics: a lipid ligand-receptor paradigm." *Current Opinion in Chemical Biology* **7:**461-8.

Rouhani, S., Facciotti, M. T., Woodcock, G., Cheung, V., Cunningham, C., Nguyen, D., Rad, B., Lunde, C. S. and Glaeser, R. M. (2002). "Crystallization of membrane proteins from media composed of connected-bilayer gels." *Biopolymers* **66:**300-316.

Roux, M., Neumann, J.-M., Hodges, R. S., Devaux, P. F. and Bloom, M. (1989). "Conformational Changes of Phospholipid Headgroups Induced by a Cationic Integral Membrane Peptide As Seen by Deuterium Magnetic Resonance." *Biochemistry* **28:**2313-2321.

Rovnyak, D., Frueh, D. P., Sastry, M., Sun, Z. Y., Stern, A. S., Hoch, J. C. and Wagner, G. (2004). "Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction." *Journal of Magnetic Resonance* **170:**15-21.

Rummel, G., Hardmeyer, A., Widmer, C., Chiu, M. L., Nollert, P., Locher, K. P., Pedruzzi, I., Landau, E. M. and Rosenbusch, J. P. (1998). "Lipidic Cubic Phases: New Matrices for the Three-Dimensional Crystallization of Membrane Proteins." *Journal of Structural Biology* **121:**82-91.

Sanders II, C. R. and Landis, G. C. (1994). "Facile Acquisition and Assignment of Oriented Sample NMR-Spectra for Bilayer Surface-Associated Proteins." *Journal of the American Chemical Society* **116:**6470-6471.

Sanders II, C. R. and Landis, G. C. (1995). "Reconstitution of Membrane Proteins into Lipid-Rich Bilayered Mixed Micelles for NMR Studies." *Biochemistry* **34:**4030-4040.

Schenkl, S., Portuondo, E., Zgrablic, G., Chergui, M., Suske, W., Dolder, M., Landau, E. M. and Haacke, S. (2003). "Compositional heterogeneity reflects partial dehydration in three-dimensional crystals of bacteriorhodopsin." *Journal of Molecular Biology* **329:**711-9.

Schibli, D. J., Montelaro, R. C. and Vogel, H. J. (2001). "The Membrane-Proximal Tryptophan-Rich Region of the HIV Flycoprotein, gp41, Forms a Well-Defined Helix in Dodecylphosphocholine Micelles." *Biochemistry* **40:**9570-9578.

Schwyzer, R. (1986). "Molecular mechanism of opioid receptor selection." *Biochemistry* **25:**6335-42.

Sennoga, C., Heron, A., Seddon, J. M., Templer, R. H. and Hankamer, B. (2003). "Membrane-protein crystallization in cubo: temperature-dependent phase behaviour of monoolein-detergent mixtures." *Acta Crystallographica Section D - Biological Crystallography* **59:**239-46.

Siegel, D. P. (1999). "The Modified Stalk Mechanism of Lamellar/Inverted Phase Transitions and Its Implications for Membrane Fusion." *Biophysical Journal* **76:**291-313.

Siegel, D. P. and Epand, R. M. (2000). "Effect of influenza hemagglutinin fusion peptide on lamellar/inverted phase transitions in dipalmitoleoylphosphatidylethanolamine: implications for membrane fusion mechanisms." *Biochimica et Biophysica Acta* **1468:**87-98.

Siekmann, B., Bunjes, H., Koch, M. H. and Westesen, K. (2002). "Preparation and structural investigations of colloidal dispersions prepared from cubic monoglyceride-water phases." *International Journal of Pharmaceutics* **244:**33-43.

Singer, S. J. and Nicolson, G. R. (1972). "The fluid mosaic model of the structure of cell membranes." *Science* **175:**720-731.

Sklenár, V., Piotto, M., Leppik, R. and Saudek, V. (1993). "Gradient-Tailored Water Suppression for H-1-N-15 Hsqc Experiments Optimized to Retain Full Sensitivity." *Journal of Magnetic Resonance Series A* **102:**241-245.

Smith, S. O., Jonas, R., Braiman, M. and Bormann, B. J. (1994). "Structure and Orientation of the Transmembrane Domain of Glycophorin-a in Lipid Bilayers." *Biochemistry* **33:**6334-6341.

Soulié, S., Neumann, J.-M., Berthomeiu, C., Moller, J. V., le Maire, M. and Forge, V. (1999). "NMR Conformational Study of the Sixth Transmembrane Segment of Sarcoplasmic Reticulum Ca^{2+}-ATPase." *Biochemistry* **38:**5813-5821.

Sparr, E., Wadsten, P., Kocherbitov, V. and Engstrom, S. (2004). "The effect of bacteriorhodopsin, detergent and hydration on the cubic-to-lamellar phase transition in the monoolein-distearoyl phosphatidyl glycerol-water system." *Biochimica et Biophysica Acta* **1665:**156-66.

Stankowski, S. and Schwarz, G. (1990). "Electrostatics of a peptide at a membrane/water interface. The pH dependence of melittin association with lipid vesicles." *Biochimica et Biophysica Acta* **1025:**164-172.

Stellwagen, E. and Shalongo, W. (1997). "Evidence for Glutamate Self-Capping Within a Peptide Helix." *Biopolymers* **43:**413-418.

Stilbs, P. (1982). "Fourier-Transform NMR Pulsed-Gradient Spin-Echo (Ft-Pgse) Self-Diffusion Measurements of Solubilization Equilibria in Sds Solutions." *Journal of Colloid and Interface Science* **87:**385-394.

Sui, H., Han, B. G., Lee, J. K., Walian, P. and Jap, B. K. (2001). "Structural basis of water-specific transport through the AQP1 water channel." *Nature* **414:**872-8.

Surewicz, W. K. and Mantsch, H. H. (1988). "Solution and membrane structure of enkephalins as studied by infrared spectroscopy." *Biochemical and Biophysical Research Communications* **150:**245-251.

Tanaka, S., Egelhaaf, S. U. and Poon, W. C. (2004). "Crystallization of a globular protein in lipid cubic phase." *Physical Review Letters* **92:**128102.

Tang, Y. C. and Deber, C. M. (2004). "Aqueous solubility and membrane interactions of hydrophobic peptides with peptoid tags." *Biopolymers* **76:**110-8.

Tessmer, M. R. and Kallick, D. A. (1997). "NMR and Structural Model of Dynorphin A (1-17) Bound to Dodecylphosphocholine Micelles." *Biochemistry* **36:**1971-1981.

Thai, K., Choi, J., Franzin, C. M. and Marassi, F. M. (2005). "Bcl-XL as a fusion protein for the high-level expression of membrane-associated proteins." *Protein Science* **14:**948-55.

Therien, A. G., Glibowicka, M. and Deber, C. M. (2002). "Expression and purification of two hydrophobic double-spanning membrane proteins derived from the cystic fibrosis transmembrane conductance regulator." *Protein Expression and Purification* **25:**81-6.

Tomich, J. M., Wallace, D., Henderson, K., Mitchell, K. E., Radke, G., Brandt, R., Ambler, C. A., Scott, A. J., Grantham, J., Sullivan, L. and Iwamoto, T. (1998). "Aqueous Solubilization of Transmembrane Peptide Sequences with Retention of Membrane Insertion and Function." *Biophysical Journal* **74:**256-267.

Tsapis, N., Reiss-Husson, F., Ober, R., Genest, M., Hodges, R. S. and Urbach, W. (2001). "Self Diffusion and Spectral Modifications of Membrane Protein, the *Rubrivivax gelatinosus* LH2 Complex, Incorporated into a Monoolein Cubic Phase." *Biophysical Journal* **81:**1613-1623.

Um, J. Y., Chung, H., Kim, K. S., Kwon, I. C. and Jeong, S. Y. (2003). "In vitro cellular interaction and absorption of dispersed cubic particles." *International Journal of Pharmaceutics* **253:**71-80.

van den Brink-van der Laan, E., Killian, J. A. and de Kruijff, B. (2004). "Nonbilayer lipids affect peripheral and integral membrane proteins via changes in the lateral pressure profile." *Biochimica et Biophysica Acta* **1666:**275-88.

Venkatraman, J., Nagana Gowda, G. A. and Balaram, P. (2002). "Structural analysis of synthetic peptide fragments from EmrE, a multidrug resistance protein, in a membrane-mimetic environment." *Biochemistry* **41:**6631-9.

Verkleij, A. J. (1984). "Lipidic Intramembranous Particles." *Biochimica et Biophysica Acta* **779:**43-63.

Vogt, T. C. B., Killian, J. A. and De Kruijff, B. (1994). "The influence of acylation on the lipid structure modulating properties of the transmembrane polypeptide gramicidin." *Biochimica et Biophysica Acta* **1193:**55-61.

Vold, R. R., Prosser, R.S., Deese, A.J. (1997). "Isotropic solutions of phospholipid bicelles: A new membrane mimetic for high-resolution NMR studies of polypeptides." *Journal of Biomolecular NMR* **9:**329-335.

Wallin, E. and von Heijne, G. (1998). "Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms." *Protein Science* **7:**1029-38.

Wang, C. and Deber, C. M. (2000). "Peptide Mimics of the M13 Coat Protein Transmembrane Segment." *The Journal of Biological Chemistry* **275:**16155-16159.

Weaver, A. J., Kemple, M. D., Brauner, J. W., Mendelsohn, R. and Prendergast, F. G. (1992). "Fluorescence, CD, Attenuated Total Reflectance (ATR) FTIR, and ^{13}C NMR Characterization of the Structure and Dynamics of Synthetic Melittin and Melittin Analogues in Lipid Environments." *Biochemistry* **31:**1301-1313.

Werten, P. J., Rémigy, H. W., de Groot, B. L., Fotiadis, D., Philippsen, A., Stahlberg, H., Grubmüller, H. and Engel, A. (2002). "Progress in the analysis of membrane protein structure and function." *FEBS Letters* **529:**65-72.

Whiles, J. A., Brasseur, R., Glover, K. J., Melacini, G., Komives, E. A. and Vold, R. R. (2001). "Orientation and effects of mastoparan X on phospholipid bicelles." *Biophysical Journal* **80:**280-93.

White, S. H. (2003). "Translocons, thermodynamics, and the folding of membrane proteins." *FEBS Letters* **555:**116-21.

White, S. H. (2004). "The progress of membrane protein structure determination." *Protein Science* **13:**1948-9.

Whitehead, T. L., Jones, L. M. and Hicks, R. P. (2001). "Effects of the incorporation of CHAPS into SDS micelles on neuropeptide-micelle binding: separation of the role of electrostatic interactions from hydrophobic interactions." *Biopolymers* **58:**593-605.

Wickner, W. (1988). "Mechanisms of Membrane Assembly: General Lessons from the Study of M13 Coat Protein and *Escherichia coli* Leader Peptidase." *Biochemistry* **27:**1086-1094.

Wienk, H. L. J., Wechselberger, R.W., Czisch, M., de Kruijff, B. (2000). "Structure, Dynamics, and Insertion of a Chloroplast Targeting Peptide in Mixed Micelles." *Biochemistry* **39:**8219-8227.

Wimley, W. C. and White, S. H. (2000). "Designing Transmembrane α-Helices That Insert Spontaneously." *Biochemistry* **39:**4432-4442.

Winter, R. and Köhling, R. (2004). "Static and time-resolved synchrotron small-angle x-ray scattering studies of lyotropic lipid mesophases, model biomembranes and proteins in solution." *Journal of Physics - Condensed Matter* **16:**S327-S352.

Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. and Sykes, B. D. (1995). "1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects." *Journal of Biomolecular NMR* **5:**67-81.

Wishart, D. S. and Sykes, B. D. (1994). "The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data." *Journal of Biomolecular NMR* **4:**171-80.

Wishart, D. S., Sykes, B. D. and Richards, F. M. (1992). "The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy." *Biochemistry* **31:**1647-51.

Wishart, D. S., Sykes, B.D., Richards, F.M. (1991). "Relationship between Nuclear Magnetic Resonance Chemical Shift and Protein Secondary Structure." *Journal of Molecular Biology* **222:**311-333.

Wood, M. J. and Komives, E. A. (1999). " Production of large quantities of isotopically labeled protein in *Pichia pastoris* by fermentation." *Journal of Biomolecular NMR* **13:**149-159.

Woolley, G. A. and Wallace, B. A. (1992). "Model Ion Channels: Gramicidin and Alamethicin." *The Journal of Membrane Biology* **129:**109-136.

Woolley, G. A. and Wallace, B. A. (1993). "Temperature Dependence of the Interaction of Alamethicin Helices in Membranes." *Biochemistry* **32:**9819-9825.

Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids. New York, John Wiley & Sons.

Yan, C., Digate, R. J. and Guiles, R. D. (1999). "NMR studies of the structure and dynamics of peptide E, an endogenous opioid peptide that binds with high affinity to multiple opioid receptor subtypes." *Biopolymers* **49:**55-70.

Yang, D., O'Brien, D. F. and Marder, S. R. (2002). "Polymerized bicontinuous cubic nanoparticles (cubosomes) from a reactive monoacylglycerol." *Journal of the American Chemical Society* **124:**13388-9.

Yau, W. M., Wimley, W. C., Gawrisch, K. and White, S. H. (1998). "The preference of tryptophan for membrane interfaces." *Biochemistry* **37:**14713-8.

Yee, A. A., Marat, K. and O'Neil, J. D. J. (1997). "The interactions with solvent, heat stability, and ^{13}C-labelling of alamethicin, an ion-channel-forming peptide." *European Journal of Biochemistry* **243:**283-291.

Yee, A. A. and O'Neil, J. D. J. (1992). "Uniform ^{15}N Labeling of a Fungal Peptide: The Structure and Dynamics of an Alamethicin by ^{15}N and ^{1}H NMR Spectroscopy." *Biochemistry* **31:**3135-3143.

Yin, Y., Zhang, F., Ling, V. and Arrowsmith, C. H. (1995). "Structural analysis and comparison of the C-terminal transport signal domains of hemolysin a and leukotoxin A." *FEBS Letters* **366:**1-5.

Zamoon, J., Nitu, F., Karim, C., Thomas, D. D. and Veglia, G. (2005). "Mapping the interaction surface of a membrane protein: Unveiling the conformational switch of phospholamban in calcium pump regulation." *Proceedings of the National Academy of Sciences of the United States of America* **102:**4747-52.

Zhang, L., Falla, T., Wu, M., Fidai, S., Burian, J., Kay, W. and Hancock, R. E. W. (1998). "Determinants of Recombinant Production of Antimicrobial Cationic Peptides and Creation of Peptide Variants in Bacteria." *Biochemical and Biophysical Research Communications* **247**:674-680.

Zhang, O. W., Kay, L. E., Olivier, J. P. and Formankay, J. D. (1994). "Backbone H-1 and N-15 Resonance Assignments of the N-Terminal Sh3 Domain of Drk in Folded and Unfolded States Using Enhanced-Sensitivity Pulsed-Field Gradient NMR Techniques." *Journal of Biomolecular NMR* **4**:845-858.

Zhang, Y. P., Lewis, R. N., Hodges, R. S. and McElhaney, R. N. (1995a). "Interaction of a peptide model of a hydrophobic transmembrane alpha-helical segment of a membrane protein with phosphatidylethanolamine bilayers: differential scanning calorimetric and Fourier transform infrared spectroscopic studies." *Biophysical Journal* **68**:847-57.

Zhang, Y.-P., Lewis, R. N. A. H., Henry, G. D., Sykes, B. D., Hodges, R. S. and McElhaney, R. N. (1995b). "Peptide Models of Helical Hydrophobic Transmembrane Segments of Membrane Proteins. 1. Studies of the Conformation, Intrabilayer Orientation, and Amide Hydrogen Exchangeability of Ac-K_2-$(LA)_{12}$-K_2-amide." *Biochemistry* **34**:2348-2361.

Zhang, Y.-P., Lewis, R. N. A. H., Hodges, R. S. and McElhaney, R. N. (1992a). "FTIR Spectroscopic Studies of the Conformation and Amide Hydrogen Exchange of a Peptide Model of the Hydrophobic Transmembrane α-Helices of Membrane Proteins." *Biochemistry* **31**:11572-11578.

Zhang, Y.-P., Lewis, R. N. A. H., Hodges, R. S. and McElhaney, R. N. (1992b). "Interaction of a Peptide Model of a Hydrophobic Transmembrane α-Helical Segment of a Membrane Protein with Phosphatidylcholine Bilayers: Differential Scanning Calorimetric and FTIR Spectroscopic Studies." *Biochemistry* **31**:11579-11588.

Zhang, Y.-P., Lewis, R. N. A. H., Hodges, R. S. and McElhaney, R. N. (2001). "Peptide Models of the Helical Hydrophobic Transmembrane Segments of Membrane Proteins: Interactions of Acetyl-K_2-$(LA)_{12}$-K_2-Amide with Phosphatidylethanolamine Bilayer Membranes." *Biochemistry* **40**:474-482.

Zubkov, S., Lennarz, W. J. and Mohanty, S. (2004). "Structural basis for the function of a minimembrane protein subunit of yeast oligosaccharyltransferase." *Proceedings of the National Academy of Sciences of the United States of America* **101**:3821-6.

Appendix A

Supplementary Materials for Chapter 3

Minimal Media Recipes

NOTE: Prepare *all* solutions using double distilled water.

Recipe for 0.5L of Media for the Preparation of M9 Agar Plates

2.5 mL	1 g/5mL	NH_4Cl
1.43 mL	3 M	NaCl
1.00 mL	1 M	$MgSO_4$
0.5 mL	50 mg/mL	Ampicillin
0.5 mL	34 mg/mL	Chloramphenicol
10 mL	20 %	Glucose (final concentration of 0.4 %)
50 µL	1 M	$CaCl_2$
0.5 mL	0.01 M	$FeCl_2$
0.5 mL	1 mg/mL	Thiamine
50 mL	10 × PO_4	Phosphate Solution

Recipe for 1 L of Optimized Minimal Media for ^{15}N-Labelled Peptide Production

1.5 g		$^{15}NH_4Cl$
2.86 mL	3 M	NaCl
2.00 mL	1 M	$MgSO_4$
1.0 mL	50 mg/mL	Ampicillin
1.0 mL	34 mg/mL	Chloramphenicol
50 mL	20 %	Glucose (final concentration of 1.0 %)
100 µL	1 M	$CaCl_2$
1 mL	0.01 M	$FeCl_2$
1 mL	1 mg/mL	Thiamine
100 mL	10 × PO_4	Phosphate Solution

Recipe for 1 L of Optimized Minimal Media for ^{13}C,^{15}N-Labelled Peptide Production

1.5 g		$^{15}NH_4Cl$
2.86 mL	3 M	NaCl
2.00 mL	1 M	$MgSO_4$
1.0 mL	50 mg/mL	Ampicillin
1.0 mL	34 mg/mL	Chloramphenicol
4 g		Glucose (final concentration of 0.4 %)
100 µL	1 M	$CaCl_2$
1 mL	0.01 M	$FeCl_2$
1 mL	1 mg/mL	Thiamine
100 mL	10 × PO_4	Phosphate Solution

Recipe for 2L of 10 × Phosphate Solution

240 g	Na_2PO_4
120 g	KH_2PO_4

* prepare solution using distilled water. Sterilize immediately by autoclaving.

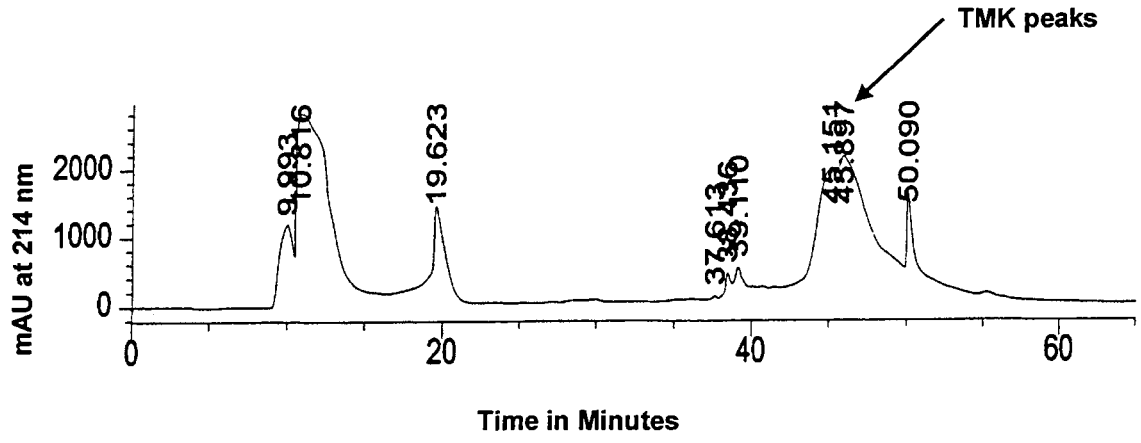


Figure A 1. HPLC Trace of CNBr-Cleaved ^{15}N-labelled KSI-TMK-Histag Fusion Protein

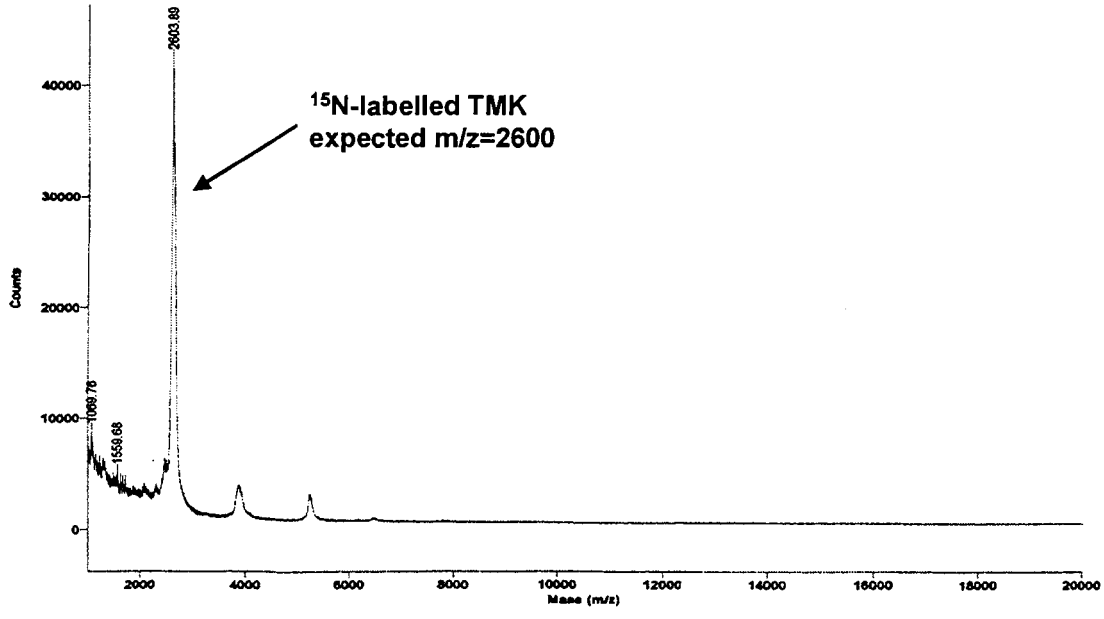


Figure A 2. MALDI-TOF Spectrum of Purified ^{15}N-labelled TMK

Appendix B

Supplementary Materials for Chapter 4

Diffusion Data Analysis Using *MestReC*

Preparation for Data Analysis

The first step in the processing of the NMR diffusion data was to edit the 'procpar' files from each dataset so that the correct values for $(\gamma\delta)^2(\Delta - \delta/3)$ would be calculated by the NMR software program *MestReC* during data analysis (Cobas and Sardina, 2003). A copy of the 'procpar' file called 'original_procpar' was made before the file was modified. The text editor *vi editor* was used to modify the files so they would not become corrupted and unreadable on computers running *UNIX* or *LINUX* operating systems. The modified 'procpar' files were saved and used in subsequent data analyses.

Data Sets Collected to Measure Lipid and Water Diffusion

In the 'procpar' files, the third occurrence of *gzlvl1* in data sets collected to measure lipid diffusion were replaced with *grad_p1*. The location within the 'procpar' files of the *gzlvl1* to be replaced with *grad_p1* is shown below (*gzlvl1* is bolded):

```
gzlvl1 1 1 32768 -32768 1 2 1 0 1 64
20 3000 4420 5840 7260 8680 10100 11520 12940 14360 15780 17200 18620
20040 21460 22880 24300 25720 27140 28560 29980
0
gzlvl0 1 1 32768 -32768 1 2 1 0 1 64
```

Data Sets Collected to Measure Peptide Diffusion

In the 'procpar' files of data sets collected to measure peptide diffusion, the third occurrence of *gzlvl7* were replaced with *grad_p1*. The location within the 'procpar' files of the *gzlvl7* to be replaced with *grad_p1* is shown below (*gzlvl7* is bolded):

```
gzlvl7 1 1 9.99999984307e+17 -9.99999984307e+17 0 2 1 0 1 64
20 3000 4420 5840 7260 8680 10100 11520 12940 14360 15780 17200 18620
20040 21460 22880 24300 25720 27140 28560 29980
0
gzlvl6 1 1 9.99999984307e+17 -9.99999984307e+17 0 2 1 0 1 64
```

Finding the Values for δ and Δ in the 'procpar' Files

The values for δ and Δ were found by searching the 'procpar' files for each experiment using a text editor that would not corrupt the files. In experiments conducted to measure lipid and water diffusion, the values for δ and Δ were the values listed for *gt1* and *BigT*, respectively. The second 'hit' in a search for *gt1* yielded the location of the value for δ which is shown below in bold:

```
gt1 3 1 14 14 14 2 1 8192 1 64
1 0.002
0
gt0 3 1 14 14 14 2 1 8192 1 64
```

The first 'hit' in a search for *BigT* yielded the location of the value for Δ which is underlined below:

```
BigT 3 1 14 14 14 2 1 8192 1 64
1 1
0
aig 2 2 2 0 0 4 1 1 1 64
```

In the experiments that were conducted to measure peptide diffusion, the values for δ and Δ were the values listed for *gt7* and *BigT2*, respectively. The second 'hit' in a search for *gt7* yielded the location of the value for δ which is shown below in bold:

```
gt7 3 1 14 14 14 2 1 8192 1 64
1 0.002
0
gt3 3 1 14 14 14 2 1 8192 1 64
```

The first 'hit' in a search for *BigT2* yielded the location of the value for Δ which is underlined below:

```
BigT2 3 1 14 14 14 2 1 8192 1 64
1 1
0
acqdim 7 1 1000000000 -1000000000 0 2 1 0 1 64
```

Analysis of NMR Spectra

Spectral data were analyzed using the 'Data Analysis' menu. In the window for 'Diffusion Data Analysis' a number of variables had to be entered manually. One of these was the G Scaling Factor (0.002147), a correction factor derived from calibration of the gradients on the spectrometer used for NMR data collection. All of the gradients listed in the 'procpar' file were automatically multiplied by this correction factor during data analysis. The other variables that needed to be verified and/or manually entered for each experiment were δ and Δ, the duration of the pulsed field gradient pulse and the time between pulsed field gradient pulses in seconds. This information was listed in the 'procpar' file for each experiment and had to be found and recorded for each data set since *MestReC* was not able to automatically extract this information from the 'procpar' file during data analysis. The spectral data were analyzed using peak intensity because this gave more reliable and reproducible results than peak integrals. Data analysis was conducted by selecting the peak of interest with the cursor. The values of the peak intensities were calculated by *MestReC* and were listed in a table as a function of $(\gamma G\delta)^2(\Delta - \delta/3)$. A plot of this data could be viewed in *MestReC*, or the data could be imported into an *Excel* spreadsheet and plotted.

A linear plot of the raw data was expected to yield an exponential decay curve. The exponent of the equation describing this decay curve would correspond to the diffusion coefficient (D) for the peak of interest in cm^2/s. If a plot of the normalized natural Log of the peak intensities were plotted as a function of the square of the gradient strength, a linear plot would result. The slope of this line would correspond to the diffusion coefficient for the peak of interest in cm^2/s.

Appendix C

Supplementary Materials for Chapter 5

Additional NMR Spectra Collected on Samples of ^{13}C, ^{15}N-labelled MetEnk in Cubic Phases

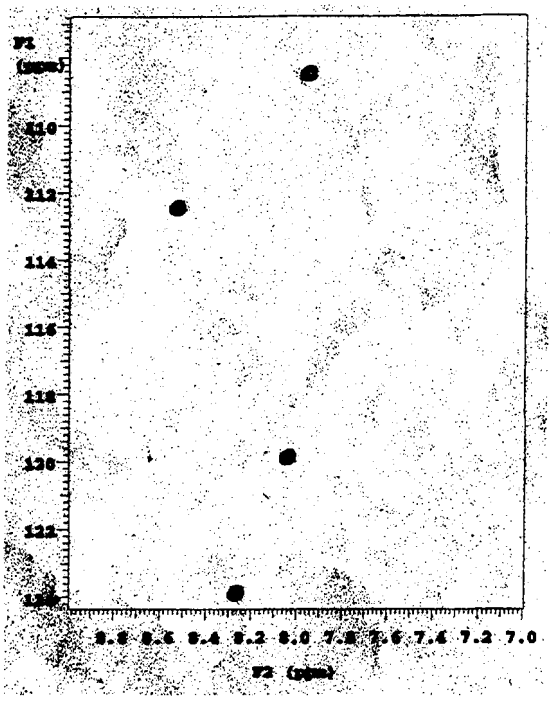


Figure C 1. **2D ^{1}H-^{15}N HSQC Spectra of MetEnk in Water**

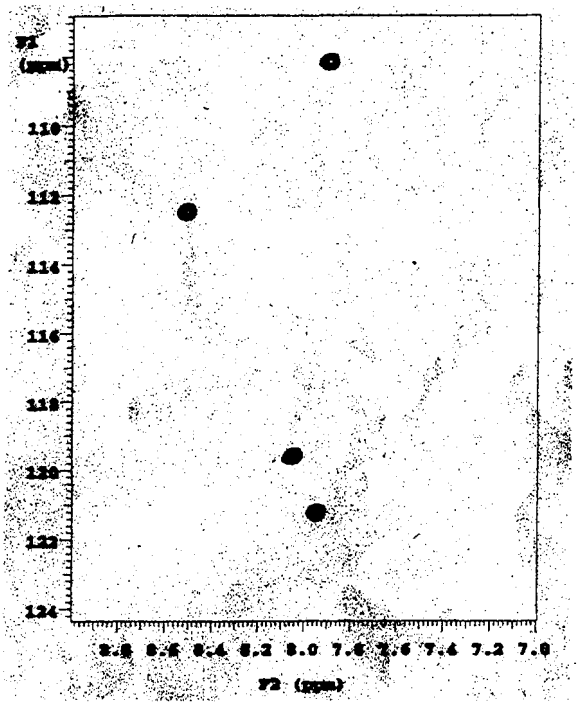


Figure C 2. **2D ^{1}H-^{15}N HSQC Spectra of MetEnk:MO Sample '3'**

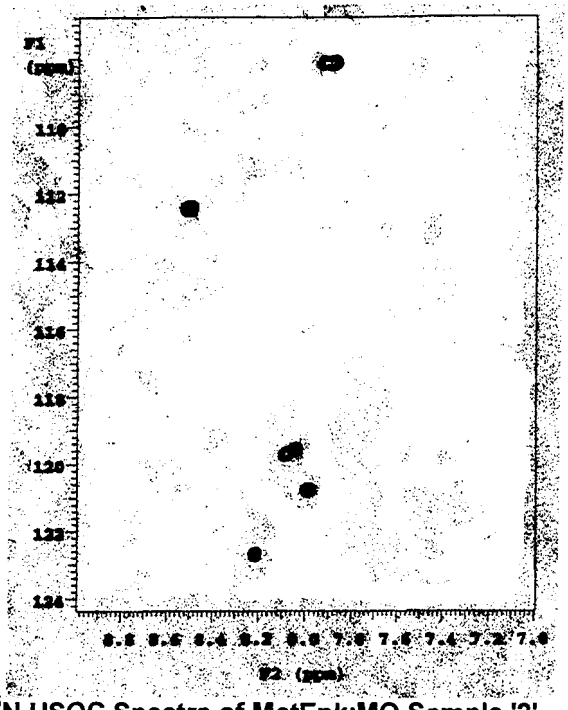


Figure C 3. 2D ^{1}H-^{15}N HSQC Spectra of MetEnk:MO Sample '2'

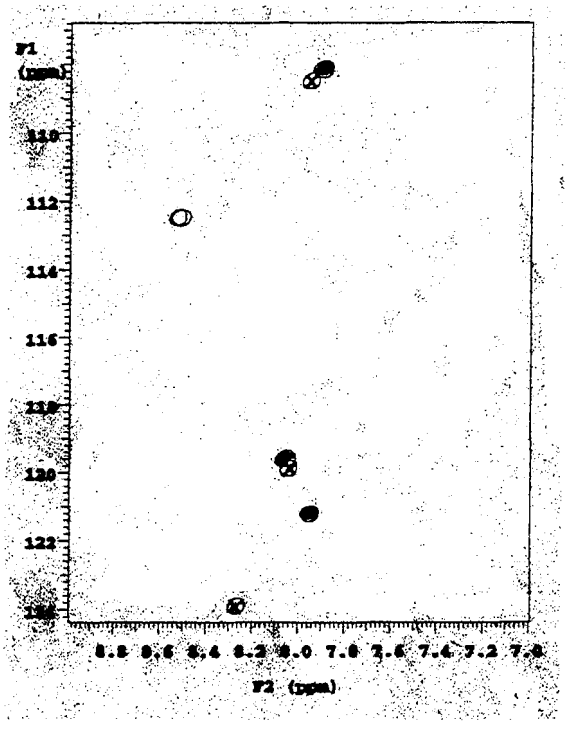


Figure C 4. Superimposition of 2D 1H-15N HSQC Spectra of MetEnk in Water and MetEnk:MO Sample '3'

Peaks denoted with an 'X' are from MetEnk in water, peaks which are filled in are from MetEnk:MO Sample '3', and the peaks which are not filled in are the superimposable peaks.

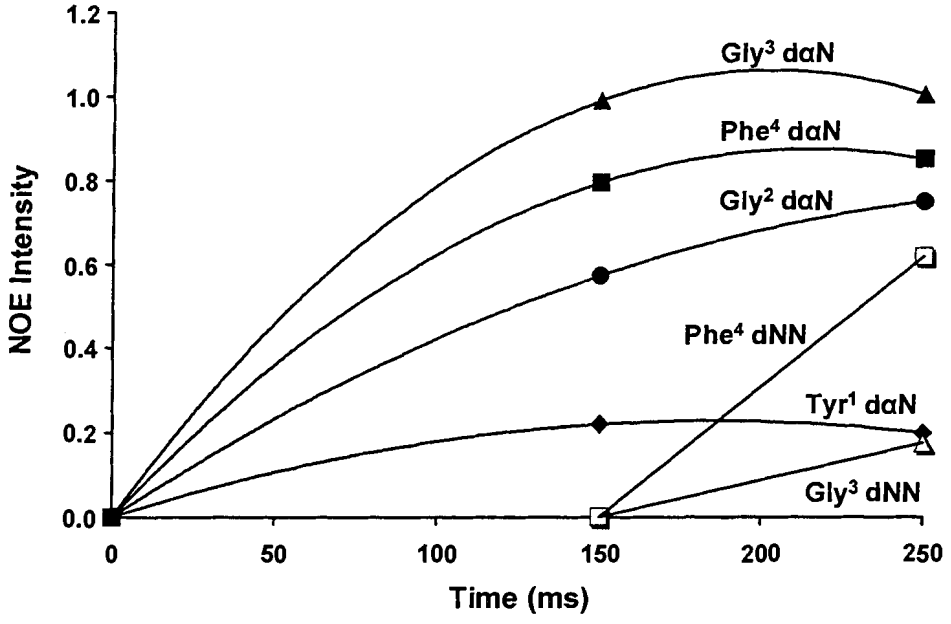


Figure C 5. NOE Build Up Curves for MetEnk:MO Sample '3' Generated Using NOE Data Collected on a 500 MHz Spectrometer at 30°C

Files and Information Related to CNS Structure Calculations

In order to run the program *CNS* for the purposes of structure calculation, it was first necessary to generate the required input files. One of these files was a topology or '.mtf' file. The topology file was generated from the primary sequence of the peptide and specified *all* of the bonds between *all* of the atoms in the molecule. Another file that was required was a '.pdb' file for the extended/unstructured conformation of MetEnk. This file named 'MetEnk_extended.pdb', was generated using the software program *DeepView* (Guex and Peitsch, 1997).

A '.tbl' file listing all of the experimentally determined NOEs constraints was prepared using *NMRView* for use in the structure calculation runs. Two '.tbl' files were prepared: one listing *all* of the aliphatic NOEs called 'original_nnoe.tbl' and one listing *all* of the aromatic NOEs called 'original_caronoe.tbl'. These files were *not* modified to account for peak ambiguities or peak redundancies.

In the file 'original_nnoe.tbl' there were 19 NOE peaks listed, of these 6 had multiple assignments. In the file 'original_caronoe.tbl', 8 NOE peaks were listed and 2 of these peaks were redundant. These '.tbl' were used as the basis for the preparation of '.tbl' files for a series of structure calculation runs that are described below. A summary of the ambiguous NOEs used in each of the structure calculation runs can be found in Table C 1. CNS was run using the 'anneal.inp' script that was supplied with the program after the appropriate input and output filenames had been specified. Upon completion of each structure calculation run an output file called 'anneal.out' was generated, in addition to the structures files in '.pdb' file format. The 'anneal.out' file provided details about each structure calculation run and the structures that were generated.

Editing of '.tbl' Files

The '.tbl' files generated by *NMRView* were edited before use in the majority of structure calculation runs (Johnson and Blevins, 1994). One change that was made to the '.tbl' files was the use of pseudoatoms (denoted with a '#'). Pseudoatoms were used when an NOE peak in an NMR spectrum was assigned to more than one atom because it was not possible to unambiguously assign the peak. An example of this would be the alpha protons of glycine where it was not possible to make stereospecific assignments based on the NMR data that was collected. When pseudoatoms were employed it was necessary to add 1.0 to dplus for methyl and methylene pseudoatoms and 2.5 to dplus for aromatic pseudoatoms.

For NOE peaks with ambiguous assignments, *NMRView* lists all of the possible assignments using 'OR' statements. The program *CNS* does not however do a very good job of interpreting data presented in this format (Booth *et al.*, 2004). It was therefore necessary to modify the '.tbl' files by removing the 'OR' statements prior to their use in structure calculation runs. Structures were then calculated using constraints derived from the unambiguous NOEs and random combinations of the ambiguous NOEs assigned to one, both or neither of the possible assignments. The NOE constraints used for the various structure calculation runs are summarized in Table C 1.

CNS Run 1

In Run 1 the original unmodified NOE list files 'original_nnoe.tbl' and 'original_caronoe.tbl' were used to calculate 50 structures for MetEnk in its cubic phase-bound form. The calculated MetEnk structures were named such that each structure could be uniquely identified and associated with a particular structure calculation run. For example, the seventeenth structure to be calculated was named 'r1_50_17.pdb' where: r1 specifies the run number, 50 the total number of structures calculated and 17 the structure number.

Table C 1. Summary of NOE Constraints Used in CNS Structure Calculations for Cubic Phase Bound MetEnk (MetEnk:MO Sample '3')

Peak ID in file: new_noe _list	Run1_50 [a]	Run2_50 [b]	Run3_50 [c]	Run4_50 [c]	Run5_50 [c]	Run6_50 [c]	Run7_50 [c]	Run8_50 [c]	Run9_50 [c]	Run10_50 [c]	Run11_50 [c]	Run12_50 [c]	Run13_50 [c]	Run14_50 [c]	Run15_50 [c]	Run16_50 [d]	Run17_100 [e]
Pk 7																	
Res4 HN Res2 HA#	OR	OR & dplus	■		■		■			■	■				■		
Res4 HN Res3 HA#			■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Pk 9																	
Res4 HN Res4 HD#	OR	OR & dplus	■	■	■				■	■							
Res4 HN Res4 HE#			■						■	■							
Pk 12																	
Res5 HN Res5 HB2	OR	OR & dplus	■	■		■				■							
Res5 HN Res5 HE1			■														
Pk 13																	
Res5 HN Res5 HB2	OR	OR & dplus	■		■												
Res5 HN Res5 HE1			■	■	■	■				■							
Pk 18																	
Res5 HN Res4 HD#	OR	OR & dplus	■	■	■				■	■	■	■	■		■	■	■
Res5 HN Res4 HE#			■						■	■			■	■	■	■	■
Pk 23																	
Res3 HN Res3 HA#	OR	OR & dplus	■		■		■		■	■							
Res3 HN Res2 HA#			■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

[a] Structures were calculated using "original_nnoe.tbl" and "original_caronoe.tbl".

[b] Structures were calculated using "original_dplus_nnoe.tbl"and "original_dplus_caronoe.tbl".

[c] Structures were calculated using the indicated subsets of NOEs (shaded boxes) derived from the file "all_nnoe.tbl" and all of the aromatic NOEs listed in the file "all_caronoe.tbl".

[d] Structures were calculated using the indicated subsets of NOEs (shaded boxes) and only aliphatic inter residue NOEs. All aromatic intraresidue NOEs were included.

[e] Structures were calculated without the use of any aliphatic or aromatic intraresidue NOEs.

CNS Run 2
Fifty structures were calculated during Run 2 using the NOE list files 'original_dplus_nnoe.tbl' and 'original_dplus_caronoe.tbl'. These '.tbl' files had been modified from the original '.tbl' files through the addition of 1.0 to dplus for methyl and methylene pseudoatoms, and 2.5 to dplus for aromatic pseudoatoms.
CNS Run 3
For Run 3, the NOE list files 'all_nnoe.tbl' and 'all_caronoe.tbl' were used to calculate 50 structures. These list files had been modified from those used in Run 2 through the removal of all 'OR' statements from the aliphatic NOE list and all redundant peaks from the aromatic NOE list. The content of these files are included in Appendix E.
CNS Run 4 → Run 15
A series of structure calculation runs were launched where the list of aliphatic NOEs was varied and the list of aromatic NOEs ('all_caronoe.tbl') was held constant. A new aliphatic NOE list file was created for each run. These files were named 'r4_nnoe.tbl' through 'r15_nnoe.tbl' and the NOEs included in each of the files are shown in Table 1.2. Fifty structures were calculated during each run.
CNS Run 16
This structure calculation run was conducted using an aliphatic NOE list ('r16_nnoe.tbl') where all intraresidue NOEs were removed, and the aromatic NOE list 'all_caronoe.tbl' (which did contain intraresidue NOEs). Fifty structures were calculated.
CNS Run 17
In Run 17 one hundred structures were calculated without the use of any intraresidue NOEs, so the only NOE list file that was used was 'r16_nnoe.tbl'.
CNS Run 18
One hundred structures were generated without the use of any NOE constraints.

Below are listed the contents of the unmodified '.tbl' files generated by *NMRView* listing the experimentally determined NOE constraints for cubic phase-bound MetEnk. The first file contains all of the aliphatic NOE data, and the second file contains the aromatic NOE data.

original_nnoe.tbl

```
assign (resid 2 and name HN) (resid 1 and name HB#)   5.000 3.2 0.0 !  new_noe_list.2
assign (resid 2 and name HN) (resid 2 and name HA#)   3.400 1.6 0.0 !  new_noe_list.3
assign (resid 4 and name HN) (resid 4 and name HB#)   2.800 1.0 0.0 !  new_noe_list.6
assign (resid 4 and name HN) (resid 2 and name HA1)   2.800 1.0 0.0 !  new_noe_list.7
or (resid 4 and name HN) (resid 2 and name HA2)
or (resid 4 and name HN) (resid 3 and name HA1)
or (resid 4 and name HN) (resid 3 and name HA2)
assign (resid 4 and name HN) (resid 4 and name HD1)   3.400 1.6 0.0 !  new_noe_list.9
or (resid 4 and name HN) (resid 4 and name HE1)
or (resid 4 and name HN) (resid 4 and name HD2)
or (resid 4 and name HN) (resid 4 and name HE2)
assign (resid 5 and name HN) (resid 5 and name HE1)   5.000 3.2 0.0 !  new_noe_list.12
or (resid 5 and name HN) (resid 5 and name HB2)
assign (resid 5 and name HN) (resid 5 and name HE1)   5.000 3.2 0.0 !  new_noe_list.13
or (resid 5 and name HN) (resid 5 and name HB2)
assign (resid 5 and name HN) (resid 5 and name HG#)   3.400 1.6 0.0 !  new_noe_list.14
assign (resid 5 and name HN) (resid 4 and name HB#)   5.000 3.2 0.0 !  new_noe_list.15
assign (resid 5 and name HN) (resid 4 and name HD1)   5.000 3.2 0.0 !  new_noe_list.18
or (resid 5 and name HN) (resid 4 and name HE1)
```

or (resid 5 and name HN) (resid 4 and name HD2)
or (resid 5 and name HN) (resid 4 and name HE2)
assign (resid 3 and name HN) (resid 3 and name HA1) 2.800 1.0 0.0 ! new_noe_list.23
or (resid 3 and name HN) (resid 3 and name HA2)
or (resid 3 and name HN) (resid 2 and name HA1)
or (resid 3 and name HN) (resid 2 and name HA2)
assign (resid 2 and name HN) (resid 1 and name HA) 3.400 1.6 0.0 ! new_noe_list.1
assign (resid 2 and name HN) (resid 3 and name HN) 5.000 3.2 0.0 ! new_noe_list.4
assign (resid 4 and name HN) (resid 3 and name HN) 3.400 1.6 0.0 ! new_noe_list.27
assign (resid 4 and name HN) (resid 4 and name HA) 3.400 1.6 0.0 ! new_noe_list.8
assign (resid 5 and name HN) (resid 4 and name HA) 2.800 1.0 0.0 ! new_noe_list.17
assign (resid 4 and name HN) (resid 5 and name HN) 3.400 1.6 0.0 ! new_noe_list.28
assign (resid 5 and name HN) (resid 5 and name HA) 2.800 1.0 0.0 ! new_noe_list.16
assign (resid 5 and name HN) (resid 5 and name HB1) 5.000 3.2 0.0 ! new_noe_list.11

original_caronoe.tbl

assign (resid 1 and name HD#) (resid 1 and name HE#) 3.400 1.6 0.0 ! all_cnoesyaro_pklist.0
assign (resid 1 and name HD#) (resid 1 and name HE#) 3.400 1.6 0.0 ! all_cnoesyaro_pklist.1
assign (resid 4 and name HE#) (resid 4 and name HZ) 5.000 3.2 0.0 ! all_cnoesyaro_pklist.3
assign (resid 4 and name HZ) (resid 4 and name HE#) 5.000 3.2 0.0 ! all_cnoesyaro_pklist.5
assign (resid 4 and name HD#) (resid 4 and name HA) 5.000 3.2 0.0 ! all_cnoesyaro_pklist.6
assign (resid 4 and name HD#) (resid 4 and name HB#) 5.000 3.2 0.0 ! all_cnoesyaro_pklist.7
assign (resid 1 and name HD#) (resid 1 and name HB#) 3.400 1.6 0.0 ! all_cnoesyaro_pklist.10
assign (resid 1 and name HD#) (resid 1 and name HA) 5.000 3.2 0.0 ! all_cnoesyaro_pklist.11

Below are listed the contents of the modified '.tbl' files used during *CNS* structure calculation Run 3. The files have been modified from their original form through the addition of 'pseudoatoms' and 'dplus' wherever there was ambiguity in the peak assignments and through the removal of 'OR' statements and redundant peak listings.

all_nnoe.tbl

assign (resid 2 and name HN) (resid 1 and name HB#) 5.000 3.2 1.0 ! new_noe_list.2
assign (resid 2 and name HN) (resid 2 and name HA#) 3.400 1.6 1.0 ! new_noe_list.3
assign (resid 4 and name HN) (resid 4 and name HB#) 2.800 1.0 1.0 ! new_noe_list.6
assign (resid 4 and name HN) (resid 2 and name HA#) 2.800 1.0 1.0 ! new_noe_list.7
assign (resid 4 and name HN) (resid 3 and name HA#) 2.800 1.0 1.0 ! new_noe_list.7
assign (resid 4 and name HN) (resid 4 and name HD#) 3.400 1.6 2.5 ! new_noe_list.9
assign (resid 4 and name HN) (resid 4 and name HE#) 3.400 1.6 2.5 ! new_noe_list.9
assign (resid 5 and name HN) (resid 5 and name HE1) 5.000 3.2 0.0 ! new_noe_list.12
assign (resid 5 and name HN) (resid 5 and name HB2) 5.000 3.2 0.0 ! new_noe_list.12
assign (resid 5 and name HN) (resid 5 and name HE1) 5.000 3.2 0.0 ! new_noe_list.13
assign (resid 5 and name HN) (resid 5 and name HB2) 5.000 3.2 0.0 ! new_noe_list.13
assign (resid 5 and name HN) (resid 5 and name HG#) 3.400 1.6 1.0 ! new_noe_list.14
assign (resid 5 and name HN) (resid 4 and name HB#) 5.000 3.2 1.0 ! new_noe_list.15
assign (resid 5 and name HN) (resid 4 and name HD#) 5.000 3.2 2.5 ! new_noe_list.18
assign (resid 5 and name HN) (resid 4 and name HE#) 5.000 3.2 2.5 ! new_noe_list.18
assign (resid 3 and name HN) (resid 3 and name HA#) 2.800 1.0 1.0 ! new_noe_list.23
assign (resid 3 and name HN) (resid 2 and name HA#) 2.800 1.0 1.0 ! new_noe_list.23
assign (resid 2 and name HN) (resid 1 and name HA) 3.400 1.6 0.0 ! new_noe_list.1
assign (resid 2 and name HN) (resid 3 and name HN) 5.000 3.2 0.0 ! new_noe_list.4

assign (resid 4 and name HN) (resid 3 and name HN)	3.400 1.6 0.0 !	new_noe_list.27
assign (resid 4 and name HN) (resid 4 and name HA)	3.400 1.6 0.0 !	new_noe_list.8
assign (resid 5 and name HN) (resid 4 and name HA)	2.800 1.0 0.0 !	new_noe_list.17
assign (resid 4 and name HN) (resid 5 and name HN)	3.400 1.6 0.0 !	new_noe_list.28
assign (resid 5 and name HN) (resid 5 and name HA)	2.800 1.0 0.0 !	new_noe_list.16
assign (resid 5 and name HN) (resid 5 and name HB1)	5.000 3.2 0.0 !	new_noe_list.11

all_caronoe.tbl

assign (resid 1 and name HD#) (resid 1 and name HE#)	3.400 1.6 2.5 !	all_cnoesyaro_pklist.1
assign (resid 4 and name HZ) (resid 4 and name HE#)	5.000 3.2 2.5 !	all_cnoesyaro_pklist.5
assign (resid 4 and name HD#) (resid 4 and name HA)	5.000 3.2 2.5 !	all_cnoesyaro_pklist.6
assign (resid 4 and name HD#) (resid 4 and name HB#)	5.000 3.2 2.5 !	all_cnoesyaro_pklist.7
assign (resid 1 and name HD#) (resid 1 and name HB#)	3.400 1.6 2.5 !	all_cnoesyaro_pklist.10
assign (resid 1 and name HD#) (resid 1 and name HA)	5.000 3.2 2.5 !	all_cnoesyaro_pklist.11

Inspection of Structures for NOE Violations

The first step in evaluating the structures generated by *CNS* was to check the structures for NOE violations. To do so a script called 'fileout_overall.awk' was run on all of the structures generated during each structure calculation run. The script was used to generate a file listing the names of all of the structures generated in a particular structure calculation run and their overall energy. These files were imported into an *Excel* workbook as separate worksheets and an energy sorted list of structures was generated for each structure calculation run. The '.pdb' files for each calculated structure were then inspected and the dihedral RMSD value was recorded, along with the presence or absence of any NOE violations. The dihedral RMSD was recorded because it is a number that should be unique to each calculated structure (it is a 5 digit number) making it possible to search the 'anneal.out' files generated during each *CNS* run for information related to a particular structure. In the case of structures that *did* contain NOE violations, the dihedral RMSD was used to search the corresponding 'anneal.out' file for information about the violated NOE(s). Four of the calculated structures were found to have NOE violations: r8_50_30.pdb, r12_50_30.pdb, r13_50_30.pdb and r14_50_30.pdb. These files were renamed: VNOE_r8_50_30.pdb, VNOE_r12_50_30.pdb, VNOE_r13_50_30.pdb and VNOE_r14_50_30.pdb and were not included in any further structural analyses.

Files Related to ENSEMBLE Analyses of Calculated Structures

Ensemble Input Files

In order to perform an analysis using *Ensemble* it was necessary to first prepare several input files. One of these files was the configuration file called 'config.par'. This file was supplied with *Ensemble* and needed to be modified to meet the needs of the analyses to be conducted. An analysis was conducted using both NOE *and* chemical shift constraints. The pool of structures that was analyzed included all of the structures calculated for cubic phase-bound MetEnk using NOE derived constraints *as well as* the 100 structures that were calculated without the use of any constraints. A copy of the 'config.par', 'cs.par' and 'noe.par' files used for the *Ensemble* analysis are included below. A file called 'all_structures_pdb.nam' which listed all of the structures to be analyzed was prepared using the utility program *makepdblist* that was provided with *Ensemble.*

A file named 'noe.con' listing of *all* of the experimentally determined NOEs was prepared in the format required by *Ensemble* and used for the analysis. A copy of this 'noe.con' file is included on the following pages. A file listing of all of the chemical shifts that had been determined for cubic phase-bound MetEnk was prepared in the format required by *Ensemble* and named 'cs.con'. A copy of 'cs.con' can be found on the following pages.

config.par

```
! This is the configuration parameter file for ENSEMBLE
! The name of this file should be passed to ENSEMBLE as the first
! command-line argument.
! All fields are required unless otherwise noted.

!------------------------------------------------------------------------

! The name of pdb list file, which contains the names of the pdb files
! in the candidate structure pool. If the full file path to the PDB list
! is not provided, it will be assumed that the file specified below is
! located in <ENSEMBLE directory>/pdblist/.

PDB_LIST         all_structures_pdb.nam

!------------------------------------------------------------------------

! The directory to be prepended to any files in the pdb list
! that do not begin with a slash.
! Leave this field empty to prepend <ENSEMBLE directory>/pdb/

PDB_DIRECTORY

!------------------------------------------------------------------------

! Nonzero to receive important email notices from the runtime environment.
! Address only required with nonzero email flag.

EMAIL_FLAG       0
EMAIL_ADDRESS    your@email.firewalls.may.cause.problems

!------------------------------------------------------------------------

! Nonzero values indicate constraint employment.
! Specific constraint configuration files must exist.

ACCESS_FLAG  0
```

```
! ATCUN_FLAG       // This constraint type is currently unavailable.
CDDECON_FLAG    0
COOP_SP_FLAG 0
CS_FLAG         1
CSI_FLAG        0
DIFFUSION_FLAG  0
ENTROPY_FLAG    0
J_FLAG          0
NIR_FLAG        0
NOE_FLAG        1
SAX_FLAG        0

!-------------------------------------------------------------------------

! Individual configuration file names are only required if the associated flag
! is nonzero. Use no directory structure for files to be found in
! <ENSEMBLE directory>/config. Otherwise, full directory from root is required.

ACCESS_CONFIG_FNAM         access.par
! ATCUN_CONFIG_FNAM            // This constraint type is currently unavailable
CDDECON_CONFIG_FNAM        cddecon.par
COOP_SP_CONFIG_FNAM        coop_sphere.par
CS_CONFIG_FNAM             cs.par
CSI_CONFIG_FNAM            csi.par
DIFFUSION_CONFIG_FNAM      diffusion.par
ENTROPY_CONFIG_FNAM        entropy.par
GRAD_DESCENT_CONFIG_FNAM   descent.par
J_CONFIG_FNAM                    j.par
NIR_CONFIG_FNAM                  nratio.par
NOE_CONFIG_FNAM                  noe.par
SIM_ANNEAL_CONFIG_FNAM           sim_anneal.par
SAX_CONFIG_FNAM                  sax.par

!-------------------------------------------------------------------------

! Select an optimization algorithm
! 1=modified gradient descent
!   *Much faster
!   *Theoretically can get stuck in local minima (not yet observed)
!   *Observed to produce round number conformer weights-- unexplained activity.
! 2=simmulated annealing  (DEFAULT)
!     *Original algorithm
!     *Avoids local minima
! Enter a single digit for a simple minimization
! Enter a string of digits to go from one minimization algorithm to another.
! There must be no spaces between digits.
! e.g. 1212 directs minimization by GD->SA->GD->SA

ALGORITHM       2

!-------------------------------------------------------------------------

! Your ENSEMBLE minimization can be re-initiated if it was terminated
! prematurely. To do this, set the rounds to be skipped here, and load
! the last saved weight file using INITIAL_WEIGHT_FILE below.
! If you were using the simulated annealing algorithm,
! rounds to skip = weight file extension.
! If you were using the gradient descent algorithm,
! rounds to skip = weight file extension * GD_SAVEMULT (descent.par)

ROUNDS_TO_SKIP   0
```

```
!-------------------------------------------------------------------------

! Option for how to set initial Ensemble weights.
! 1=random; 2=even distribution; 3=from file.
! Weight file name only required if it will be utilized.

INITIAL_RANDOM_EVEN_FILE     1
INITIAL_WEIGHT_FILE

!-------------------------------------------------------------------------

! If you are loading weights from a file, ENSEMBLE will predict the file
! before weight normalization if you set DELAY_NORMALIZATION non-zero integer.
! If delay_normalization is zero, the ensemble will be normalized prior to
! prediction.

DELAY_NORMALIZATION          0

!-------------------------------------------------------------------------

! Base name of the files to save ensemble weights.
! Intermediate weight files will be appended by .### (rounds)
! Final weight file will be appended by .ens

OUTPUT_FILE                  ensemble

!-------------------------------------------------------------------------

! Nonzero to save the ensemble weights after each iteration.
! Zero to save only after completion.
! Note:If there is an error during your run, and if you have saved intermediate
! ensembles,you can begin another run by loading in you previously best results

SAVE_INTERMEDIATE_ENSEMBLES   1

!-------------------------------------------------------------------------

! Nonzero to compress weight files after they are written.
! You must specify the file compressor from root in ZIP_COMMAND
! Set ZIP_WEIGHT_FILES to zero for no compression.

ZIP_WEIGHT_FILES   0
ZIP_COMMAND

!-------------------------------------------------------------------------

! With verbose piping set to a non-zero value, all information
! that derives from a child process is printed (e.g. shiftx, cns).
! Note that even when this verbose mode is off (recommended) error messages
! from cns are transmitted by recognition of the "ERR" string.

VERBOSE_PIPING               0

!-------------------------------------------------------------------------

! With verbose conformer prediction set to a non-zero value, all information
! about all predictions will be displayed during setup.
! If you have a large conformer pool,this will result in a very large log file.
! Default is zero.

VERBOSE_CONFORMER_PREDICTION     0
!-------------------------------------------------------------------------
```

```
! Use nonzero value to allow conformer predictions to be loaded only when the
! prediction file was created after the pdb file.
! Note that the solvent accessibility data is stored within the pdb file
! (in the B-factor column) and ENSEMBLE is unable to determine its validity.

FILE_DATE_CHECKING          0

!-------------------------------------------------------------------------------

! ENSEMBLE will load extraneous conformer values and manipulate the data to
! assist ensemble solution analysis. If you would like to track the weighted
! average of ca_ca RMSDs of conformers from the folded structure, create a
! data file containing one ca_ca RMSD value per line, the order corresponding
! to that of the pdb files in the pdblist file, and name this file from root
! after the CA_CA_RMSD_FILE specifier.
! Leave this field blank if you will not use this feature.

CA_CA_RMSD_FILE

! If you would like to track the weighted average of conformer radius of
! gyration, create a data file containing one radius of gyration value per
! line, the order corresponding to that of the pdb files in the pdblist file,
! and name this file from root after the RGYR_FILE specifier.
! Leave this field blank if you will not use this feature.

RGYR_FILE

! If you would like to track the weighted average of one other data type,
! create a data file containing one data point per line, the order
! corresponding to that of the pdb files in the pdblist file, and name this
! file from root after the ABSTRACT_FILE specifier.
! Leave this field blank if you will not use this feature.

ABSTRACT_FILE

!-------------------------------------------------------------------------------

! A Nonzero RCO_FLAG causes the program to calculate the population average
! residue-specific relative contact order and the population average
! residue-specific number of contacts.
! This does not constrain your data set.
! Set RCO_FLAG to zero to disable these calculations.

RCO_FLAG        0

! Use no directory structure for files to be found in
! <ENSEMBLE directory>/config/ Otherwise, full directory from root is required.

RCO_CONFIG_FNAM      rco.par

!-------------------------------------------------------------------------------

! When ENSEMBLE creates histograms, it first finds minimum value and the
! maximum values. These bounds are then divided equally into the number of bins
! specified here. Default = 20.

HISTOGRAM_COLUMNS        20

!-------------------------------------------------------------------------------

! To ensure that all the pdb files are consistent, set this parameter
```

```
! to a non-zero value. To disable the check for pdb consistency, set this
! value to 0. (Default = 1)

CHECK_PDB_CONSISTENCY      1

!------------------------------------------------------------------------------

! There is an option to allow ENSEMBLE to create statistic files
! for each constraint/data type. In doing so, the conformer pool minimum,
! maximum, average, and standard deviation values will be calculated for each
! each data value (of a particular data type).
! To enable this feature, set the value below to 1.
! Otherwise, set the value to zero. (Default = 1)

CREATE_STATISTICS_FILES    1

!------------------------------------------------------------------------------

! There is an option to allow ENSEMBLE to create data spread files
! for most constraint/data types. In doing so, a separate folder will be
! created for each applicable data type. Within that folder, there will
! be a separate data spread file for each data value of that type.
! A data spread file will show how comformer values are distributed over a
! range of values.
! To enable this feature, set the value below to 1.
! Otherwise, set the value to zero. (Default = 1)

CREATE_DATA_SPREAD_FILES   1

!------------------------------------------------------------------------------
```

noe.par

```
! This is the configuration parameter file used to define the NOE
! constraint environment.
! This file is only utilized if the NOE_FLAG is nonzero.
! The NOE_FLAG is part of the main configuration file.

!------------------------------------------------------------------------------

! The name of the file that contains NOE constraints.
! The NOE constraint file has a required syntax,
! see the README file for details. If the full path to the constraint
! file is not provided, it will be assumed to be located in the
! <ENSEMBLE directory>/con/ directory.

NOE_FILE_NAME    noe.con

!------------------------------------------------------------------------------

! Binary file loading overrides and disallows regular individual file loading
! and saving. Binary file saving can occur after loading regular individual
! prediction files or after data creation. Enter a nonzero integer to load/save
! binary files, and then enter the file name from root.
! Binary loading disables binary saving.
! Binary files will be tested for ensemble size, constraint number and file
! size. However, the controls are possibly leaky.
! Proposed file name convention:
! noe.bin_pdblist.nam_noe.con

NOE_LOAD_BINARY    0
```

```
NOE_SAVE_BINARY    0
NOE_BINARY_FILE_NAME    <ENSEMBLE directory>/predict/noe/noe.bin_pdb.nam_noe.con

!------------------------------------------------------------------------------

! Linear modifier of the NOE penalization coefficient.
! This represents your faith that this type of constraint is based in real
! characteristics of the ensemble or the relative importance of this constraint
! type. It represents something about the constraint type as a group.
! This term is collective for all NOEs, so increasing the number of NOE
! constraints requires an increasing NOE_FAITH to retain correlative importance
! of individual NOE constraints and those of other constraint types.
! Use decimal or scientific notation to specify a real number.
! Default=2.5

NOE_FAITH       2.5

!------------------------------------------------------------------------------
! The energy term is modified as follows: RAW_NOE_ENERGY * NOE_FAITH
!------------------------------------------------------------------------------

! Due to the nature of the NOE interaction, conformer contributions are
! averaged to the inverse sixth power (or in fast timescale R^-3).
! Use 2 for r^-6 averaging (default), and 1 for r^-3 averaging.
! default = 2 (for r^-6 averaging)
! Enter an integer (1 or 2)

NOE_R3_R6       2

!------------------------------------------------------------------------------
```

cs.par

```
! This is the configuration parameter file for ENSEMBLE chemical shift
! constraints. This file is only utilized if the CS_FLAG is nonzero.
! The CS_FLAG is part of the main configuration file.

!------------------------------------------------------------------------------

! The name of the file that contains chemical shift constraints.
! The chemical shift constraint file has a required syntax,
! see the README file for details. If the full path is not provided,
! the file will be assumed to be located in the <ENSEMBLE directory>/con/
! directory.

CS_FILE_NAME      cs.con

!------------------------------------------------------------------------------

! Binary file loading overrides and disallows regular individual file loading
! and saving. Binary file saving can occur after loading regular individual
! prediction files or after data creation. Enter a nonzero integer to load/save
! binary files, and then enter the file name from root.
! Binary loading disables binary saving.
! Binary files will be tested for ensemble size, constraint number and file
! size. However, the controls are possibly leaky.
! Proposed file name convention:
! cs.bin_pdblist.nam_cs.con

CS_LOAD_BINARY      0
CS_SAVE_BINARY       0
```

```
CS_BINARY_FILE_NAME
/disk3/hpc/lsingha/Current_Analyses/ensemble/predict/cs/cs.bin_pdb.nam_cs.con

!------------------------------------------------------------------------------

! Linear modifier of the chemical shift penalization coefficient.
! This represents your faith that this type of constraint is based in real
! characteristics of the ensemble or the relative importance of this constraint
! type. It represents something about the constraint type as a group.
! This term is collective for all chemical shift constraints, so
! increasing the number of chemical shift constraints requires an increasing
! CS_FAITH to retain correlative importance of individual chemical shift
! constraints and those of other constraint types.
! Use decimal or scientific notation to specify a real number.
! Default=2.0

CS_FAITH          1.0

!------------------------------------------------------------------------------

! Linear modifier of the chemical shift penalization coefficient for individual
! types.
! Use decimal or scientific notation to specify a real number.
! Default = 1.0 for all.

CS_HA_FAITH       1.0
CS_HN_FAITH       0.25
CS_N15_FAITH      0.5
CS_CA_FAITH       2.0
CS_CB_FAITH       2.0
CS_CO_FAITH       1.0

!------------------------------------------------------------------------------
! The energy term is modified as follows:
! ((RAW_HA_ENERGY * CS_HA_FAITH) + (RAW_HN_ENERGY * CS_HN_FAITH) +
!  (RAW_N15_ENERGY * CS_N15_FAITH) + (RAW_CA_ENERGY * CS_CA_FAITH) +
!  (RAW_CB_ENERGY * CS_CB_FAITH) + (RAW_CO_ENERGY * CS_CO_FAITH)) * CS_FAITH
!------------------------------------------------------------------------------

! The maximum deviation of experimental result from predicted value
! of an individual conformer that is accepted to be real.
! Deviation values above this maximum are considered to involve human error.
! If this restriction is not desired, the maximum values may be set
! to arbitrarily large values.
! Use decimal or scientific notation to specify a real number.

CS_MAX_HA_DEV     100
CS_MAX_HN_DEV     100
CS_MAX_N15_DEV    200
CS_MAX_CA_DEV     200
CS_MAX_CB_DEV     200
CS_MAX_CO_DEV     200

!------------------------------------------------------------------------------

! The experimental absolute error.  Ensemble values are allowed to deviate from
! experimental values by this value with no energetic penalty.
! Use decimal or scientific notation to specify a real number.

CS_HA_ERROR       0.02
CS_HN_ERROR       0.02
CS_N15_ERROR      0.1
```

```
CS_CA_ERROR       0.1
CS_CB_ERROR       0.1
CS_CO_ERROR       0.1

!------------------------------------------------------------------------------

! Permissions are used to control file creation deletion and use from
! within the runtime environment.
! Nonzero indicates permission granted; zero denies permission.
! The create_cs_permission determines whether the ensemble program can
! call shiftx to create a cs prediction.

CS_LOAD_PERMISSION        1
CS_SAVE_PERMISSION        0
CS_CREATE_PERMISSION      0
CS_OVERWRITE_PERMISSION   0

!------------------------------------------------------------------------------

! If the permissions are given for loading cs files,
! please specify a directory to be searched for files.
! Leave blank for <ENSEMBLE directory>/out/

CS_DIRECTORY

!------------------------------------------------------------------------------

! Where to find the program ShiftX.
! Leave it blank to default to <ENSEMBLE directory>

SHIFTX_DIRECTORY    /disk3/hpc/lsingha/Current_Analyses/ensemble/shiftx/

!------------------------------------------------------------------------------
```

noe.con

```
! This file contains NOE restraints for MetEnk bound to monoolein cubic phases
! The listed restraints are those listed in "all_nnoe.tbl" and
! "all_caronoe.tbl" used in Run_3 of structure calculations using CNS
! Sequential NOEs
1       HA      2       HN      3.4     1.6     0       1       !
new_noe_list.1
1       HB1     2       HN      OR
1       HB2     2       HN      5       3.2     0       1       !
new_noe_list.2
2       HN      3       HN      5       3.2     0       1       !
new_noe_list.4
3       HN      4       HN      3.4     1.6     0       1       !
new_noe_list.27
4       HN      5       HN      3.4     1.6     0       1       !
new_noe_list.28
4       HA      5       HN      2.8     1       0       1       !
new_noe_list.17
4       HB1     5       HN      OR
4       HB2     5       HN      5       3.2     0       1       !
new_noe_list.15
! Medium Range NOEs
! Long Range NOEs
! Ambiguous NOEs
1       HD1     1       HA      OR
```

```
1       HD2     1       HA      5       3.2     0       1       !
all_cnoesyaro_pklist.11
1       HD1     1       HB1     OR
1       HD1     1       HB2     OR
1       HD2     1       HB1     OR
1       HD2     1       HB2     3.4     1.6     0       1       !
all_cnoesyaro_pklist.10
1       HD1     1       HE1     OR
1       HD1     1       HE2     OR
1       HD2     1       HE1     OR
1       HD2     1       HE2     3.4     1.6     0       1       !
all_cnoesyaro_pklist.1
2       HN      2       HA1     OR
2       HN      2       HA2     3.4     1.6     0       1       !
new_noe_list.3
2       HA1     4       HN      OR
2       HA2     4       HN      OR
3       HA1     4       HN      OR
3       HA2     4       HN      2.8     1       0       1       !
new_noe_list.7
2       HA1     3       HN      OR
2       HA2     3       HN      OR
3       HA1     3       HN      OR
3       HA2     3       HN      2.8     1       0       1       !
new_noe_list.23
4       HN      4       HA      3.4     1.6     0       1       !
new_noe_list.8
4       HN      4       HB1     OR
4       HN      4       HB2     2.8     1       0       1       !
new_noe_list.6
4       HN      4       HD1     OR
4       HN      4       HD2     OR
4       HN      4       HE1     OR
4       HN      4       HE2     3.4     1.6     0       1       !
new_noe_list.9
4       HD1     4       HA      OR
4       HD2     4       HA      5       3.2     0       1       !
all_cnoesyaro_pklist.6
4       HD1     4       HB1     OR
4       HD1     4       HB2     OR
4       HD2     4       HB1     OR
4       HD2     4       HB2     5       3.2     0       1       !
all_cnoesyaro_pklist.7
4       HZ      4       HE1     OR
4       HZ      4       HE2     5       3.2     0       1       !
all_cnoesyaro_pklist.5
4       HD1     5       HN      OR
4       HD2     5       HN      OR
4       HE1     5       HN      OR
4       HE2     5       HN      5       3.2     0       1       !
new_noe_list.18
5       HN      5       HA      2.8     1       0       1       !
new_noe_list.16
5       HN      5       HB1     OR
5       HN      5       HB2     OR
5       HN      5       HE1     5       3.2     0       1       !
new_noe_list.11
5       HN      5       HB1     OR
5       HN      5       HB2     OR
5       HN      5       HE1     5       3.2     0       1       !
new_noe_list.12
5       HN      5       HB1     OR
```

```
5       HN      5       HB2     OR
5       HN      5       HE1     5       3.2     0       1       !
new_noe_list.13
5       HN      5       HG1     OR
5       HN      5       HG2     3.4     1.6     0       1       !
new_noe_list.14
```

cs.con

```
! This file contains chemical shift data for MetEnk bound to monoolein
! cubic phases
! The chemical shifts listed are from the 15N-hsqc, 13C-hsqc and 13C-
!  aromatic hsqc experiments
Y     1.CA  57.511      1.0
Y     1.HA  4.246 1.0
Y     1.CB  38.819      1.0
Y     1.HB2 3.140 1.0
Y     1.HB1 3.140 1.0
Y     1.CD1 133.499     1.0
Y     1.HD1 7.121 1.0
Y     1.CE1 118.637     1.0
Y     1.HE1 6.824 1.0
Y     1.CE2 118.637     1.0
Y     1.HE2 6.824 1.0
Y     1.CD2 133.499     1.0
Y     1.HD2 7.121 1.0
G     2.N   112.515     1.0
G     2.HN  8.536 1.0
G     2.CA  45.177      1.0
G     2.HA2 3.954 1.0
G     2.HA1 3.910 1.0
G     3.N   108.188     1.0
G     3.HN  7.922 1.0
G     3.CA  45.177      1.0
G     3.HA2 3.954 1.0
G     3.HA1 3.910 1.0
F     4.N   119.609     1.0
F     4.HN  8.077 1.0
F     4.CA  57.802      1.0
F     4.HA  4.666 1.0
F     4.CB  40.008      1.0
F     4.HB2 3.095 1.0
F     4.HB1 3.095 1.0
F     4.CD1 132.097     1.0
F     4.HD1 7.254 1.0
F     4.CE1 131.075     1.0
F     4.HE1 7.254 1.0
F     4.CZ  129.276     1.0
F     4.HZ  7.149 1.0
F     4.CE2 131.075     1.0
F     4.HE2 7.254 1.0
F     4.CD2 132.097     1.0
F     4.HD2 7.254 1.0
M     5.N   121.228     1.0
M     5.HN  7.976 1.0
M     5.CA  55.460      1.0
```

```
M      5.HA   4.267 1.0
M      5.CB   32.952       1.0
M      5.HB2  2.081 1.0
M      5.HB1  1.940 1.0
M      5.CG   32.465       1.0
M      5.HG2  2.329 1.0
M      5.HG1  2.271 1.0
M      5.CE   17.163       1.0
M      5.HE1  2.028 1.0
```

Ensemble Output Files

Each *Ensemble* run generated a weight file, prediction files, data spread files and conformer pool statistics files. The files that were used to evaluate the results of the *Ensemble* runs were the weight files. The weight files listed each PDB filename in the conformer pool, along with the associated weight of that conformer in the form of a decimal number (where the weights of all of the conformers in a run summed to 1).