Studies of the Conserved Hydrophobic Region of the Prion Protein in a Lipid Environment by NMR

Daniel Buijs

Thesis submitted to the
Faculty of Graduate Studies and Research
In partial fulfillment of the requirements
For the M.Sc. degree in Chemistry

Department of Chemistry
Faculty of Science
Carleton University

© Daniel Buijs, Ottawa, Canada, 2009

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-52004-8
Our file Notre référence
ISBN: 978-0-494-52004-8

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

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

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

The progression of prion disease involves a conformational shift of normally folded prion protein, PrP^C, into an infectious, pathogenic form, PrP^{Sc}. Previous work has suggested that the conserved hydrophobic region (CHR) of the prion protein interacts directly with the cell membrane. A peptide encompassing the CHR of the prion protein was produced and studied by nuclear magnetic resonance (NMR) spectroscopy in dodecylphosphocholine (DPC) micelles intended to mimic a membrane environment. The peptide studied, PrP(111-136), showed significant conformational differences when prepared in DPC micelles compared with water. Structural studies by NMR were conducted and a preliminary ensemble of structures calculated for PrP(111-136). The structure set reported here shows an alpha helical structure, which is not observed in reported NMR structures of the prion protein in water. It is found that PrP(111-136) interacts with DPC micelles in a specific way, and that future studies to better characterize this interaction are required.

Acknowledgements

I'd like to thank Geneviève Gingras and Simon Sauvé for their work in the lab and support and encouragement. Thanks also to Lisa Waldron and Terry Cyr for mass spectrometry services, Michel Girard and Barry Lorbetskie for the use of their HPLC.

Special thanks to my supervisor Yves Aubin, without whom, the work presented here would not have been possible.

Thanks also to Melinda Wong and my parents Rosanne and Joseph Buijs for their support, patience, and constant encouragement.

Table of Contents

Abstract	ii
Acknowledgements.....	iii
Table of Contents	iv
List of Tables	v
List of Figures.....	vi
Introduction.....	1
Transmissible Spongiform Encephalopathies.....	1
Prion Protein	4
Nuclear Magnetic Resonance Spectroscopy.....	11
Structure Determination	18
Sample considerations	19
Materials and Methods.....	21
Preparation of Expression Vectors.....	21
Expression and Isotope Labeling	21
Purification.....	23
HPLC Purification.....	24
NMR Sample Preparation in Bicelles	25
NMR Sample Preparation in Micelles.....	26
NMR Experiments.....	26
HSQC	27
¹⁵ N-edited NOESY	27
¹⁵ N-edited TOCSY	28
3D HNCO and HN(CA)CO.....	29
3D HNCACB and CBCA(CO)NH.....	29
Structure Calculation.....	29
Results and Discussion	31
Expression and Purification of Prion Peptides.....	31
DMPC/DHPC Bicelles and DPC Micelles	34
Structure Calculation.....	47
¹³ C, ¹⁵ N-Labeled PrP(110-136).....	52
Conclusions and Future Work	56
Bibliography	58

List of Tables

Table 1: Oligonucleotides used to assemble the expression cassette.....	22
Table 2: Expected and observed molecular masses as confirmed by mass spectrometry.	34
Table 3: Tentative Chemical Shift Assignments of PrP(111-136) used in Structure Calculations.....	46
Table 4: Statistics from structure calculation of HuPrP(111-136).....	52

List of Figures

Figure 1: The solution structure of human PrP ^C as determined by solution NMR.	6
Figure 2: A schematic of the prion peptide expression constructs.....	22
Figure 3: SDS-PAGE Gels showing expression and purification of PrP(110-136).....	33
Figure 4: An HPLC chromatogram of the TEV cleavage reaction.....	35
Figure 5: HSQC of PrP(110-129) in water and DPC micelles	36
Figure 6: Comparison of glycine amide resonances	36
Figure 8: HSQC spectra of PrP(111-136) in DMPC/DHPC bicelles and DPC micelles. .	38
Figure 9: HSQC spectra of PrP(110-136) in LMPG and SDS at varying pH	40
Figure 10: The HSQC of PrP(111-136) in DPC micelles at different pH.....	42
Figure 11: A partial strip plot of the ¹⁵ N-edited NOESY experiment, demonstrating sequential assignment through NOEs to adjacent amides	44
Figure 12: Schematic of expected and observed NOEs in an alpha helix.....	48
Figure 13: Superimposed backbone structures calculated for PrP(111-136).	50
Figure 14: HSQC of PrP(110-136) in DPC micelles at pH 7.6.	54

Introduction

Transmissible Spongiform Encephalopathies

Public attention has been focused in recent years on a group of neurological diseases found in several different species of mammals, including humans. These diseases, which cause a wide array of neurological symptoms, are related in that they have all been correlated with the presence of an abnormally folded protein in the brain tissue. Post-mortem examination of the brain tissue of affected individuals shows holes throughout the affected areas, from which they get their name: spongiform encephalopathies(1).

These diseases were initially identified individually, based on behavioral diagnoses. Spongiform encephalopathies can be divided into three broad groups, based on the patterns of their appearance in the populations they affect. These groups are sporadic spongiform encephalopathies, familial spongiform encephalopathies, and transmissible spongiform encephalopathies(2).

Spongiform encephalopathies are said to be of sporadic origin when they cannot be attributed to either familial or transmissible sources. They are largely a diagnosis of exclusion and some researchers argue whether there are in fact any truly sporadic cases of disease. Nevertheless, classical Creutzfeldt-Jakob disease in humans is classified as a sporadic prion disease. Sporadic prion disease may be caused by an acquired somatic mutation in the gene for the prion protein or by spontaneous conversion of normally

folded prion protein into a pathogenic form(3). Classical Creutzfeldt-Jakob disease has a yearly mortality rate of approximately 1 per million people in Canada(4).

Familial spongiform encephalopathies are hereditary and have been linked to particular mutations in the gene for the prion protein. The familial spongiform encephalopathies include Gerstmann-Sträussler-Scheinker Syndrome (GSS), Fatal Familial Insomnia (FFI), and Familial Creutzfeldt-Jakob disease (fCJD)(3). Each of these disease phenotypes has been associated with several particular prion genotypes, including mutations and natural polymorphisms(5). Among others, GSS has been associated with A117V, meaning that the alanine (A) normally found at position 117 has been replaced with a valine (V), and G131V, and fCJD and FFI are both influenced by the polymorphism at position 129, methionine in this position being associated with FFI and valine being associated with fCJD.(5) All three of these positions lie within the Conserved Hydrophobic Region (CHR) of the prion protein, spanning residues 113 to 128(3).

Transmissible spongiform encephalopathies (TSEs) follow a more predictable etiology, being transmitted between individuals. The transmissible nature of TSEs was recognized well before the causative agent was identified. The oldest known TSE is scrapie in sheep(3). This disease was known to be transmissible among flocks of sheep and was characterized by animals scraping themselves against rocks or trees, sometimes removing their coats in the process. To date, scrapie has been confined to sheep, although the infectious agent is quite hardy, with flocks being infected even after a pasture had lain

fallow for many years(6). Of much more concern to human health is bovine spongiform encephalopathy (BSE), which has been observed to jump species to mice, pigs, mink, monkeys under laboratory conditions, and presumably to humans as variant Creutzfeldt-Jakob Disease (vCJD)(3). Human to human transmission of TSE has been observed in iatrogenic Creutzfeldt-Jakob disease (iCJD), associated with prion contaminated dura mater grafts and human-sourced human growth hormone(3). A TSE seen only in humans is kuru, identified among the Fore people of Papua New Guinea. Animal forms of TSE have also been identified in mink, elk, deer, domestic cats, kudu, nyala, and oryx(3).

The mechanism of transmission of most TSEs is through ingestion of infected tissue and ultimately of a pathogenic form of the prion protein. As discussed below, the infectious form of the prion protein is extremely resistant to heat, pH, and proteolytic enzymes, and thus survives the trip through the gastrointestinal tract of the animals that are susceptible to infection. There is no evidence of transmission through the air, through contact with bodily fluids, or through sexual contact, although the mechanism of transmission of Chronic Wasting Disease (CWD) in elk is currently unknown, with no evidence of cannibalism evident in this species(7).

All forms of prion disease are associated with progressive neurodegeneration, followed by death. Patients and animals who have died of prion disease are found to have multiple holes in their brain tissue, associated with plaques of protein. These plaques have been found to contain a particular form of the prion protein.

Prion Protein

It was only many years after the discovery of TSEs, with the advent of modern laboratory medicine that they were recognized as having similar brain pathologies, and they were all correlated to the presence of a misfolded protein, the prion protein(8). The prion protein has been identified as the causative agent of all TSEs(9). All species affected by TSEs have their own version of the prion protein, and this protein is found to be highly conserved across species. In fact, a version of the prion protein is present in almost all higher organisms.

The prion protein has been found to exist in multiple conformations. The cellular form of the prion protein, PrP^C, is normally the predominant form of prion found in organisms. A schematic of the amino acid sequence of the human prion protein is shown in Figure 1A. The most common human allele, with methionine at position 129 is shown. The protein consists of a long unstructured N-terminal containing so-called octa-repeat domains which have been found to have the ability to bind copper, and the conserved hydrophobic region (CHR), followed by a globular domain that is found to form three alpha helices and a short section of beta-sheet in solution, ending with a glycosphosphatidylinositol (GPI) anchor, attaching the protein to the cell membrane at the C-terminal. Also included in the globular domain are two glycosylation sites and a single disulfide bridge. The CHR has always been found to adopt a random coil configuration in nuclear magnetic resonance (NMR) solution structures. In organisms, PrP^C is a soluble monomer, which is sensitive to proteases and can be denatured by heat, like most proteins. The solution structures of the cellular form of the prion protein from

many species have been solved by solution NMR(10-15). The structure of the folded domain of human PrP^C, as studied by solution NMR, is shown in Figure 1B(14).

The scrapie form of the prion protein, PrP^{Sc}, is an amyloid, which is insoluble in water, and resistant to most proteases and heat denaturation. The precise structure of PrP^{Sc} has not yet been elucidated, however FTIR studies indicate that it has more beta-sheet content than PrP^C (16). The scrapie form of prion has been found in amyloid plaques located in the brain tissue of patients and animals with transmissible spongiform encephalopathies. This form of the prion protein is the agent responsible for transmissible spongiform encephalopathies and is capable of initiating and propagating a conformational change of PrP^C into PrP^{Sc}(3).

The CHR of the prion protein, located immediately prior to the first strand of beta-sheet in PrP^C is highly conserved between species and mutations in this region have been associated with different prion diseases or resistance to TSE(9).

Figure 1: The solution structure of human PrP^C as determined by solution NMR.

Alpha helices are shown in blue and beta sheet in red. A: The amino acid sequence shown indicates, in addition to alpha helix and beta sheet, random coil (grey), turns (green), and the single disulfide bridge (yellow). The conserved hydrophobic region (CHR) (17) is in green, while PrP(110-136) is shown in green and orange. B: The structure shown includes the folded domain (121-230). The unstructured region in the N-terminal (21-120) was found not to alter the folding of this domain.

Lingappa and coworkers studied interactions between the prion protein and the cell membrane(9). It was previously discovered that a small proportion of prion protein produced in cell-free expression systems remained associated with the membrane after cleavage of the GPI anchor(18), suggesting that it spans the membrane, and the Lingappa group was interested in studying these artificial forms of the prion protein. They found that mutations in the CHR that made this region more hydrophobic resulted in a greater proportion of the prion protein being produced in transmembrane form *in vitro*. Intrigued by these results, they introduced transgenic versions of the prion protein bearing these mutations into mice that did not express any endogenous prion protein. These mice, expressing the mutations K110I and H111I (KH->II), or A113V, A115V, and A118V (AV3), all showed clinical symptoms of neurodegeneration. Separately, Lingappa and coworkers found the same predisposition towards transmembrane forms *in vitro* in human PrP(A117V), a mutation associated with GSS in humans, and confirmed this finding in a preserved brain section from a GSS patient who was confirmed to bear the A117V genotype.

The work of Lingappa and coworkers demonstrates that interaction between the prion protein and the cell membrane is associated with clinical symptoms of disease and that mutations in the CHR that render this region more hydrophobic not only predispose the prion protein towards the transmembrane form, but are also associated with disease. Additionally, the highly conserved nature of this region, with respect to the rest of the protein, both within species and within individuals(3) and the association of mutations in

this region with human and animal disease suggest that structural studies of the CHR of the prion protein in a lipid environment are warranted.

Interactions between the CHR of the prion protein and a membrane may cause local changes in the structure of this region of the prion protein, which may impact the global structure of the whole protein. The objective of this project was to better characterize the interactions between the CHR of the prion protein and phospholipid micelles intended to mimic a biological membrane. In order to better characterize the structural behaviour of the CHR of the human prion protein in a membrane environment, a peptide encompassing the CHR was studied by solution NMR in phospholipid micelles.

Phospholipids consist of long hydrophobic tails and a hydrophilic head group. Lysophospholipids in solution form spherical structures in water called micelles, orienting their fatty tails inward and their polar or charged head groups outward to make contact with water. These spherical assemblies have an analogous arrangement to phospholipid bilayers that make up cell membrane. Mixtures of different phospholipids have also been developed comprising a phospholipid with a hydrophobic chain length approximately equal to those found in biological membranes and a phospholipid with a shorter chain length in a fixed ratio. These mixtures form disc shaped structures called bicelles, or bilayered micelles(19). Both micelle and bicelle systems have been used for structural studies of membrane proteins by solution NMR(20).

Structural studies of various prion peptides encompassing part of the CHR have been reported(21; 22), studying the conformational behaviour of peptides including PrP(113-127), PrP(113-120), and PrP(106-126) in organic solvents or sodium dodecyl sulfate (SDS) micelles under varying conditions and reported helical tendencies of these peptides by circular dichroism (CD) spectroscopy. Vold and co-workers studied a synthetic peptide, PrP(110-136), that encompasses the entire CHR and the first strand of the beta-sheet observed in the solution structure of the human prion protein. In this peptide, only the glycine residues were ^{15}N -enriched in DMPC/DHPC bicelles (23). While this allowed the group to study the conformational behaviour of the peptide, it did not provide sufficient NMR information to determine a high resolution structure. They observed that the glycine resonances of PrP(110-136) recorded in water shift drastically when the peptide is prepared in bicelles, suggesting the presence of structure. They also found that the structure was generally helical by CD.

In light of the results reported by Lingappa and coworkers, we had an interest in studying the conformational behaviour of the entire CHR without including the residues that participate in the first beta-sheet, in an environment that more closely mimics biological membranes than organic solvent or SDS micelles. A peptide spanning residues 110-129 would encompass the entire CHR, thus excluding the polar residues included in the Vold peptide, and could be studied in phospholipid micelles at physiological pH to more closely model the natural environment of the prion protein.

We developed a bacterial expression system to enable the production of samples of PrP(110-129) and PrP(110-136) with uniform ^{15}N labeling and/or ^{13}C , ^{15}N labeling, allowing a complete and detailed analysis of the structure of the CHR in a membrane-like environment.

Nuclear Magnetic Resonance Spectroscopy

Protein structure in a lipid environment can be studied with NMR. By preparing a protein sample in phospholipids that form micelles, a membrane environment can be simulated(24). Solution NMR is limited by particle size, with larger proteins more difficult to study than smaller ones. NMR spectroscopy of peptides in solution is well understood and quite common in the literature. The study of a peptide containing the CHR of the prion protein in lipid micelles could proceed using conventional techniques to study protein structure by NMR. The major difference is that the peptide-micelle complex would behave like a much larger protein, roughly equivalent to a protein of the same mass as the entire complex.

NMR Spectroscopy is an extremely powerful analytical technique that is sensitive to protein structure. At its heart, NMR detects and records the resonance frequencies of individual NMR-sensitive nuclei(25). These resonance frequencies, referred to as chemical shifts, are highly sensitive to the magnetic environment of the nucleus, which are in turn affected by the electrons surrounding the nucleus and the behavior of any other sensitive nuclei nearby. A folded protein experiences interactions between atoms and forms intra-molecular hydrogen bonds that are not observed in an unfolded protein. The specific pattern of interactions and the locations of the hydrogen bonds are unique to a particular structure. Because NMR is sensitive to individual nuclei, the information it provides can be used to calculate atomic-level protein structures(26).

The most commonly used NMR sensitive nucleus is the proton. With their high natural abundance (99.9884% of all hydrogen)(27), and importance in biological systems, protons are very useful for NMR studies of proteins. A basic one-dimensional proton NMR experiment can detect the signals of all protons present in the sample. The signals generated by each proton are additive however, meaning that the strongest NMR signal recorded from a protein sample dissolved in water will be from the solvent, overpowering the resonances of interest in the protein. Specialized NMR experiments have been developed to filter out or suppress solvent signals(25).

In addition to the need to isolate protein resonances from the solvent, the sheer number of unique resonances in a protein sample make it impossible to resolve each individual resonance in a one-dimensional experiment. This difficulty is present in any sufficiently complex sample and for this reason multidimensional NMR experiments were developed(25). Multidimensional NMR experiments allow individual resonances to be resolved in additional dimensions based on specific interactions with other nuclei, correlating a resonance in one dimension with one or more other resonances in another dimension. Multidimensional experiments allow very similar or identical resonances to be completely resolved in two or more dimensions, which would not otherwise be possible with one-dimensional NMR. The spectra from multidimensional NMR experiments are depicted as contour maps, allowing resonances to be more clearly visualized.

Multidimensional experiments have been developed that separate and group the proton, nitrogen and carbon resonances of protein samples in structurally relevant ways. A protein has one amide group per peptide bond, consisting of a proton-nitrogen amide pair. A selective NMR experiment that records only the resonances of protons that are directly bound to ^{15}N has been developed, called the ^1H - ^{15}N Heteronuclear Single Quantum Coherence (HSQC) experiment(25). The two-dimensional spectra produced by this experiment show a single peak for each amide pair in the sample, correlating the proton resonance frequency in the direct dimension with the nitrogen resonance frequency in the indirect dimension. When used with a uniformly ^{15}N -labeled protein sample, the HSQC spectrum will show a peak for each amino acid residue in the sample (with the exception of proline residues and amide protons that are in rapid exchange with the solvent). The HSQC can detect any proton bound to an ^{15}N nucleus, including the terminal amines on lysine, arginine, glutamine and asparagines residues, although these resonances are usually easily distinguished from amide resonances due to their highly characteristic nitrogen chemical shifts. The HSQC experiment is a very useful tool during the optimization step of solution conditions of samples (see section on sample preparation). The resonances of proton-nitrogen pairs are very sensitive to local environment, and can be used as qualitative indicators of the conformation of the peptide(28). This experiment can also be used to monitor sample stability and optimize NMR parameters such as pulse lengths and relaxation delays.

The ^{15}N isotope has a low natural abundance of 0.38%(27), which makes it desirable to artificially enrich protein samples used for structural NMR studies with this

isotope. Using a bacterial expression system, it is possible to uniformly incorporate ^{15}N and/or ^{13}C into any recombinant protein produced, by using growth media with a single source of nitrogen and a single source of carbon, such as M9 minimal media (NH_4Cl and glucose, respectively). If these simple sources of nitrogen and carbon in the growth media are enriched with the appropriate isotope, every protein the bacteria express will be uniformly labeled with these isotopes. The use of ^{15}N -labeled peptide for NMR studies in lipid micelles allows ^{15}N -edited experiments to be performed, which are very effective in filtering out resonances of nuclei that are not bound to ^{15}N . Without the use of ^{15}N -edited experiments, the massive over-abundance of lipid with respect to peptide would make it very difficult to detect signals from the peptide.

The three-dimensional NMR experiments used in this work exploit interactions between nuclei that occur through bonds (scalar coupling) or through space (nuclear Overhauser enhancement)(25). NMR-sensitive nuclei that share electrons through bonds will interact with each other in a phenomenon called scalar coupling. The simplest NMR experiment that exploits this interaction is called COrrrelation SpectroscopY (COSY), which correlates the resonances of each pair of nuclei that are directly coupled to each other(25). An extension of the COSY experiment used here is TOtal Correlation SpectroscopY (TOCSY), which allows magnetization energy to be relayed through a chain of coupled nuclei, correlating the directly detected nucleus to each nucleus in the chain, called a spin system. The TOCSY spectrum of a protein sample can correlate the resonances of each side-chain proton to the corresponding amide proton resonance (29). The three-dimensional version of the TOCSY experiment used here, called an ^{15}N -edited

TOCSY, uses an HSQC to record only those TOCSY resonances that have been relayed from the amide protons(26).

The ^{15}N -edited TOCSY experiment, which can correlate the resonances of each side-chain to the chemical shift of the amide proton, is the preferred data source to determine the amino acid type of each amide proton. The data from this experiment can be used to identify the amino acid type of each unassigned residue by comparing the pattern of resonances observed with the chemical shifts reported for the corresponding amino acid in a random coil(30).

When a ^{15}N , ^{13}C , doubly labeled sample is available, triple resonance experiments greatly simplify the task of resonance assignment. Intra-residue triple resonance experiments correlate the amide proton and nitrogen resonances to one or more carbon or hydrogen resonances in the preceding residue, while inter- and intra- residue triple resonance experiments correlate the amide proton and nitrogen resonances to one or more carbon and hydrogen resonances in both the amide residue (i) and in the preceding residue (i-1). The combination of an intra-residue experiment with an inter- and intra-residue experiment allows the resonances of each residue to be assigned unambiguously not just by amino acid type, but also in the correct sequence.

The triple resonance experiments used to confirm the backbone resonance assignments of PrP(110-136) included the HNCO, the HN(CA)CO, the CBCA(CO)NH, and the HNCACB(31). These experiments are named for the resonances they record. The

HNCO and CBCA(CO)NH are inter-residue experiments, correlating the amide proton and nitrogen resonances of each residue with the carbonyl carbon or beta and alpha carbons of the preceding residue, respectively(29). The HN(CA)CO and HNCACB experiments are inter- and intra-residue experiments which correlate the amide proton and nitrogen resonances of each residue to the carbonyl carbon or alpha and beta carbons, respectively, of both the amide residue and the preceding residue(29).

In addition to interacting through bonds, NMR sensitive nuclei also interact through space, in a distance-dependent way. The measurable phenomenon exploited in these experiments is called the Nuclear Overhauser Enhancement (NOE), and it is observed due to dipole-dipole interactions between sensitive nuclei through space. The simplest experiment that exploits this interaction is called Nuclear Overhauser Enhancement Spectroscopy (NOESY), which records the resonances of each pair of sensitive nuclei that are within 5 Å of each other in space(25). As with the TOCSY, in this study the NOESY was combined with an HSQC in an ^{15}N -edited NOESY experiment, detecting only those protons which are within 5 Å of the amide protons(26). In rare circumstances, an NOE may not be detected between two nuclei that are within 5 Å or an NOE will be detected between two nuclei that are further apart than 5 Å, due to a phenomenon called spin diffusion(29).

While approximate chemical shift ranges for the nuclei in amino acid residues have been published(32), the precise resonances of each NMR-sensitive nucleus in a protein sample are strongly influenced by the protein structure and must still be assigned

manually. If the protein has a relatively large number of amino acids, is in a relatively uniform environment, or contains regions with many amino acids of the same type in close proximity to each other, NMR resonance assignment can be complicated by overlapping peaks. Different nuclei that have indistinguishable chemical shifts are referred to as degenerate. The use of multi-dimensional experiments or the use of ^{13}C experiments in conjunction with a ^{13}C , ^{15}N doubly labeled sample can help overcome the challenges associated with highly degenerate spectra.

The first step of data analysis in a protein structure calculation project is to assign each resonance (or peak) to a particular pair of nuclei. Each NMR sensitive nucleus in a sample will have a unique chemical shift, characteristic of the magnetic environment around it. In the experiments recorded with PrP(111-136), the chemical shifts of each amide proton-nitrogen pair were recorded, and correlated with the chemical shifts of protons that were near the amide proton in space (^{15}N -edited NOESY), or protons that were in the same spin system as the amide proton (^{15}N -edited TOCSY). Each amide pair obtained from the peaks on the HSQC corresponds to an amino acid residue in the peptide. The indirect proton dimension of the ^{15}N -edited NOESY contains the chemical shifts of protons near the amide proton in space, usually including the side-chain protons, as well as the adjacent amides and any other protons that may be close to the parent amide as a result of secondary structure. The indirect proton dimension of the ^{15}N -edited TOCSY contains the chemical shifts of protons that are in the same spin-system as the parent amide, usually comprising the protons of the same amino acid residue as the parent.

Structure Determination

A protein structure can be uniquely defined as the three-dimensional co-ordinates occupied by each of its component atoms. The goal of experimental protein structure determination is to define the subset of physically possible structures that are consistent with the experimentally derived data. It is assumed that the desired subset occupies an energy minimum within the set of all possible structures defined by experimentally defined data, however it is not feasible to systematically evaluate each mathematically possible structure. Multiple models are constructed that satisfy all the constraints, which are then allowed to evolve in a molecular simulation, moving the models toward the largest local energy minimum. If the system is sufficiently constrained, all models will converge toward a single global energy minimum, consistent with the physically possible structures at a given energy level (i.e. 37°C). If the system is under-constrained, the models will tend toward different local energy minima, which will not be representative of the global minimum. If the system is over-constrained, bias will be introduced into the system and may lead to convergence of all models at a single local energy minimum which does not correspond to the true structure.

Current NMR techniques can be used to extract structural information about a protein in the form of discrete constraints. With sufficient constraints, the set of possible structures can be restricted to the point where randomly chosen structures from within the set will converge on a single high-resolution structure in molecular simulations. Protein NMR experiments can provide distance constraints between individual atoms, in the form

of NOEs and torsion angle restraints, predicted from the chemical shift values of the backbone and side chain nuclei(29).

In situations where the chemical shifts of a protein are highly degenerate, it can be impossible to unambiguously assign NOEs prior to structure calculation. The ARIA program can consider several different types of constraints in calculating structure, including distance constraints from NOEs and distance constraints from hydrogen bonds, among others. ARIA is designed to automatically assign NOE cross peaks in an iterative manner based upon the energy of the calculated structures.

Sample considerations

The application of multidimensional NMR techniques for structure determination requires the protein of interest to be soluble, mono-dispersed and stable for prolonged periods of time under experimental conditions. The solution conditions, including the type of buffer, the pH, the ionic strength, and co-solutes must be optimized for each protein in order to obtain the properties described above.

Protein size is a limiting factor in structure determination for two reasons. Firstly, the sheer number of NMR active nuclei makes peak resolution and assignment very challenging. These limitations can be overcome with higher dimensionality experiments and selective labeling. The second reason is that larger molecules tumble more slowly in solution. This leads to more rapid loss of the NMR signal, resulting in less intense, broader peaks.

For this project, small peptides encompassing the CHR of the prion protein were studied in phospholipid micelles, intended to model the interaction of this region of the prion protein with the cell membrane. While the size of the peptide itself is relatively small, the effective size of the peptide/micelle complex is much larger, leading to slower tumbling in solution, and the same overall NMR behaviour of a larger protein, in the range of 30-35 kDa. Although the study of this large system was technically challenging, by studying this peptide and its interactions with a model membrane system, new insights may be gained into the behaviour of the full-length prion protein when it interacts with physiological membranes.

Materials and Methods

Preparation of Expression Vectors

Expression plasmids containing PrP(110-129) and PrP(110-136) were designed and produced by Ms. Geneviève Gingras by annealing and ligation of synthetic oligonucleotides corresponding to the correct sequences (see Table 1) into a modified pET19b plasmid (Novagen). The complete construct included a ten-histidine tag (10xHIS) to facilitate purification, glutathione-S-transferase, a linker containing the recognition site for Tobacco Etch Virus protease (TEV), and the coding sequence for the peptide of interest (see Figure 2).

Expression and Isotope Labeling

The appropriate expression vector was freshly transformed into chemically competent *E. coli* BL21(DE3) cells (Stratagene) for each expression culture. A starter culture of 3 mL LB broth was inoculated and incubated for approximately 3 hours at 37°C. Once this culture reached an optical density at 600 nm (OD_{600}) of approximately 0.6, 1 mL was transferred to 200 mL of M9 minimal media (see below), which was incubated overnight. After approximately 15 hours this inoculum was diluted into 1L of pre-warmed M9 minimal media at 37°C. This expression culture was incubated until it reached a cell density of $OD_{600} \approx 0.8$ at which point protein synthesis was induced by the addition of 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG). Cells were harvested 3 hours post-induction by centrifugation at 3000 x g for 15 minutes and frozen at -20°C.

Table 1: Oligonucleotides used to assemble the expression cassette

Oligonucleotide	Sequence
PrP(110-129) Forward	GATCCGAGAA CCTCTATTTC CAGGGTAAAC ACGTCGCAGG TGCCGCGGCA GCAGGAGCAG TGGTAGGGGG GTTAGGCGGT TACCTTTAAC
PrP(129-110) Reverse	TCGAGTTAAA GGTAACCGCC TAACCCCCCT ACCACTGCTC CTGCTGCCGC GGCACCTGCG ACGTGTTTAC CCTGGAAATA GAGGTTCTCG
PrP(110-119) Forward 5' end	GATCCGAGAA CCTCTATTTC CAGGGTAAAC ACATGGCAGG TGCCGCGGCA GCAGG
PrP(119-136) Forward 3' end	AGCAGTGGTA GGGGGGTTAG GCGGTTACAT GTTAGGCAGC GCAATGTCGC GTTAAC
PrP(136-126) Reverse 5' end	TAACCCCCCT ACCACTGCTC CTGCTGCCGC GGCACCTGCC ATGTGTTTAC CCTGGAAATA GAGGTTCTCG
PrP(126-110) Reverse 3' end	TCGAGTTAAC GCGACATTGC GCTGCCTAAC ATGTAACCGC C

Figure 2: A schematic of the prion peptide expression constructs

The ten-histidine tag (red), glutathione-S-transferase (yellow), the linker containing the TEV protease cleavage site (green) and the PrP peptide (PrP(110-129) or PrP(110-136)) (blue) are shown.

M9 Minimal Media consists of a salt solution which is sterilized by autoclaving (6g/L Na_2HPO_4 anhydrous, 3g/L KH_2PO_4 anhydrous, 0.5g/L NaCl, 1g/L NH_4Cl (or $^{15}\text{NH}_4\text{Cl}$), adjusted to pH 7.3 with 10 N NaOH (approximately 0.8 mL/L media)), and a filter sterilized supplement, added just before use (15 mL/L 20% glucose (or $^{13}\text{C}_6$ -glucose), 2 mL/L 1M MgSO_4 , 1 mL/L 0.1 M CaCl_2 , 10 mL/L 100X MEM Vitamins (Invitrogen), 100 mg/L Ampicillin). The supplement was prepared and stored at 4 °C, protected from light until just before use. Uniform isotope labeling with nitrogen-15 (^{15}N) and carbon-13 (^{13}C) was accomplished by using ^{15}N -ammonium chloride and glucose with $^{13}\text{C}_6$ -glucose, respectively.

Purification

Frozen cells were thawed in lysis buffer (10 mM Tris, 100 mM Phosphate, 5 mM beta-mercaptoethanol, 5 mM MgCl_2 , pH 8.0, 1 tablet/50mL Complete™ Protease Inhibitors(Roche), 20 mg/mL lysozyme and 50 µg/mL DNase) on ice. The cells were disrupted by sonication on ice with a 400W Branson Sonicator at 50% duty cycle, 30% amplitude, 10 minutes sonication time. After sonication, cell debris was removed by centrifugation at 50k x g at 8°C for 1 hour. Lysis and centrifugation was repeated once. Successive supernatants were retained and pooled.

The pooled supernatant was incubated at room temperature for approximately 1 hour with Qiagen SuperFlow Ni-NTA resin with mild agitation. The resin was then poured into a column and was washed with buffer to remove unbound impurities, and

with 60 mM imidazole to remove weakly bound impurities. The target protein was eluted with 250 mM imidazole.

The purification was monitored by SDS-PAGE and fractions containing the fusion protein were pooled and concentrated by ultrafiltration to approximately 50 mL total volume, at which point the protein was quantified by absorbance at 280 nm, using molecular mass of the fusion protein and the theoretical molar extinction coefficient calculated by the ProtParam utility(33) for PrP(110-129) (31 704 g/mol and $46\,090\text{ M}^{-1}\cdot\text{cm}^{-1}$, respectively) and (PrP(110-136) (32 403 g/mol and $45\,840\text{ M}^{-1}\cdot\text{cm}^{-1}$, respectively). The typical yield was 7 mg fusion protein per L of M9 media. The molecular masses were adjusted to account for uniform ^{15}N or ^{13}C labeling when appropriate.

After dialysis or buffer exchange into the TEV protease cleavage buffer (50 mM Tris-HCl (pH 8.0), 0.5 mM EDTA, 1 mM DTT), the fusion protein was incubated at 4°C overnight with 25 U AcTEV (Invitrogen)/mg fusion protein to release the peptide of interest.

HPLC Purification

The cleaved peptide was purified by High Performance Liquid Chromatography (HPLC) on a semi-preparative C4 Jupiter (Phenomenex) column with a linear gradient from 95% water, 5% Acetonitrile and 0.06% trifluoroacetic acid to 95% acetonitrile, 5% water and 0.06% trifluoroacetic acid over 15 minutes at a flow rate of 4 mL/min. The

recovered peptide was lyophilized to dryness and stored at -20°C until needed for NMR experiments.

The identity of each peptide was initially confirmed by positive ion, electrospray and/or matrix-assisted laser desorption ionization, time-of-flight mass spectrometry by Ms. Lisa Waldron. Subsequent identity and purity were evaluated by HPLC and NMR. The peptides were quantified by measuring the absorbance at 280 nm and calculating the concentration with the molecular mass and theoretical extinction coefficient calculated by the ProtParam utility(33) for PrP(110-129) (1 846.1 g/mol and 1 490 M⁻¹·cm⁻¹, respectively) and (PrP(110-136) (2 548.9 g/mol and 1 490 M⁻¹·cm⁻¹, respectively). The molecular masses were adjusted to account for uniform ¹⁵N or ¹³C labeling when appropriate.

NMR Sample Preparation in Bicelles

The samples prepared in 1,2-dicaproyl-*sn*-glycerol-3-phosphocholine (DHPC) and 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) bicelles were prepared by dissolving the lyophilized peptide in pure H₂O, removing a small aliquot for quantification by UV as described above, followed by the addition of solid DMPC, vortexing to homogeneity, addition of 25% w/w aqueous DHPC, vortexing until dissolved, addition of 1 M HEPES pH 7.5 and finally the addition of D₂O to arrive at a 500 µL sample containing 2 mg peptide, 15% w/w total lipid with a molar ratio of 1

DMPC/2 DHPC, 5% v/v D₂O, 25 mM HEPES, and approximately 1:100 peptide to DMPC molar ratio.

NMR Sample Preparation in Micelles

The NMR samples prepared with dodecylphosphocholine (DPC) were prepared by dissolving the lyophilized peptide in pure H₂O, removing a small aliquot for quantification by UV as described above, followed by addition of D₂O to 5% v/v, sodium phosphate buffer to 10 mM, a two hundred-fold molar excess of DPC with regard to peptide, and pH adjusted to 7.6 with 10 N NaOH_(aq). The sample of PrP(111-136) used for structural calculations contained approximately 0.5 mg peptide and 14 mg DPC (approximately 200-fold molar excess, 90 mM) in 450 μ L total volume with 10 mM sodium phosphate at pH 7.6, containing 3.1% w/w lipid.

NMR samples prepared with sodium dodecyl sulfate (SDS), n-octyl glucoside, and lysomyristoylphosphatidylglycerol (LMPG) were prepared as described above for DPC, adjusting the amount of lipid added to maintain an approximately 100-fold molar excess of phospholipid with regard to the amount of peptide used.

NMR Experiments

All spectra were recorded at 310 K, on a Bruker Avance 700 MHz equipped with cryogenic probehead or a Bruker Avance III 600 MHz equipped with a cryogenic probehead. The spectra recorded were ¹H-¹⁵N Heteronuclear Single Quantum Coherence (HSQC), ¹⁵N-edited NOESY, ¹⁵N-edited TOCSY, 3D HNCO, 3D HN(CA)CO, 3D

CBCA(CO)NH, and 3D HNCACB. All experiments used a relaxation delay of 1 second between transients.

HSQC

The ^1H - ^{15}N HSQC experiments were recorded at 700 MHz or 600 MHz with the standard pulse sequence using a water flip-back pulse, gradient coherence selection and echo-antiecho detection in the indirection dimension (34). A total of 1024 complex points per FID were collected in the direct dimension (F2) with a spectral window of 14 ppm. The indirect dimension (F1) had 64 complex points with a spectral window of 32 ppm for ^{15}N . Eight transients per data point were collected.

The HSQC data were processed with nmrPipe (35). After deconvolution of the water signal, each FID was multiplied by a 30° shifted sine square followed by zero-filling and Fourier transformation into the frequency domain. The indirect dimension was linear predicted to double the number of datapoints prior to multiplication of a cosine window function, followed by zero-filling and Fourier transformation. Data were visualized with NMRView software(36).

^{15}N -edited NOESY

The ^{15}N -edited NOESY experiment was recorded at 700 MHz with a standard pulse sequence (37). The parameters used were 2048 complex points in the direct dimension, 40 complex points in the ^{15}N dimension, and 120 complex points in the

indirect ^1H dimension. The carrier and spectral windows for the indirect dimensions were 121.5 ppm and 17.0 ppm in the ^{15}N dimension and 4.658 ppm and 11.00 ppm in the ^1H dimension, respectively. The ^{15}N -edited NOESY was recorded twice, with mixing times of 50 ms and 150 ms. A total of 16 transients per data point were recorded and the total acquisition time was 55 hours for each experiment.

The data were subjected to a 63 degree shifted sine square bell window function, zero-filling, Fourier transform into the frequency domain, phase correction, extraction of the region of interest (7.0 ppm to 10.0 ppm), and a polynomial baseline correction. Processing in the ^1H indirect dimension included a 30 degree shifted sine bell window function, followed by zero-filling, Fourier transform into the frequency dimension and a polynomial baseline correction. Processing in the ^{15}N direction involved forward and backward linear prediction, a 30 degree shifted sine bell window function, followed by zero-filling, Fourier transform into the frequency dimension.

^{15}N -edited TOCSY

The ^{15}N -edited TOCSY experiment was recorded with a standard pulse sequence (37) at 700 MHz using the same spectral parameters used for the NOESY described above. The ^{15}N -edited TOCSY was recorded with a mixing time of 60 ms. Eight transients per data point were collected and the total acquisition time was 28 hours. The ^{15}N -edited TOCSY data were processed in the same way as the ^{15}N -edited NOESY.

3D HNCO and HN(CA)CO

The 3D HNCO and 3D HN(CA)CO experiments were recorded at 700 MHz with standard pulse sequences(31), using the same spectral parameters used for NOESY and TOCSY experiments in the direct and indirect nitrogen dimensions and with 256 complex points and a carrier and spectral window of 176 ppm and 10 ppm in the indirect carbon dimension. A total of 16 transients per data point were collected. The data for these experiments were processed in the same was as the NOESY and TOCSY experiments.

3D HNCACB and CBCA(CO)NH

The 3D HNCACB and 3D CBCA(CO)NH experiments were recorded at 600 MHz with standard pulse sequences(31), using the same spectral parameters used for NOESY and TOCSY experiments in the direct dimension, with 80 (HNCACB) or 40 (CBCA(CO)NH) complex points in the indirect nitrogen dimension and 128 complex points and a spectral window of 60 ppm in the indirect carbon dimension. A total of 16 transients per data point were collected. The data for these experiments were processed in the same was as the NOESY and TOCSY experiments.

Structure Calculation

Data analysis and resonance assignment were performed using NMRview(36) and CCPNMR Analysis(38) software packages. Spectra, chemical shift lists, and partially assigned peak lists were converted into a suitable format with the CCPNMR Format Converter(38). Structures were calculated using ARIA 2.1(39). Constraints used were the

partially assigned NOESY peak list, and a hydrogen bond table. The default parameters of ARIA were used, including considering all existing assignments as ambiguous and tolerance values of 0.02 ppm in the direct proton dimension, 0.04 ppm in the indirect proton dimension, and 0.5 ppm in the nitrogen dimension. An ensemble of three hundred structures was calculated, with the twenty lowest energy structures selected for water refinement. The structure set generated by ARIA was inspected graphically and images prepared with the software package Chimera(40).

Results and Discussion

Expression and Purification of Prion Peptides

The expression and purification system described here was developed to produce PrP(110-129), however it was also used to produce PrP(110-136) and PrP(111-136) when PrP(110-129) did not demonstrate similar properties to those reported by the Vold group for PrP(110-136). While the initial goal of this project was to determine the structure of PrP(110-136) in DMPC/DHPC micelles as studied by Vold and coworkers, for reasons discussed later the peptide that was produced and studied was in fact PrP(111-136). What follows is a description and discussion of the expression and purification system used to produce PrP(110-129), PrP(110-136) and PrP(111-136).

The peptides studied here were expressed as fusion proteins with Glutathione-S-transferase (GST). The use of an N-terminal fusion partner is required to protect the unfolded peptide from proteolytic degradation inside the bacteria (41). The expression vectors containing the sequence for PrP(110-129) or PrP(110-136) are based on the pET19b plasmid (Novagen). This plasmid contains a multiple cloning site downstream from the promoter sequence for T7 polymerase. Additionally, the sequence was designed to contain an N-terminal ten-histidine tag, glutathione-S-transferase (GST), and a short linker sequence containing the recognition sequence of the tobacco etch virus (TEV) protease. When transformed into *E. coli* BL21(DE3), this plasmid expresses high levels of the protein of interest (see Figure 3). The expressed fusion protein was then purified using nickel affinity chromatography. The peptide was released from the purified fusion

protein by cleavage with TEV, and finally purified by HPLC. The TEV cleavage was designed to release the peptide of interest with an additional N-terminal glycine, which is not expected to have a significant structural effect. The amino acid sequence of the expected PrP(110-136) peptide is GKHMAGAAAAGAVVGGLGGYMLGSAMSR, while the sequence of PrP(110-129) ends after the second methionine and PrP(111-136) does not contain the N-terminal GK residues.

An HPLC chromatogram of the TEV cleavage mixture is shown in Figure 4. The identity of each peak was initially confirmed by positive ion, electrospray ionization, time of flight mass spectroscopy. Analysis revealed that a small amount of a shorter peptide, PrP(111-136), was also produced. The TEV reaction conditions were modified to minimize the production of this byproduct by changing the reaction temperature from 30°C to 4°C and the reaction time from 2 hours to 15 hours. Upon verification by MS however, the ^{15}N -labeled samples used for structure determination were confirmed to be PrP(111-136), while the ^{15}N , ^{13}C -labeled samples produced later were confirmed to be PrP(110-136) (see Table 2). It is suspected that the production of this byproduct was the result of lot to lot variability in the commercially sourced TEV protease. The precise cause of the variability is currently being pursued by the Aubin lab with the vendor.

Figure 3: SDS-PAGE Gels showing expression and purification of PrP(110-136)

A: A 20% homogeneous SDS-PAGE gel showing the results of cell lysis. The lanes from left to right: Mark 12 molecular mass marker (Invitrogen) (M), Supernatant 1, Pellet 1, Supernatant 2, and Pellet 2.

B: A 20% homogeneous SDS-PAGE gel showing the results of nickel affinity chromatography. The lanes from left to right: Mark 12 Molecular marker (Invitrogen) (M), Flowthrough (FT), Wash (W), Elution fractions corresponding to the beginning of elution (E1), the elution peak maximum (E2), and trailing fractions (E3, E4, and E5).

Table 2: Expected and observed molecular masses as confirmed by mass spectrometry

Expected mass of PrP(110-136)	¹⁵ N-labeled samples	¹⁵ N, ¹³ C-labeled samples
Unlabelled: 2548.9 Da		
¹⁵ N-labeled: 2582.9 Da		
¹⁵ N, ¹³ C-labeled: 2689.9 Da		2692 Da
Expected mass of PrP(111-136)		
Unlabelled: 2363.7 Da		
¹⁵ N-labeled: 2394.7 Da	2397 Da	
¹⁵ N, ¹³ C-labeled: 2493.7 Da		

DMPC/DHPC Bicelles and DPC Micelles

When prepared in DPC micelles or SDS micelles, PrP(110-129) samples did not show drastic shifts in the glycine amide resonances comparable to those reported by the Vold group for PrP(110-136) and suggestive of secondary structure (see Figure 5). Since it was possible that PrP(110-129) was too short to interact with lipid micelles, we chose to study a longer peptide, PrP(110-136). The precise nature of the interactions between PrP(110-129) and phospholipid micelles are being further investigated by the Aubin group and are outside the scope of the work presented here.

Figure 4: An HPLC chromatogram of the TEV cleavage reaction.

The inset shows PrP(110-136), PrP(111-136), TEV, and the mixture of uncleaved GST-PrP(110-136) and cleaved GST.

Figure 5: HSQC of PrP(110-129) in water and DPC micelles

Figure 6: Comparison of glycine amide resonances

The glycine amide resonances of PrP(110-136) in DMPC/DHPC bicelles at pH 7.5 as reported by Glover *et al.* (23) (blue), compared with the observed glycine resonances of PrP(111-136) in DPC at pH 7.6 (green) and PrP(110-136) in DPC at pH 7.6 (red).

The experimental conditions used by the Vold group were reproduced with PrP(111-136) and the buffer fixed at pH 7.5 and the resulting HSQC is shown in Figure 7A. While a spread of glycine resonances similar to those reported by the Vold group (see Figure 6) is discernable in the nitrogen range of 103 ppm to 111 ppm, the quality of the spectrum is not optimal. Because they were able to observe glycine resonances in a relatively narrow nitrogen spectral window, the Vold group was able to acquire more indirect points over a narrower spectral window, producing similar digital resolution in a shorter time than possible with a uniformly ^{15}N labeled peptide.

It became apparent that a short acquisition time was essential when it was discovered that PrP(111-136) in DMPC/DHPC bicelles was not stable. The quality of HSQC spectra of PrP(111-136) prepared this way decay rapidly, becoming indistinguishable from noise less than an hour after preparation. This system appears to reproduce the conformational behaviour reported by Vold and coworkers, but is not suitable for structural studies because the sample is not stable for long enough to record three-dimensional NMR experiments.

Figure 7: HSQC spectra of PrP(111-136) in DMPC/DHPC bicelles and DPC micelles.

A: PrP(111-136) in 1:2 DMPC/DHPC micelles at approximately pH 7.0, demonstrating similar glycine resonances to those reported by the Vold group(23). B: PrP(111-136) in 90 mM DPC micelles at pH 7.6.

It was necessary to find a phospholipid system that reproduced the structural behaviour observed in DMPC/DHPC bicelles but also allowed all or most amide resonances to be resolved and was sufficiently stable to allow three-dimensional NMR experiments to be recorded. When prepared in DPC micelles, a similar dispersion of glycine amide proton chemical shifts was observed for PrP(111-136) (see Figure 7B), and additionally the signal to noise ratio of the HSQC spectra is greatly improved, most of the amide resonances were resolved, and the samples were sufficiently stable to allow structural NMR experiments to be recorded.

Additional detergents were investigated, namely sodium dodecyl sulfate (SDS), n-octyl glucoside, and lysomyristoylphosphatidylglycerol (LMPG). The PrP(110-136) peptide shows interactions with SDS and LMPG micelles, the latter showing larger chemical shift dispersions, as shown in Figure 8, while no interaction is apparent with n-octyl glucoside (data not shown). None of the chemical shift perturbations observed in these detergents are as large as those observed in DPC. The fact that PrP(110-129) did not interact in the same way suggests that residues 130-136 are important for the observed interaction.

Figure 8: HSQC spectra of PrP(110-136) in LMPG and SDS at varying pH

We found that PrP(111-136) prepared in DPC shows a similar dependence upon pH as reported by the Vold group for PrP(110-136) in DMPC/DHPC, with evidence of conformational heterogeneity at lower pH. All samples prepared in DPC micelles were pH adjusted and verified to be at pH 7.6 unless otherwise specified. Figure 9 shows the HSQC spectra of PrP(111-136) in DPC micelles at different pH. At acidic pH, more than the 25 expected resonances are present in the HSQC spectrum, indicating that the peptide exists in multiple conformations. At pH 7.6, the HSQC spectrum shows only 25 resonances, with any increase in pH causing the disappearance of resonances due to rapid exchange of the amide protons with water. These resonances are also present at pH 6.3 and pH 7.0. In light of these findings with PrP(111-136) in DPC, it is likely that the pH of the sample shown in Figure 3A is lower than pH 7.5. The pH behaviour of PrP(111-136) in DMPC/DHPC was not explored further in order to focus on structural studies of PrP(111-136) in DPC. The pH behaviour observed with PrP(111-136) in DPC micelles is similar to the findings of the Vold group.

Figure 9: The HSQC of PrP(111-136) in DPC micelles at different pH.

Resonance Assignment

The Vold group reported chemical shifts for the glycine amide pairs of PrP(110-136) prepared in DMPC/DHPC at pH 7.5(23). They were able to unambiguously assign each glycine amide pair by incorporating a unique $^{14}\text{N}/^{15}\text{N}$ isotope ratio into each glycine residue and comparing the HSQC signal intensities to those of an identical peptide with an $^{14}\text{N}/^{15}\text{N}$ isotope ratio of 1 in each glycine. The values reported in the literature (see Figure 6) are not identical to the chemical shifts reported here, but the general pattern of the chemical shifts is sufficient to unambiguously assign the amide pairs of 6 of the 7 glycine residues, allowing the sole remaining glycine to be assigned as glycine 114. The amide proton resonance of glycine 114 in PrP(111-136) in DPC is significantly different than the value reported for PrP(110-136) in DMPC/DHPC, which is likely due to the fact that our peptide was shorter than the Vold peptide by 1 residue at the N-terminal.

From the glycine amide assignments, adjacent amide protons were assigned based on NOEs between amide protons, in a manner illustrated in Figure 10. The amide resonances of many residues were more challenging to assign, due to the absence of data in the TOCSY. Therefore it was necessary to use the NOESY spectrum to assign the remaining resonances. However this approach suffers from the drawback that NOEs can occur between the amide proton of interest and amide protons with similar or degenerate chemical shifts to the adjacent residues.

Figure 10: A partial strip plot of the ^{15}N -edited NOESY experiment, demonstrating sequential assignment through NOEs to adjacent amides

The experiment recorded with PrP(111-136) suffered from a lack of sensitivity however, due to a low peptide concentration in the sample, and did not detect all side-chain resonances, the most notable absence being the glycine alpha protons. The assignments used for structure calculations of PrP(111-136) are shown in Table 3. Notably, histidine 111 and methionine 112 were not observed. The presence of NOE crosspeaks between the two amide resonances assigned to R136 suggests that it is in exchange between two conformations. The assignment of arginine 136 used for the first round of structure calculations was the most intense of the two peaks. It is expected that histidine 111 will not be detected because terminal amines are protonated at pH 7.6 and likely to be in rapid exchange with water. Additionally, the sidechain of arginine 136 is not expected to be observed by HSQC due to rapid exchange of the protons with water.

The risk of using the NOESY spectrum as the primary source for assignment is that strong NOEs from protons that are near amide protons in space, but not in the same spin system can complicate assignment. When adjacent residues are of the same type and experiencing similar magnetic environments, chemical shifts are also highly similar. In this case, the alpha protons of the alanines and glycines were highly degenerate, due to their close proximity to each other, the highly homogeneous environment provided by the lipid, and thus in some cases the assignments could only be considered tentative.

Table 3: Tentative Chemical Shift Assignments of PrP(111-136) used in Structure Calculations

Residue	Amide Resonances (¹⁵ N, ¹ H, ppm)	Alpha Resonance(s) (ppm)	Other (ppm)
Ala113*	118.6, 8.264	3.891	1.433 (β)
Gly114	105.5, 8.809	3.644, 3.519	
Ala115*	121.2, 9.192	4.021	1.435 (β)
Ala116*	123.0, 8.080	4.027	1.605 (β)
Ala117*	119.5, 8.270	3.889	1.443 (β)
Ala118	119.1, 8.166	4.259	1.356 (β)
Gly119	103.2, 8.423	3.525	
Ala120	123.3, 8.404	4.175	1.312 (β)
Val121	117.5, 7.948	3.544	2.165 (β)
Val122	117.6, 7.696	3.524	2.031 (β)
Gly123	105.8, 9.134	4.027	
Gly124	111.1, 8.641	4.031, 3.486	
Leu125	122.0, 8.151	4.108	1.843 (β)
Gly126	105.7, 8.427	3.766, 3.554	
Gly127	109.5, 9.721	4.422, 3.796	
Tyr128	123.1, 8.702	4.109	3.151 (β), 2.994 (β), 8.702 (2,6 H), 6.841(3,5H)
Met129	119.6, 8.370	4.026	2.073 (β)
Leu130	119.8, 8.347	4.109	1.872 (β)
Gly131	106.6, 9.050	3.301, 3.215	
Ser132	114.6, 7.915	3.989	3.463 (β)
Ala133	123.2, 7.619	4.179	1.502 (β)
Met134	115.1, 7.925	4.444	2.116 (β)
Ser135	115.2, 7.689	4.450	3.904 (β)
Arg136	127.1, 7.587	4.156	1.734, 1.623 (β)
Marked (*) assignments were tentative.			

Structure Calculation

It was initially expected that the backbone assignments and the ^{15}N -edited NOESY peak list would be sufficient for ARIA to generate a high-resolution ensemble of structures for PrP(111-136) in DPC. When run in this way however, the structures calculated failed to converge. It is likely that this is due to the highly degenerate chemical shifts of the amide and alpha protons, and assignment errors among alanines 113, 115, 116 and 117.

A repetitive pattern of NOEs between alpha protons ($\text{H}\alpha_i$) and amide protons three (HN_{i+3}) and four residues (HN_{i+4}) downstream emerged upon further analysis of the data. The Vold group reported that PrP(110-136) in DMPC-DHPC bicelles at pH 7.5 has helical character by CD. Based on this finding and the observed pattern of NOEs, the presence of a hydrogen bond between the carbonyl oxygen (CO_i) and the amide proton four residues downstream (HN_{i+4}) was inferred when both NOEs ($d_{\alpha\text{N}}(i,i+3)$ and $d_{\alpha\text{N}}(i,i+4)$) (42) were observed (See Figure 11). These inferred hydrogen bonds were used as additional constraints for structure calculation.

HuPrP 110-136

Figure 11: Schematic of expected and observed NOEs in an alpha helix.

A representation of an ideal alpha helix (top), demonstrating the network of NOEs that are possible between amide (attached to blue nitrogens) and alpha protons (attached to grey carbons). The dashed lines denote expected NOEs between the amide proton at left and alpha protons three and four residues downstream in green and orange, respectively. At bottom, a map of the observed NOEs (based on tentative assignment) between amide and alpha protons in huPrP(111-136). Hydrogen bonds between an amide at position i and the carbonyl oxygen at position $i-4$ when NOEs were observed between the amide at position i and the alpha carbons at positions $i-3$ and $i-4$.

It was recognized that the use of hydrogen bond constraints may over-constrain the molecular system during structure calculations and potentially bias the outcome. The degeneracy of the chemical shifts however, under-constrained the system, as evidenced by the failure of the structures to converge when only constrained by unassigned NOEs. Forcing ARIA to use manual assignments in light of the degeneracy would have introduced bias as well.

The resulting ensemble of structures (Figure 12) shows a high degree of heterogeneity in the N- and C-Terminal regions from residues 111-115 due to the lack of chemical shift assignments for residues 111 and 112, and raised the possibility of assignment errors for alanines 113, 115, 116 and 117, and due to an insufficient number of constraints. The ensemble shown in Figure 12 includes all structures selected for refinement which do not have NOE violations $> 0.5 \text{ \AA}$.

Figure 12: Superimposed backbone structures calculated for PrP(111-136).

The 18 lowest energy structures without NOE violations $> 0.5 \text{ \AA}$ are shown, oriented with the N-termini at the top.

The statistics of the structure calculation are shown in Table 4. Compared to the average structure, the root mean square deviation (RMSD) of $0.84 \text{ \AA} \pm 0.62 \text{ \AA}$ indicates that this ensemble of structures is preliminary and suggests that the available NOEs that can be unambiguously assigned are not sufficient. The four highest energy structures in the set contain all four of the distance violations $>0.5 \text{ \AA}$, as well as visually obvious distortions in the central helical region. Upon closer examination, only the two highest energy structures have NOE violations $> 0.5 \text{ \AA}$ and the most frequent NOE violations in these high-energy structures are between amide protons and side-chain beta protons (123 $\rightarrow$ 120, 122 $\rightarrow$ 121, 134 $\rightarrow$ 133, 123 $\rightarrow$ 122, 132 $\rightarrow$ 135, and 122 $\rightarrow$ 122), indicating that there may be assignment errors in these beta protons. With the two highest energy structures removed, the most frequent NOE violations are between HN116 and HN113 and HN116 and HA113, which raised the possibility of assignment errors among alanines 113, 115, 116, and 117.

The overall tendency towards an alpha helical structure however indicates that the peptide does adopt a secondary structure in DPC micelles that is very different and more ordered than its conformation in water. This ensemble is biased by the use of hydrogen bond constraints, which most likely over-constrained the system. The RMSD of the generated structure set indicates that this is a preliminary structure, and that more data is required to confirm the assignments and increase the number of distance constraints. In order to generate a high-resolution, unbiased structure set, the molecular system will have to be better constrained, using unbiased constraints, directly derived from experimental observations.

Table 4: Statistics from structure calculation of HuPrP(111-136)

Total number of constraints	194
RMSD from experimental distance restraints	
127 Unambiguous NOEs (Å)	0.042 (+/- 0.035)
49 Ambiguous NOEs (Å)	0.032 (+/- 0.040)
18 H Bonds (Å)	0.036 (+/- 0.034)
RMSD from average structure (Å), backbone atoms	0.84 (+/- 0.62)
#NOE violations in 20 lowest energy structures > 0.5 Å	4

^{13}C , ^{15}N -Labeled PrP(110-136)

The initial ensemble of structures presented above indicates that the resonances of PrP(111-136) are too degenerate to calculate an unbiased structure with ^{15}N -edited data alone, and that the tentative assignments in Table 3 may not be correct for a few residues, particularly the highly degenerate alanines at the N-terminus. In order to obtain more reliable assignments and to extract additional experimental constraints, a doubly labeled ^{13}C , ^{15}N sample of PrP(110-136) was produced.

The same procedure used to produce ^{15}N -labeled PrP(111-136) was followed to produce ^{13}C , ^{15}N -labeled PrP(110-136), with the exception that the glucose used in the M9 minimal media was $^{13}\text{C}_6$ -glucose. The HSQC spectrum of PrP(110-136) in DPC at pH 7.6 is shown in Figure 13. The drastic differences in the amide proton resonances of residues alanine 113 (Figure 13 assignment) and glycine 114 between Figure 7 and

Figure 13 made it clear that these were different peptides, and mass spectrometry confirmed that the peptide sample used for structure determination above is PrP(111-136) and the subsequently produced doubly labeled sample shown in Figure 13 is PrP(110-136).

In addition to the differences in chemical shifts between PrP(110-136) and PrP(111-136) shown in Figures 7 and 13, PrP(110-136) samples prepared in DPC micelles appear to be more stable. Samples of PrP(111-136) in DPC were stable enough for the collection of three-dimensional NMR experiments, but the signal-to-noise ratio of the NMR spectra eventually declined after several weeks. Samples of PrP(110-136) prepared in DPC, by comparison, appear to be highly stable, showing comparable HSQC signal to noise after several months.

Figure 13: HSQC of PrP(110-136) in DPC micelles at pH 7.6.

This spectrum was recorded with a ^{13}C , ^{15}N doubly labeled sample.

The collection of ^1H , ^{13}C , ^{15}N triple resonance experiments has provided a mechanism to resolve assignment errors by correlating inter- and intra-residue experiments. The corrected amide assignments are shown in Figure 13. Additionally, the amide, alpha, and beta proton resonances, combined with the alpha, beta and carbonyl carbon resonances can be used to predict the phi and psi torsion angles based upon a database of high-resolution structures using the TALOS software package (43). The addition of ^{13}C -edited proton-proton distances will allow better definition not only of the carbon backbone, but also of the side chains.

A peptide with only ^{15}N -labelling was prepared initially because it provides higher resolution ^{15}N -edited spectra than ^{13}C , ^{15}N doubly labeled samples, since there is no need for ^{13}C decoupling. The structure calculation using ^{15}N data only and using hydrogen bond constraints was useful in helping to identify the presence of assignment errors. The molecular system of prion peptides in lipid micelles was more complex than originally anticipated and merits further study with more robust techniques than would ordinarily be considered for studying a small peptide by NMR.

Conclusions and Future Work

This study found that PrP(111-136) and PrP(110-136) interact with DPC micelles and that the nature of the interaction is pH-dependant. similar to the interactions reported in the literature between PrP(110-136) and DMPC/DHPC bicelles. This particular interaction is characterized by an extreme dispersion of the glycine amide proton resonances, which does not occur with PrP(110-129) in DPC or with PrP(110-136) in SDS, n-octyl glucoside, or LPMG.

A preliminary ensemble of structures was calculated for PrP(111-136) in DPC micelles at pH 7.6, suggesting that the conformation of the peptide is helical, but requiring more data to generate an unbiased, high-resolution structure. The NMR data from ^{15}N -edited experiments was highly degenerate, and as a result, additional constraints will be required, which will be available from ^{13}C , ^{15}N doubly labeled samples. The addition of torsion angle constraints, and ^{13}C -edited proton-proton distance constraints should allow for the calculation of a high-resolution ensemble of structures.

The striking difference in the structural behaviour between PrP(111-136) and PrP(110-129) was unexpected and should be better characterized. It appears that these differences are due solely on the presence of seven predominantly polar amino acids, including the charged arginine 136. It was unexpected that these amino acids would make a predominantly hydrophobic peptide more likely to adopt secondary structure in a lipid environment. Equally interesting is the fact that residues 128-131 have been shown by the

Wuthrich group(14) to participate in a beta-sheet in the solution structure of the globular domain. It cannot be concluded, however, that PrP(110-129) does not interact with DMPC/DHPC bicelles or DPC micelles, only that the interaction, if it exists, does not induce the same type of radical conformational shift seen in PrP(111-136) and PrP(110-136).

Future work to better characterize the interactions of PrP(110-136) and PrP(110-129) with DPC micelles should be undertaken. Studies with paramagnetic relaxation agents are currently underway to define the surfaces of interaction in both peptides, which will help determine whether PrP(110-136) (and potentially PrP(111-136)) are completely buried in the micelle and their orientation within it, and whether PrP(110-129) interacts with the surface of the micelle.

Structural studies with PrP(110-136) in DPC which include the mutations studied by the Lingappa group would be very interesting. Identifying differences between the way wildtype PrP(110-136) and mutant PrP(110-136) interact with membrane may provide insight into the predisposition of hydrophobic mutants to become embedded in the membrane, and ultimately how these transmembrane configurations of the prion protein contribute to the neurological symptoms reported by Lingappa and coworkers.

Bibliography

1. Siedler H, Malamud N. Creutzfeldt-Jakob's disease. Clinicopathologic report of 15 cases and review of the literature (with special reference to a related disorder designated as subacute spongiform encephalopathy). *J Neuropathol Exp Neurol*. 1963 Jul;22:381-402.
2. Kovács GG, Budka H. Prion diseases: from protein to cell pathology. *Am J Pathol*. 2008 Mar;172(3):555–565.
3. Prusiner SB. Prion diseases and the BSE crisis. *Science*. 1997 Oct 10;278(5336):245-251.
4. Ladogana A, Puopolo M, Croes EA, Budka H, Jarius C, Collins S, Klug GM, Sutcliffe T, Giulivi A, Alperovitch A, Delasnerie-Laupretre N, Brandel J, Poser S, Kretzschmar H, Rietveld I, Mitrova E, Cuesta JDP, Martinez-Martin P, Glatzel M, Aguzzi A, Knight R, Ward H, Pocchiari M, van Duijn CM, Will RG, Zerr I. Mortality from Creutzfeldt-Jakob disease and related disorders in Europe, Australia, and Canada. *Neurology*. 2005 May 10;64(9):1586-91.
5. Kovács GG, Trabattoni G, Hainfellner JA, Ironside JW, Knight RSG, Budka H. Mutations of the prion protein gene phenotypic spectrum. *J Neurol*. 2002 Nov ;249(11):1567-82.
6. Schneider K, Fangerau H, Michaelsen B, Raab WH. The early history of the transmissible spongiform encephalopathies exemplified by scrapie. *Brain Res Bull*. 2008 Dec 16;77(6):343-55.
7. Miller MW, Williams ES, Hobbs NT, Wolfe LL. Environmental sources of prion transmission in mule deer. *Emerging Infectious Diseases*. 2004;10(6):1003-1006.
8. Prusiner SB. Prions: novel infectious pathogens. *Adv Virus Res*. 1984;29:1-56.
9. Hegde RS, Mastrianni JA, Scott MR, DeFea KA, Tremblay P, Torchia M, DeArmond SJ, Prusiner SB, Lingappa VR. A transmembrane form of the prion protein in neurodegenerative disease. *Science*. 1998 Feb 6;279(5352):827-834.
10. Lysek DA, Schorn C, Nivon LG, Esteve-Moya V, Christen B, Calzolari L, von Schroetter C, Fiorito F, Herrmann T, Güntert P, Wüthrich K. Prion protein NMR structures of cats, dogs, pigs, and sheep. *Proc Natl Acad Sci U S A*. 2005 Jan 18;102(3):640-5.
11. Calzolari L, Lysek DA, Pérez DR, Güntert P, Wüthrich K. Prion protein NMR structures of chickens, turtles, and frogs. *Proc Natl Acad Sci U S A*. 2005 Jan 18;102(3):651-5.

12. Gossert AD, Bonjour S, Lysek DA, Fiorito F, Wüthrich K. Prion protein NMR structures of elk and of mouse/elk hybrids. *Proc Natl Acad Sci U S A*. 2005 Jan 18;102(3):646-50.
13. López Garcia F, Zahn R, Riek R, Wüthrich K. NMR structure of the bovine prion protein. *Proc Natl Acad Sci U S A*. 2000 Jul 18;97(15):8334-9.
14. Zahn R, Liu A, Lühns T, Riek R, von Schroetter C, López García F, Billeter M, Calzolari L, Wider G, Wüthrich K. NMR solution structure of the human prion protein. *Proc Natl Acad Sci U S A*. 2000 Jan 4;97(1):145-50.
15. Riek R, Wider G, Billeter M, Hornemann S, Glockshuber R, Wüthrich K. Prion protein NMR structure and familial human spongiform encephalopathies. *Proc Natl Acad Sci U S A*. 1998 Sep 29;95(20):11667–11672.
16. Baldwin MA, Pan KM, Nguyen J, Huang Z, Groth D, Serban A, Gasset M, Mehlhorn I, Fletterick RJ, Cohen FE. Spectroscopic characterization of conformational differences between PrPC and PrPSc: an alpha-helix to beta-sheet transition. *Philos Trans R Soc Lond B Biol Sci*. 1994 Mar 29;343(1306):435-41.
17. Smith CJ, Drake AF, Banfield BA, Bloomberg GB, Palmer MS, Clarke AR, Collinge J. Conformational properties of the prion octa-repeat and hydrophobic sequences. *FEBS Lett*. 1997 Apr 1;405(3):378-84.
18. Hay B, Barry RA, Lieberburg I, Prusiner SB, Lingappa VR. Biogenesis and transmembrane orientation of the cellular isoform of the scrapie prion protein [published erratum appears in *Mol Cell Biol* 1987 May;7(5):2035]. *Mol Cell Biol*. 1987 Feb;7(2):914-20.
19. Baleja JD. Structure determination of membrane-associated proteins from nuclear magnetic resonance data. *Analytical Biochemistry*. 2001 Jan 1;288(1):1-15.
20. Poget SF, Girvin ME. Solution NMR of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better. *Biochim Biophys Acta*. 2007 Dec;1768(12):3098-106.
21. Satheeshkumar KS, Jayakumar R. Conformational polymorphism of the amyloidogenic peptide homologous to residues 113-127 of the prion protein. *Biophys J*. 2003 Jul;85(1):473-83.
22. Di Natale G, Impellizzeri G, Pappalardo G. Conformational properties of peptide fragments homologous to the 106-114 and 106-126 residues of the human prion protein: a CD and NMR spectroscopic study. *Org Biomol Chem*. 2005 Feb 7;3(3):490-7.

23. Glover KJ, Whiles JA, Wood MJ, Melacini G, Komives EA, Vold RR. Conformational dimorphism and transmembrane orientation of prion protein residues 110-136 in bicelles. *Biochemistry*. 2001 Nov 6;40(44):13137-42.
24. Marassi FM, Opella SJ. NMR structural studies of membrane proteins. *Curr Opin Struct Biol*. 1998 Oct;8(5):640-8.
25. Claridge T. High-resolution NMR techniques in organic chemistry. 1st ed. Pergamon; 1999.
26. Evans JNS. Biomolecular NMR spectroscopy. illustrated edition. Oxford University Press, USA; 1995.
27. Bohlke JK, De Laeter JR, De Bievre P, Hidaka H, Peiser HS, Rosman KJR, Taylor PDP. Isotopic compositions of the elements, 2001. *Journal of Physical and Chemical Reference Data*. 2005;34(1):57-68.
28. Huth JR, Bewley CA, Jackson BM, Hinnebusch AG, Clore GM, Gronenborn AM. Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR. *Protein Sci*. 1997 Nov;6(11):2359-64.
29. Cavanagh J, Fairbrother WJ, III AGP, Skelton NJ, Rance M. Protein NMR spectroscopy: principles and practice. 2nd ed. Academic Press; 2006.
30. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR*. 1995 Jan;5(1):67-81.
31. Parella T. Pulse Program Catalogue: II. Biomolecular NMR experiments. In: *NMRGuide4.1 - Topspin 2.0*. Bruker Biospin; 2006.
32. Wishart DS, Nip AM. Protein chemical shift analysis: a practical guide. *Biochem Cell Biol*. 1998 ;76(2-3):153-63.
33. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M, Appel R, Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker J, editor(s). *The proteomics protocols handbook*. Humana Press; 2005. p. 571-607.
34. Parella T. Pulse Program Catalogue: I. 1D & 2D NMR Experiments. In: *NMRGuide4.1 - Topspin 2.0*. Bruker Biospin; 2006.
35. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR*. 1995 Nov 1;6(3):277-293.

36. Johnson BA, Blevins RA. NMR View: A computer program for the visualization and analysis of NMR data. *Journal of Biomolecular NMR*. 1994;4(5):603-614.
37. Zhang O, Kay LE, Olivier JP, Forman-Kay JD. Backbone ¹H and ¹⁵N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J Biomol NMR*. 1994 Nov;4(6):845-58.
38. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. 2005 Jun 1;59(4):687-96.
39. Rieping W, Habeck M, Bardiaux B, Bernard A, Malliavin TE, Nilges M. ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics*. 2007 Feb 1;23(3):381-2.
40. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*. 2004 Oct;25(13):1605-12.
41. Varshavsky A. The N-end rule: functions, mysteries, uses. *Proc Natl Acad Sci U S A*. 1996 Oct 29;93(22):12142-9.
42. Wagner G, Braun W, Havel TF, Schaumann T, Go N, Wüthrich K. Protein structures in solution by nuclear magnetic resonance and distance geometry : the polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *Journal of Molecular Biology*. 1987 Aug 5;196(3):611-639.
43. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*. 1999 Mar;13(3):289-302.