Journal of the Korean Magnetic Resonance Society **1998**, *2*, 85-105

Solution Structure of a Prion Protein: Implications for Infectivity

He Liu[‡], Shauna Farr-Jones[‡], Nikolai Ulyanov[‡], Manuel Llinas[§], Susan Marqusee[§], Fred E. Cohen[‡,⊥,#], Stanley B. Prusiner[¶,∥,⊥] and Thomas L. James[‡,*]

Departments of Pharmaceutical Chemistry[‡], Neurology[¶], Medicine[⊥], and Biochemistry & Biophysics[∥], University of California, San Francisco, CA 94143
and Department of Molecular and Cell Biology[§], University of California, Berkeley, CA 94720
Received September 15, 1998

Abstract: Prions cause neurodegenerative diseases in animals and humans. The scrapie prion protein (PrP^{Sc}) is the major – possibly only – component of the infectious prion and is generated from the cellular isoform (PrP^{C}) by a conformational change. Limited proteolysis of PrP^{Sc} produces a polypeptide comprised primarily of residues 90 to 231, which retains infectivity. The three-dimensional structure of rPrP(90-231), a recombinant protein resembling PrP^{C} with the Syrian hamster (SHa) sequence, was solved using multidimensional NMR. Low-resolution structures of rPrP(90-231), synthetic peptides up to 56 residues, a longer (29-231, full-length) protein with SHa sequence, and a shorter (121-231) protein with the mouse sequence have been previously reported. We report here further structure refinement of rPrP(90-231) and dynamic features of the protein. Consideration of these features in the context of published data suggests regions of conformational heterogeneity, structural elements involved in the $PrP^{C} \rightarrow PrP^{Sc}$ transformation, and possible structural features related to a species barrier to transmission of prion diseases.

[1]Abbreviations: PrP, prion protein; PrP^{Sc}, scrapie PrP isoform; PrP^{C}, cellular PrP isoform; Mo, mouse; SHa, Syrian hamster; GPI, glycosylphosphatidylinositol; CD, circular dichroism; 1D, one-dimensional; 2D, two-dimensional; 3D, three-dimensional; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; HSQC, heteronuclear single-quantum coherence; T_1, spin-lattice relaxation time; T_2, spin-spin relaxation time; PFG, pulsed field gradient ; CSI, chemical shift index; RMSD, atomic root-

mean-square-deviations; BSE, bovine spongiform encephalopathy; CJD, Creutzfeldt-Jacob disease; GSS, Gerstmann-Sträussler-Scheinker syndrome; FFI, fatal familiar insomnia

INTRODUCTION

Prions are infectious particles that cause fatal neurodegenerative disorders, a newly identified class of diseases in humans and animals that manifest sporadic, inherited, and infectious illnesses.[1] The four neurologic disorders in humans are kuru, Creutzfeldt-Jakob disease (CJD), Gerstmann-Sträussler-Scheinker disease (GSS), and fatal familial insomnia (FFI).[2-4] Familial CJD, GSS and FFI are autosomal dominant diseases caused by mutations in the PrP gene that are transmissible to experimental animals. In animals, bovine spongiform encephalopathy, or "mad cow" disease, has caused more than 160,000 cattle deaths in Great Britain and is thought to be caused by a meat and bone meal dietary supplement containing prion-contaminated offal from sheep and cattle.[5] Recent reports suggest that bovine prions from mad cows have been transmitted to humans.[6-8] In contrast to viruses and viroids, prions do not contain a nucleic acid genome encoding their progeny. Rather, prions are composed largely, if not entirely, of a modified host-encoded glycoprotein denoted PrP^{Sc}. Through a post-translational process, PrP^{Sc} is formed from the normal, cellular PrP isoform designated PrP^{C}. Attempts to identify a post-translational chemical modification responsible for conversion of PrP^{C} into PrP^{Sc} were unsuccessful; both PrP^{C} and PrP^{Sc} possess a GPI anchor at the C-terminus, a disulfide bridge linking residues 179 and 214, and are glycosylated at Asp^{181} and Asp^{197}, but no covalent chemical differences between the two isoforms have been found.[9] The glycosylphosphatidylinositol (GPI) anchor of PrP^{C} apparently targets it to caveolar-like structures within or adjacent to the plasma membrane where PrP^{C} is either degraded or converted into PrP^{Sc}.[10]

Although the two isoforms of PrP have identical chemical properties, they possess very different physicochemical properties. Spectroscopic studies indicate that PrP^{C} is high in α-helix content and nearly devoid of β-sheet, while PrP^{Sc} has about 40% β-sheet.[11,12] The properties of PrP^{Sc} differ considerably from those of PrP^{C}. Although PrP^{C} is readily degraded by proteases, limited proteolysis of PrP^{Sc} yields an N-terminally truncated protein comprised primarily of residues 90-231; this truncated PrP retains scrapie infectivity.[13] It appears that PrP^{Sc} propagates

prion diseases by catalyzing the conversion of PrP^C to PrP^{Sc}.[1] This is supported by the observation that transgenic mice lacking the PrP gene do not develop scrapie at all, regardless of exposure to PrP^{Sc}.

Determination of the conformations of PrP^C and PrP^{Sc} is crucial for understanding the molecular basis of prion diseases. The secondary structure and conformational transition of peptides consisting of residues 90-145, and 109-141, including the most conserved hydrophobic stretch including residues 113-128 AGAAAAGAVVGGLGGY, were studied by NMR.[14] The peptides could be induced to form α-helical structures in aqueous solution in the presence of some organic solvents or detergents, or to form intermolecular β-sheet with physiological salt added. Remarkably, experiments seeking to simulate the interaction between PrP^C and PrP^{Sc} by mixing the 90-145 peptide with PrP^C have revealed that PrP^C could be induced to form protease-resistant intermolecular β-sheet.[15] The peptide studies and the infectivity of the protease-resistant core PrP stimulated our choice to determine the structure and dynamics of rPrP(90-231), which is composed of the same sequence but corresponds to the cellular form of the protein.

Recombinant PrP, labeled uniformly with ^{15}N or $^{15}N/^{13}C$ in order to carry out multidimensional heteronuclear NMR studies, was produced via expression in *E. coli*. The recombinant prion protein was purified and refolded; spectroscopic analyses indicate that the refolded prion protein has α-helical character similar to the normal cellular isoform PrP^C. A low-resolution structure of the rPrP(90-231), refolded to resemble the cellular form, has been reported.[16] The core domain of SHa rPrP(90-231) has a structure similar to a shorter recombinant mouse prion protein, MoPrP(121-231),[17,18] *i.e.*, the core contains three α-helices and two short antiparallel β-strands. The primary but not only difference between SHa rPrP(90-231) and Mo PrP(121-231) is that the former exhibits additional structure even in the region where the sequences are coincident, *e.g.*, an additional turn of helical structure at the beginning of the second helix in SHa rPrP(90-231), the third helix (C-terminus) is three turns longer than that in MoPrP(121-231), and residues 167-176, which is not defined in MoPrP(121-231), form a well-defined loop.

The structure and dynamics of the recombinant "full-length" SHa prion protein rPrP(29-231) has also been studied and compared with SHa rPrP(90-231).[19] rPrP(29-231) has identical core structure based on chemical shifts and NOE connectivities, and similar backbone dynamics to rPrP(90-231), with the N-terminal residues 29-124, including octarepeats, highly flexible and mostly unstructured. A study of full-length mouse rPrP(23-231) also concluded that the N-terminus was disordered.[20]

Further structure refinement of SHa rPrP(90-231) together with a presentation of backbone dynamics and hydrogen-deuterium exchange of amide protons is included here. The refined structure deviates little from the previously published low-resolution structure,[16] especially for secondary structure elements and overall folding but, with more restraints assigned, the backbone atomic root-mean-square-deviation (RMSD) is improves from 2.0 Å to 0.67 Å. Comparison of SHa rPrP(90-231) and MoPrP(121-231) suggests regions of relative stability and structural differences which may pertain to species specificity of prion diseases.

METHODS

NMR Samples and Experiments

Solutions of rPrP(90-231), uniformly ^{15}N-labeled or ^{15}N/^{13}C-labeled rPrP(90-231), were concentrated to a final protein concentration of 1 mM, 20 mM sodium acetate, pH 5.2, 0.005% sodium azide in 10% D_2O/90% H_2O for NMR studies. Self-diffusion experiments were performed on the 1 mM sample as well as a 0.33 mM sample prepared as above except that the solvent contained 99% D_2O.

The following NMR spectrometers were used: Bruker AMX500, DRX600, RMX750 and Varian UnityPlus600, each equipped with a 5-mm ^{13}C/^{15}N/^{1}H triple resonance probe and pulsed field gradients (PFG). Three-dimensional heteronuclear triple resonance experiments,[21] HNCA,[22] HNCACB,[23] CBCA(CO)NH,[24] HNCO,[22] and HBHA(CO)NH,[25] that correlate resonances from backbone HN, ^{15}N, H$^{\alpha}$, ^{13}C$^{\alpha}$ and ^{13}CO, and sidechain H$^{\beta}$ and ^{13}C$^{\beta}$, were acquired for backbone resonance assignments as described previously [16] HNCA, HNCACB, and CBCA(CO)NH spectra were acquired at 500 MHz. HNCO and HBHA(CO)NH spectra were acquired at 600 MHz.

For assignment of sidechain aliphatic ^{13}C and ^{1}H resonances, the 3D heteronuclear experiments C(CO)NH-TOCSY[26] and HCCH-TOCSY[27] were recorded at 600 MHz with isotropic mixing times of 11 and 22 ms. An ^{15}N-edited NOESY-HSQC spectrum with 50 ms mixing time was acquired at 600 MHz proton frequency and used to assist assignments complicated by overlap and spin diffusion. All spectra were referenced relative to 3-(trimethylsilyl)-tetradeutero-sodium propionate (TSP) at 0 ppm.

To obtain distance restraints for structure calculation, 100 ms ^{15}N-edited NOESY-HSQC[28] and ^{13}C-edited NOESY-HSQC spectra[29] were acquired at 750 MHz for protons. J coupling constants were measured using a 3D HNHA experiment for backbone ϕ torsion angle restraints.[30]

Spin-lattice relaxation time (T_1), spin-spin relaxation time (T_2) and heteronuclear $^{15}N\{^1H\}$ values were obtained using spectra acquired via pulse sequences previously described.[31] NMR was utilized to measure the self-diffusion constant for rPrP(90-231) in D_2O buffer at 600 MHz using the pulsed field gradient method.[32] We determined the diffusion constant of a 1 mM sample of rPrp90-231 at pH 5.2 (the same sample used for some of the structural studies), a 0.1 mM sample obtained by diluting into D_2O buffer a sample that had been used for structural studies, and a sample at 0.33 mM which had never been concentrated to greater than 0.33 mM protein. A plot of ln (amplitude) *vs.* gradient strength squared was linear over that range for all three samples.

Amide proton-deuterium exchange was monitored using ^{15}N-labeled protein. ^{15}N-HSQC spectra were collected at 25 °C over a period of two weeks. At the earliest time point measured, 48 amide protons could be detected. The observed exchange rates (k_{obs}) were calculated by fitting peak intensities as a function of exchange time to a single exponential. Protection factors (P) were calculated as $P=(k_{rc}/k_{obs})$ where k_{rc} are the known intrinsic exchange rates for a residue in a specific tripeptide context.[33]

Nuclear Resonances Assignment

SPARKY, the NMR spectral display and analysis program developed at UCSF, was used to assign NMR spectra. SPARKY is interfaced with the molecular graphics program MidasPlus to make the display and examination of assignments in the multiple spectra and interim structure convenient. The program also does assignment bookkeeping. (See http://picasso.ucsf.edu /software.html for information about SPARKY.) Backbone ^1H, ^{13}C and ^{15}N resonances of rPrP(90-231) were assigned principally via the connectivity of sequential backbone amides using the triple resonance experiments HNCACB and CBCA(CO)NH. As HNCA has better sensitivity than HNCACB, it is used to establish backbone connectivities when sensitivity loss in the HNCACB spectrum renders assignments difficult. With the backbone amide protons and nitrogens assigned, the ^{13}C resonance of CO can be assigned readily using the HNCO experiment, as well the H^α and H^β protons using HBHA(CO)NH. Assignment of these latter two spectra also confirms the other backbone resonance assignments.

The C(CO)NH-TOCSY spectrum allows correlation of backbone amides to sidechain ^{13}C's, permitting assignment of C^γ's. ^1H and ^{13}C resonance assignments of aliphatic sidechains were completed with HCCH-TOCSY, which provides correlations among the ^{13}C's of the

sidechains through direct as well as multiple-relayed magnetization transfers.

Assignments for amino groups of asparagine and glutamine residues were made using HNCA, HNCACB, and CBCA(CO)NH spectra by correlating sidechain amide proton and nitrogen with intraresidue C^{α} and C^{β}'s for Asn, and C^{β}'s and C^{γ}'s for Glu. The ^{15}N NOESY-HSQC spectrum was used to assist the assignment for partially overlapped amino groups, using the cross-peaks to H^{β} protons or H^{γ} protons as additional information for identifying the amino groups of Asn or Glu, respectively. The $\gamma 2$ and $\delta 1$ methyl groups in isoleucine residues were assigned based on ^{13}C-edited NOESY-HSQC cross-peaks that resulted from two-bond proton-heteronuclear scalar interactions (J couplings). The $\gamma 2$ methyl has a cross-peak to H^{β} due to two-bond H^{β}- C^{β}- $C^{\gamma 2}$ coupling. The $\delta 1$ methyl has a cross-peak to the $\gamma 1$ methylene due to two-bond $H^{\gamma 1}$- $C^{\gamma 1}$- $C^{\delta 1}$ J coupling.

Assignment of NMR Structural Restraints

SPARKY was also used for assigning NOE cross-peaks. The program has tools to search for 3D symmetric peaks for assignments, makes assignment suggestions and lists alternative assignments based on chemical shifts, symmetric peaks, or model structures. Its bookkeeping provisions permit listing of alternative assignments with annotations during the long process of making protein resonance assignments. A total of 1895 NOE cross-peaks were assigned unambiguously in ^{15}N- and ^{13}C-edited NOESY spectra. Ambiguous assignments with degeneracy less than 3 were made for an additional 162 peaks. Additional NOE restraints were assigned using ARIA during structure refinement.[34] The total number of unambiguous NOE restraints after ARIA refinement was 2778, and the number of ambiguous restraints was 306.

NOE intensities were classified as strong, medium, weak, or very weak, respectively corresponding to distance restraints of 1.7-2.8 Å, 2.0-3.5 Å, 3.0-5.0 Å, or 3.5-6.0 Å. The distance calibration is based on NOE peaks that are assigned to the approximately known distances [such as $d_{\alpha N}(i,i+3)$] in the α-helices identified by chemical shift indices and preliminary structure calculation. Wide distance error bounds were used to reflect experimental errors as well as the fact that spin diffusion was not taken into account. For equivalent atoms in methyl groups and aromatic rings, averaging of the NOE contributions is used in the ARIA/X-PLOR calculation, so the correction to the distance bounds is not necessary.

Three-bond $^{3}J_{H^{N}H}\alpha$ scalar coupling constants were measured using the 3D HNHA spectrum that correlates intraresidue H^{N}, N, and H^{α},[35] to provide information about the dihedral backbone

angle ϕ. J coupling constant error bounds ranged from ±0.44 to 0.89 Hz. Sixty-four J coupling restraints were used directly in structure refinement.

Forty-one hydrogen bonding restraints were identified for the helices. Hydrogen bonding was determined based on NOEs between the amide proton and α protons at residues i-3 and i-4 [$d_{\alpha N}(i,i+3)$ and $d_{\alpha N}(i,i+4)$], an NOE between an α proton and residue (i+3) β protons [$d_{\alpha\beta}(i,i+3)$], and observation of weak amide proton-solvent exchange peaks. Three hydrogen bond restraints were used for the short antiparallel β-sheet. These hydrogen bonds were determined based on the observed interstrand NOE connectivity patterns. Some of the more inert H-bonds were confirmed by the hydrogen-deuterium exchange. However, the ends of helices are less stable, so hydrogen bonding cannot be detected by the exchange experiment. Hydrogen bond distance restraints, d_{HN-O} = 1.5-2.3 Å and d_{N-O} = 2.5-3.3 Å, were used for each hydrogen bond.

Structure Refinement

Initial structures of rPrP(90-231) were calculated using a modified version of X-PLOR 3.851, compiled with the ARIA (Ambiguous Restraints for Iterative Assignment) routines,[34] which allows the use of ambiguous distance restraints and also provides the capability to assign automatically ambiguous NOE restraints in an iterative fashion during structure refinement. Manually assigned NOE restraints were used to generate 100 structures from an extended polypeptide chain corresponding to the sequence of rPrP(90-231) in iteration zero. The best 20 structures, with the lowest total energy, were used in further iterations. The interface of ARIA routines with the X-PLOR simulated annealing calculation makes it convenient to assign ambiguous NOE restraints based on the best structures obtained at each stage of the refinement. Among the 20 structures calculated at each level of iteration, the best 10 were used to calibrate NOE peaks and assign ambiguous restraints. With decreasing level of assignment ambiguity in the iterative procedure, more restraints could be assigned unambiguously. Following ARIA, 2778 unambiguous and 306 ambiguous restraints, 63 J coupling constants and 44 hydrogen bond restraints were used to calculate 100 structures with the X-PLOR restrained simulated annealing protocol. The best 25 structures were selected based on total energies.

The 25 structures calculated with the X-PLOR simulated annealing protocol were each subjected to 25 ps of restrained molecular dynamics in a 6 Å shell of water using the AMBER 4.1 force field for better simulation of nonbonded interactions.[36] Only unambiguous restraints (assigned either manually or by ARIA) were used. The unstructured N-terminal residues 90-111 were removed before solvation of the protein with the 6 Å water shell. The resulting

structures have been deposited in the Brookhaven Protein Data Bank with accession code 2prp. Table 1 lists statistics pertinent to the refined structures.

RESULTS AND DISCUSSION

Characterization

For proper folding, the disulfide linking C179 and C214 must be maintained.[37] Circular dichroism confirmed that under conditions employed for our study, rPrP(90-231) was predominately α-helical, as is native PrP^C. The linearity of the pulse field gradient experiments over a large range of gradient strengths indicates that the rPrP(90-231) oligomeric state is homogeneous. The rPrP(90-231) protein had the same self-diffusion constant, $0.6 \times 10^{-6} \pm 0.1 \times 10^{-6}$ cm^2/s, at concentrations of 0.1 mM, 0.33 mM and 1.0 mM at 30 °C. This value is slightly high, but not nearly large enough to imply dimer formation. However, the protein is not spherical (*vide infra*); furthermore, the unstructured and solvated N-terminus would be expected to increase the diffusion constant relative to a compact spherical protein of the same molecular weight. It is known from sedimentation equilibrium experiments that at a concentration of 0.025 mM, rPrP(90-231) in the α form is monomeric.[37] Because of the different concentrations required for NMR and equilibrium sedimentation experiments, we are not able to determine the sedimentation distribution of the NMR sample. However, since the same diffusion coefficient is observed for the NMR sample over the range 0.1 – 1.0 mM and the oligomeric state is homogeneous, the self-diffusion results strongly suggest that the NMR sample of rPrP(90-231) at 1 mM is monomeric.

Secondary Structure

Deviation of C^α, CO, and H^α chemical shift values relative to random coil values are indicative of protein secondary structure and can be conveniently described by a Chemical Shift Index (CSI),[38,39] and for rPrP(90-231), provided the first glimpse of secondary structure.[16] The first 30 residues of the N-terminus exhibit only small chemical shift deviations from random coil and manifest no systematic shifts except for C^α resonances. This region also has weak NOEs with few medium- and long-range NOEs, as well as narrow NMR linewidths. Together this suggests that the first 30 residues are largely unstructured in solution. Although the N-terminal residues apparently lack stable secondary structure, transient structure may exist. For example, some medium and weak long-range NOEs identified for residues 96-101 and 112-128.

The chemical shift indices and NOE connectivities for rPrP(90-231) lead to the secondary

structure diagram shown in Figure 1. The chemical shift deviations indicate three α-helices spanning residues 144-156, 172-194, and 200-227, which are termed helices A, B and C, respectively.

These helices indicated by CSI are supported by $d_{NN}(i,i+1)$, $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$, and $d_{\alpha N}(i,i+4)$ connectivities. The refined structure of rPrP(90-231) contains three helices that span essentially the same residues as predicted by CSI (*vide infra*). CSI analysis suggests two short β-strands may exist. Residues 159-164 show β character consistently in all C^{α}, CO, H^{α} chemical

Table 1. Structure Refinement of rPrP(90-231). Structural Restraints Input and Statistics Pertaining to the Resulting Ensemble of 25 Structures

Number of distance restraints	
Unambiguous	2778
Ambiguous	306
Number of torsion angle restraints from $^{3}J_{H^{N}H}\alpha$	63
Number of hydrogen bond restraints	44
Energy* (kcal/mol)	
E_{total}	-741 ± 27
E_{vdw}	-761 ± 20
$E_{constraint}$	430 ± 34
E_{AMBER}	-1171 ± 61
Residual distance deviation, Å	0.032 ± 0.002
Residual $^{3}J_{H^{N}H}\alpha$ deviation, Hz	0.059 ± 0.015
Atomic RMSD of residues 125-227†, Å	
Backbone (N,C^{α},C)	0.67
all heavy atoms	1.35

* Energy was calculated using the AMBER 4.1 force field with the AMBER program. Restrained molecular dynamics and restrained energy minimization were carried out for residues 112-231 with a 6 Å shell of water. Water molecules were removed to calculate the energy terms (± standard deviations) presented.

† Average parameters for 25 refined structures, ± standard deviations.

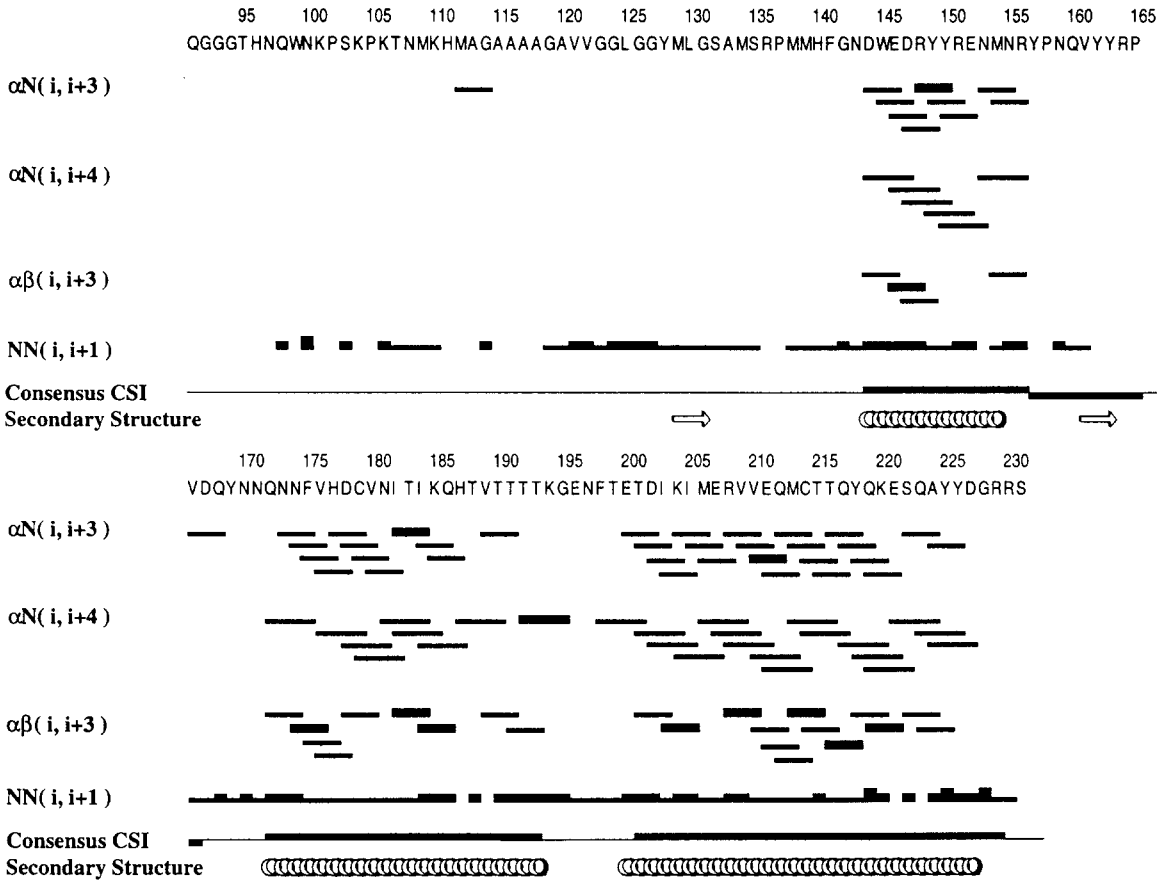

Fig. 1. Secondary structure diagram for rPrP(90-231). NOE connectivities are denoted by lines, where the thickness qualitatively represents the relative intensity (weak, medium or strong) of the NOE cross-peaks, and i designates the residue number for rPrP. $\alpha N(i, i+3)$ denotes an NOE between the α-proton of residue i and the amide proton of residue $i+3$. The other NOEs are similarly defined. For the consensus chemical shift index,[38] contiguous up bars indicate α-helix and down bars indicate β-strand. Regions of secondary structure are depicted by helices for α-helices and broad arrows for β-strands.

shift data, but the complementary strand, residues 128-133, exhibits much weaker β character. According to C^{α} chemical shifts, only residues 129 and 130 shifts imply β character, but ^{13}CO chemical shifts suggest the β-strand may span residues 127-134, while $H^{\alpha}CSI$ values indicate that residues 129-132 are shifted to β. The analysis of NOE cross-peaks shows that hydrogen bonds exist between M129 O and Y163 H^{N}, between M129 H^{N} and Y163 O, and between V161 O and G131 H^{N}. For example, Figure 2 shows the cross-strand NOE peaks due to interaction of L130 H^{α} and Y162 H^{α}. The two β-strands form an antiparallel β-sheet. According to the structures obtained using NMR restraints and structure refinement (*vide infra*), the small β-sheet forms most consistently for residues 129-131 and 161-163, while other residues in the sequences 159-164 and 128-133 exhibit only some β-sheet characteristics.

Amide hydrogen-deuterium exchange experiments were carried out to determine the regions of rPrP(90-231) most resistant to exchange. All secondary structure elements in rPrP(90-231) contain amide protons protected from exchange with water, *i.e.*, have detectable resonances in the ^{15}N-HSQC experiment after 20 minutes of exchange with deuterium oxide. Helix A (residues 144-154) amides are protected. The amide protons of residues 175 to 186 in helix B are protected from proton/deuterium exchange. However, lack of protection through to the C-terminus of helix B (residue 194) suggests that this end of the helix is more dynamic permitting amide proton exchange with solvent. The limited helix stability for residues 187-194 is also manifest in the strong solvent exchange peaks for amides in the ^{15}N edited NOESY-HSQC spectrum. However, residues 187-194 do consistently form a helix in the structure calculations using the structural restraints in this region. Residues 199-221 in helix C are also well protected according to the H/D exchange data, but the C-terminal end of this helix, residues 222-227, lacks probes for H/D exchange. The extension of helix C to D227 is supported by medium-range $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$, and $d_{\alpha N}(i,i+4)$ connectivities (see for example Figure 3) as well as the chemical shift indices. Long-range NOEs from V166 to Y218, E221, S222 and Y225, and from Q172 to Y215 and Q219 also indicate that the region 222-227 is helical.

In the rPrP(90-231) β-sheet region, H/D exchange data indicate that strand 2 (S2) is more protected than strand 1 (S1) of the β-sheet. Residues 161-164 are protected from exchange, but only the L130 amide is sufficiently stable in S1. Note that loop residue V166 is protected. The sidechain of V166 interacts extensively with helix C [16] and also is the beginning of a helical turn formed by residues 166-168.

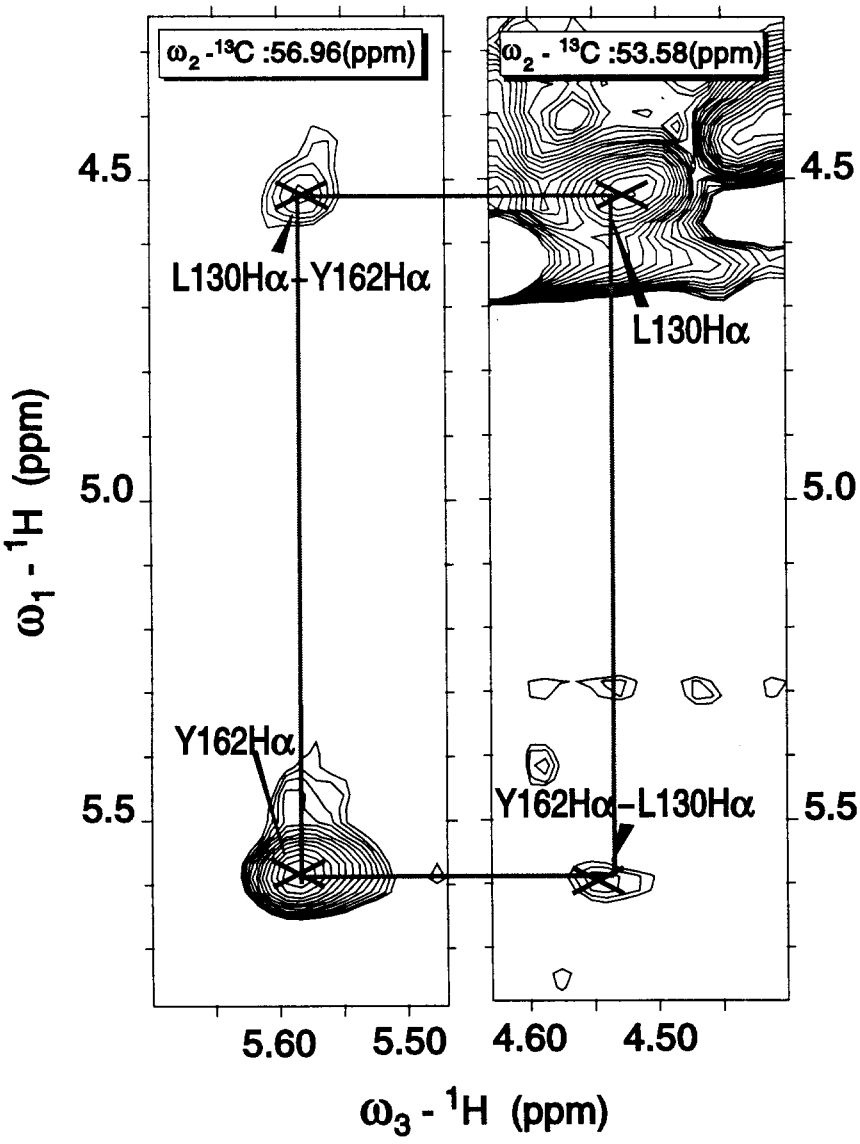

Fig. 2. NOE cross-peaks (symmetric mirror peaks) in ^{13}C-edited NOESY spectra between L130 H^{α} and Y162 H^{α} illustrating the formation of antiparallel β-sheet.

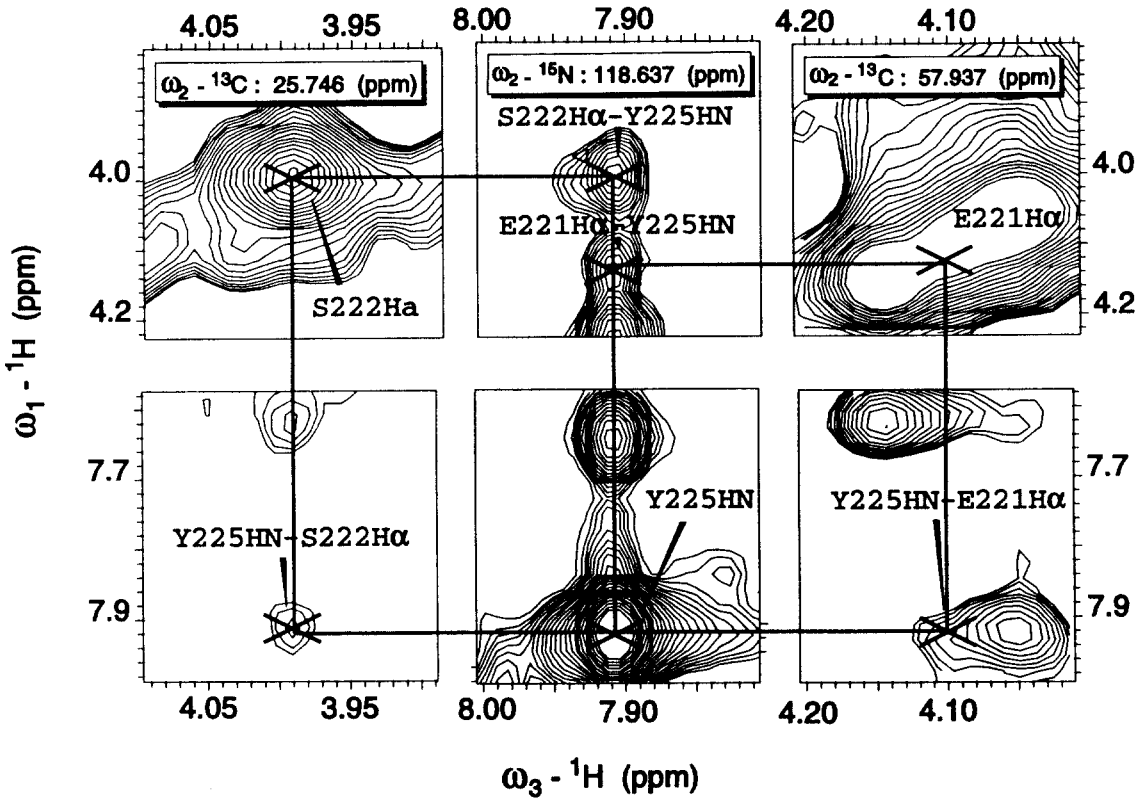

Fig. 3. NOE cross-peaks in ^{15}N- and ^{13}C-edited NOESY spectra illustrating the extension of helix C. Specifically, cross-peaks from Y225 H^N to E221 H^α and S222 H^α are shown. Both are shown with the symmetric mirror peaks in the same ^{13}C-edited NOESY-HSQC spectrum.

Three-dimensional Structure

Structure refinement via ARIA/X-PLOR and AMBER yielded an ensemble of twenty-five rPrP(90-231) structures. The 25 structures were superimposed by minimizing the RMSD of backbone CO, C^{α} and N atoms of residues 125-228 to the average structure. An RMSD of 0.67 Å for backbone atoms and 1.35 Å for all heavy atoms was found for the ensemble. The 25 structures were evaluated with PROCHECK,[40] the structure with the lowest AMBER energy and good structural parameters is shown in Figure 4. A Ramachandran plot for residues 125-228 of the final 25 structures of rPrP(90-231) reveals that the backbone torsion angles ϕ and φ fall within the areas favorable for α-helix or β-sheet structure (86.6% in most favorable, 11.6% in allowable, and 1.3% in generously allowable regions). However, ϕ and φ angles for residues in less defined loops display a greater variety; for residues 112-231, 79.7% of the values fall within the most favorable regions of the Ramachandran plot and an additional 15.2% fall within allowed regions.

The rPrP(90-231) structure has a core region consisting of helices A, B and C. Helix A includes residues 144-153. At the C-terminus of helix A, a hydrogen-bonded turn extends to residue Y157. All 25 structures, except one, exhibit this turn, which is identified as a 3/10 helix by the Kabsch & Sander algorithm.[41] Helix B covers residues 172 through 193. The N-terminus of helix B is capped by N171, but the C-terminus continues as a turn to residue K194, being capped by G195. As noted above, the helical structure of residues 187-194 is not stable. With inadequate NOE restraints in this region (partially due to overlap of four sequential threonines), helical structure is obtained only with ambiguous medium-range NOE restraints assigned by automated ARIA/X-PLOR iterative calculations; the chemical shift index, however, is also consistent with helix B extended through T193. Helix C is the longest helix in the structure. It spans residues 200-227; it is the most stable helix in rPrP(90-231).

Strands S1 and S2, shown with green and blue arrows, respectively, in Figure 4, comprise a small β-sheet spanning residues 129-131 and 161-163 in the respective strands. It seems likely that extended β-strands may exist at least transiently. Residues 159-164 form a good □ structure in the 25-structure ensemble, but the C-terminus of the S1 strand, residues 128-134, is distorted, in accord with CSI and hydrogen-deuterium exchange data. There are extensive cross-strand NOE connectivities between the extended strands (colored green and blue in Figure 4), but there is no obvious NOE cross-peak pattern to support a longer regular antiparallel β-sheet. The flexibility of S1 is supported by all our data. The G131 amide resonance exhibits a significantly broader linewidth (44 Hz vs. 24-26 Hz for most core residues) which is quite weak. This implies the existence of multiple exchanging conformations. We note that when G131 points away from

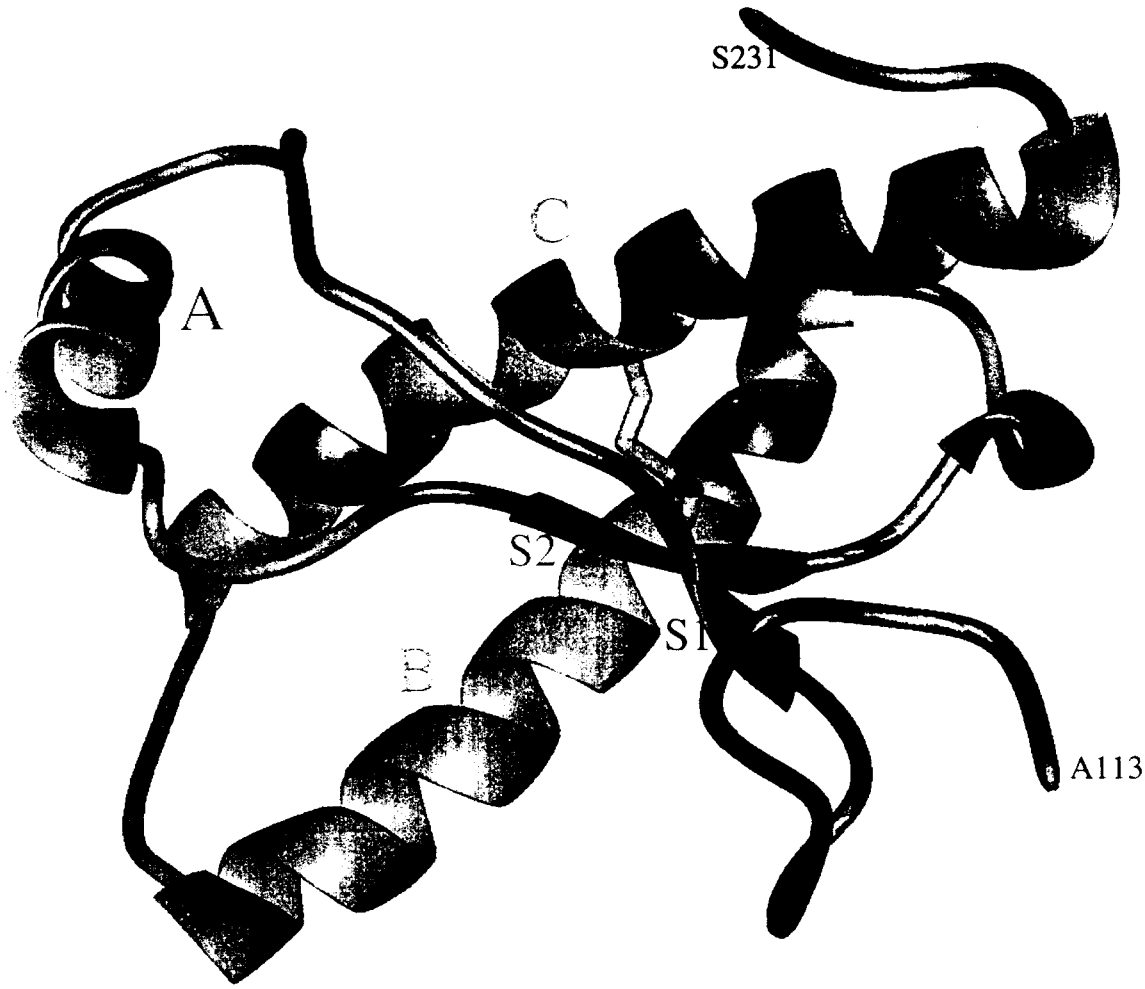

Fig. 4. Ribbon representation of a rPrP(90-231) structure. The three helices A, B and C are colored magenta. Helices B and C are linked by a disulfide bond (yellow) formed between residues C179 and C214. Strands S1 and S2 of the β-sheet were shown as green and blue arrows, respectively. Residues participating in the extended antiparallel interstrand interactions of the β-strands were also colored green and blue. A helical turn, represented by a short gray ribbon, was identified for residues 165-169 between S2 and helix B. The red segment is the hydrophobic cluster of residues 113-128. The loops and the unstructured N-terminus are colored gray.

V161, S132 is nearly within H-bonding distance of V161 N-H. The V161 backbone also presents two conformations in the ensemble of refined structures, but neither puts the amide within H-bonding distance of G132 on S1. There is also evidence for more than one conformation for V161 in the ^{15}N HSQC spectrum which has two peaks for the V161 amide.

Residues 165-169 between S2 and helix B (gray ribbon in Figure 4) are recognized by the Kabsch & Sander algorithm as forming either a 3/10 helix turn (18/25 structures) or a hydrogen-bonded turn (7/25 structures). In our present refined structure, the loop 165-171 is well-defined, with all NMR restraints between the loop and helix C satisfied, as well as those within the loop. As noted above, the stability of this turn has resulted in the V166 amide being substantially protected from exchange with bulk water.

The sequence spanning residues 113-128 (red in Figure 4) is composed of glycines and hydrophobic residues. This region seems to lack regular secondary structure, but there is evidence of structure. Weak medium- and long-range NOEs were evident for this region, but many are ambiguous. About 30 medium- or long-range NOEs could be assigned for this hydrophobic region and used in structure calculations. Residues 113-115 are largely disordered; residues 116-118 also exhibit less order but may be involved in a turn. Weak or transient interactions involving methyl groups of V121 or V122 are evident in NOESY-HSQC spectra for this region resulting in a pronounced bend or turn evident in the 25-structure ensemble. There are two or three bends or turns for this hydrophobic segment. The most consistent hydrogen-bonded turn involves residues 116-123, and the most consistent bend occurs for residues 124-127. The existence of some persistent structure in the hydrophobic cluster region is also manifested by resonances shifted significantly from random coil values: methyl groups of L125 are shifted upfield about 0.2 and 0.4 ppm from ring current effects emanating from Y162 and Y128. Long-range connectivities indicate that the hydrophobic cluster has interactions with β-strand S1 and with helix B via contacts between the sidechains of L125, I182 and Q186. In the hydrophobic cluster, residues 121-128 are relatively well-ordered, but two amide signals are observed in the ^{15}N-edited NOESY spectrum for A120 indicating that A120 exists in at least two conformations. It is possible that this same conformational exchange results in the ^{15}N HSQC spectral signal broadening noted above for Gly131 (44 Hz) as well as Met129 (32 Hz) and Leu130 (32 Hz) and S2 residues Tyr162 (33 Hz) and Arg164 (33 Hz). The amide linewidths for most residues in the protein core are 24-25 Hz. Conformational heterogeneity in the hydrophobic cluster and the adjacent irregular β-sheet may play a crucial role in the $PrP^{C} \rightarrow PrP^{Sc}$ transformation.

Protein Backbone Dynamics

To characterize the rapid backbone dynamics of rPrP(90-231), the ^{15}N NMR longitudinal relaxation rates R1 ($1/T_1$), transverse relaxation rates R2 ($1/T_2$) and $^{15}N\{^1H\}$ NOE values for the amides were measured at 600 MHz. The relaxation data were analyzed using MODELFREE, version 3.0.[42] We found that it was not possible to describe satisfactorily the backbone dynamics using the popular model-free style of analysis, even when fast and slow molecular motions were assumed. Most probably is due to the assumption that protein tumbles isotropically. A structured protein core would certainly not tumble isotropically if it is dragging a long minimally-structured, but hydrated, N-terminus. The pronounced asymmetric shape of the core portion of the protein may also be troublesome. However, the raw ^{15}N relaxation parameter values indicate that the hydrophobic cluster is more dynamic than the core.

Interaction of Protein X with the Prion Protein

Studies using mouse neuroblastoma cells transfected with mutated chimeric Hu/MOPrP genes identified four residues as forming an epitope which interacts with a molecular chaperone, tentatively named protein X, that assists in conversion of cellular PrP^C to the scrapie form, PrP^{Sc}: Q168 in the stable turn between S2 and helix B, Q172 at the beginning of helix B, and T215 and Q219 on helix C.[43] Q168, as part of the helical turn, is not in close contact with the other three residues that appear to be involved in protein X binding (Figure 4). However, upon binding to protein X, it is quite possible that Q168 moves into close proximity to Q172, T215 and Q219 which already form a nice surface for binding.

Comparison of rPrP(90-231) with Recombinant MoPrP(121-231)

There are only six sequence differences between Syrian hamster and mouse PrP sequences in the region 121-231: residues 139 (M›I), 155 (N›Y), 170 (N›S), 203 (I›V), 205 (I›M), and 215 (T›V).[44] Our comparison between the MoPrP(121-231) structure [17,18] [Brookhaven Protein Data Bank accession code 1AG2] and the corresponding region of the SHa rPrP(90-231) structure reveals that the overall folding of these two proteins is the same, with an RMSD of 1.8 Å for the backbone atoms of residues 128-214, *i.e.*, without the hydrophobic cluster or the three-turn extension to helix C, both of which we find with SHa rPrP(90-231) but which are not reported in MoPrP(121-231). Residues 167-176 were reported to be disordered in MoPrP(121-231) [17] (later revised to 167-171 [18]), but they are well-resolved in the spectra and resulting structures of rPrP(90-231) with linewidths comparable to those from other structured regions of the protein.

Three of the six sequence differences between MoPrP(121-231) and SHa rPrP(90-231), *i.e.*, at residues 155, 203 and 205, apparently do not have any structural effect. However, changes from T215 to V215 and S170 to N170, comparing the hamster sequence to the mouse-A sequence, may be substantially responsible for the structural differences evident in the C-terminus of helix C and loop 165-171. Substitution of V for T (which distinguishes the Mo-A from human and SHa sequences) may alone result in the structural change between the Mo-A and SHa proteins. In the C-terminal structure of the mouse protein, Y225 and Y226 are involved in a hydrophobic interaction with V215. Residues Q172, V215 and Q219, which are apparently involved in protein X binding (*vide supra*), are also partially covered by these Y225 and Y226. If the V215-Y225/Y226 hydrophobic interaction impedes the extension of helix C, it can also destabilize the 165-171 loop. In contrast to our observations with SHa rPrP(90-231), no long-range NOE connectivities reflecting interaction between V166 and the C-terminus of helix C has been reported for MoPrP(121-231).

The structure of the loop between helix A and β-strand S1, part of S1, and the first residue D144 of helix A, also differ between MoPrP(121-231) and SHa rPrP(90-231). This region includes residue 139, which is I in Mo and Hu sequences, and M in the SHa sequence. Thus we find structural differences at two of the three locations (residues 139, 155 and 170) which have been suggested to play a role in species specificity of prion diseases.[45]

Further Relationship of Structure to $PrP^C \rightarrow PrP^{Sc}$ transformation

The region 113-131 which essentially consists of the hydrophobic cluster is the most conserved sequence among mammals, indicating its probable importance for the normal function of PrP^C. However, that same region may be essential for transformation into the infectious form of the prion protein as well. This is supported by the fact that proteolytic cleavage yielding sequence 121-231 does not lead to infectivity, while the protease-resistant core sequence 90-231 does. Protease-resistant cores associated with different pathological phenotypes of prion diseases always contain this hydrophobic cluster, although with a mutation inducing β conversion in some cases; for example, A117V can lead to GSS. Our biophysical studies of peptide fragments from PrP indicate that residues 90-145 and 109-141 can form α-helical structure or β aggregates depending on solution conditions, but three shorter peptides containing either part or none of the hydrophobic core residues show little tendency for the $PrP^C \rightarrow PrP^{Sc}$ conversion.[14]

In the SHa rPrP(90-231) structure, M129 immediately precedes the β-strand S1. We previously remarked that the polymorphism of residue 129, being a valine in a significant number of humans, which is correlated with disease phenotype expressed by the D178N mutation can be

explained by the structure.[16] Methionine is a strong α-helix forming residue, while valine is the strongest β forming residue. The presence of valine at 129 could further stabilize S1, which may be important in expression of different disease phenotypes. As another example, note that GSS patients with different point mutations in the hydrophobic cluster region manifest different clinical phenotypes. The mutation A117V leads to GSS patients with dementia when residue 129 is methionine, but a family of GSS patients with valine at both residues 117 and 129 exhibited ataxia.[46] These different clinical manifestations suggest that the different prion diseases may result from different conformations of PrP^{Sc} being present.

Acknowledgements

We gratefully acknowledge Darlene Groth for preparation of the protein samples and Manuel Llinas for performing H/D exchange experiments. We are also very grateful to Dr. David G. Donne for acquisition of some of the NMR spectra at The Scripps Research Institute and to Drs. Peter E. Wright and H. Jane Dyson for access to those facilities. This work was supported by National Institutes of Health grant AG10770.

REFERENCES

1. S. B. Prusiner, and M. R. Scott, *Annu. Rev. Genet.* **31**, 139 (1997).
2. D. C. Gajdusek, *Science* **197**, 943 (1977).
3. C. L. Masters, D. C. Gajdusek, and C. J. Gibbs, Jr., *Brain* **104**, 559 (1981).
4. R. Medori, H.-J. Tritschler, A. LeBlanc, F. Villare, V. Manetto, H. Y. Chen, R. Xue, S. Leal, P. Montagna, P. Cortelli, P. Tinuper, P. Avoni, M. Mochi, A. Baruzzi, J. J. Hauw, J. Ott, E. Lugaresi, L. Autilio-Gambetti, and P. Gambetti, *N. Engl. J. Med.* **326**, 444 (1992).
5. R. M. Anderson, C. A. Donnelly, N. M. Ferguson, M. E. J. Woolhouse, C. J. Watt, H. J. Udy, S. MaWhinney, S. P. Dunstan, T. R. E. Southwood, J. W. Wilesmith, J. B. M. Ryan, L. J. Hoinville, J. E. Hillerton, A. R. Austin, and G. A. H. Wells, *Nature* **382**, 779 (1996).
6. R. G. Will, J. W. Ironside, M. Zeidler, S. N. Cousens, K. Estibeiro, A. Alperovitch, S. Poser, M. Pocchiari, A. Hofman, and P. G. Smith, *Lancet.* **347**, 921 (1996).
7. G. Chazot, E. Broussolle, C. I. Lapras, T. Blättler, A. Aguzzi, and N. Kopp, *Lancet.* **347**, 1181 (1996).
8. J. Collinge, K. C. L. Sidle, J. Meads, J. Ironside, and A. F. Hill, *Nature* **383**, 685 (1996).
9. N. Stahl, M. A. Baldwin, D. B. Teplow, L. Hood, B. W. Gibson, A. L. Burlingame, and S. B. Prusiner, *Biochemistry* **32**, 1991 (1993).

10. A. Taraboulos, M. Scott, A. Semenov, D. Avrahami, L. Laszlo, and S. B. Prusiner, *J. Cell Biol.* **129**, 121 (1995).
11. B. W. Caughey, A. Dong, K. S. Bhat, D. Ernst, S. F. Hayes, and W. S. Caughey, *Biochemistry* **30**, 7672 (1991).
12. K.-M. Pan, M. Baldwin, J. Nguyen, M. Gasset, A. Serban, D. Groth, I. Mehlhorn, Z. Huang, R. J. Fletterick, F. E. Cohen, and S. B. Prusiner, *Proc. Natl. Acad. Sci.* **90**, 10962 (1993).
13. S. B. Prusiner, D. C. Bolton, D. F. Groth, K. A. Bowman, S. P. Cochran, and M. P. McKinley, *Biochemistry* **21**, 6942 (1982).
14. H. Zhang, K. Kaneko, J. T. Nguyen, T. L. Livshits, M. A. Baldwin, F. E. Cohen, T. L. James, and S. B. Prusiner, *J. Mol. Biol.* **250**, 514 (1995).
15. K. Kaneko, H. Wille, I. Mehlhorn, H. Zhang, H. Ball, F. E. Cohen, M. A. Baldwin, and S. B. Prusiner, *J Mol. Biol.* **270**, 574 (1997).
16. T. L. James, H. Liu, N. B. Ulyanov, S. Farr-Jones, H. Zhang, D. G. Donne, K. Kaneko, D. Groth, I. Mehlhorn, S. B. Prusiner, and F. E. Cohen, *Proc. Natl. Acad. Sci.* **94**, 10086 (1997).
17. R. Riek, S. Hornemann, G. Wider, M. Billeter, R. Glockshuber, and K. Wüthrich, *Nature* **382**, 180-182 (1996).
18. M. Billeter, R. Riek, G. Wider, S. Hornemann, R. Glockshuber, and K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **94**, 7281 (1997).
19. D. G. Donne, J. H. Viles, D. Groth, I. Mehlhorn, T. L. James, F. E. Cohen, S. B. Prusiner, P. E. Wright, and H. J. Dyson, *Proc. Nat. Acad. Sci. USA* **94**, 13452 (1997).
20. R. Riek, S. Hornemann, G. Wider, R. Glockshuber, and K. Wüthrich, *FEBS Letters* **413**, 282 (1997).
21. D. R. Muhandiram, and L. E. Kay, *J. Magn. Reson B* **103**, 203 (1994).
22. S. Grzesiek and A. Bax, *J. Magn. Reson.* **96**, 432 (1992).
23. M. Wittekind, and L. Mueller, *J. Magn. Reson. B* **101**, 201 (1993).
24. S. Grzesiek, and A. Bax, *J. Am. Chem. Soc.* **114**, 6291 (1992).
25. S. Grzesiek, and A. Bax, *J. Biomol. NMR* **3**, 185 (1993).
26. T. M. Logan, E. T. Olejniczak, R. X. Xu, and S. W. Fesik, *J. Biomol. NMR* **3**, 225 (1993).
27. L. E. Kay, G.-Y. Xu, A. U. Singer, D. R. Muhandiram, and J. D. Forman-Kay, *J. Magn. Reson. B* **101**, 333 (1993).
28. D. Marion, P. C. Driscoll, L. E. Kay, P. T. Wingfield, A. Bax, A. M. Gronenborn, and G. M. Clore, *Biochemistry* **28**, 6150 (1989).

29. D. R. Muhandiram, N. A. Farrow, G. Y. Xu, S. H. Smallcombe, and L. E. Kay, *J. Magn. Reson. B* **102**, 317 (1993).
30. H. Kuboniwa, S. Grzesiek, F. Delaglio, and A. Bax, *J. Biomol. NMR* **4**, 871 (1994).
31. N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay, *Biochemistry* **33**, 5984 (1994).
32. A. S. Altieri, D. P. Hinton, and R. A. Byrd, *J. Am. Chem. Soc.* **117**, 7566 (1995).
33. Y. Bai, J. S. Milne, L. Mayne, and S. W. Englander, *Proteins* **17**, 75 (1993).
34. M. Nilges, M. J. Macias, S. I. O'Donoghue, and H. Oschkinat, *J. Mol. Biol.* **269**, 408 (1997).
35. G. W. Vuister, and A. Bax, *J. Am. Chem. Soc.* **115**, 7772 (1993).
36. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.* **117**, 5179 (1995).
37. H. Zhang, J. Stöckel, I. Mehlhorn, D. Groth, M. A. Baldwin, S. B. Prusiner, T. L. James, and F. E. Cohen, *Biochemistry* **36**, 3543 (1997).
38. D. S. Wishart, and B. D. Sykes, *Meth. Enzymol.* **239**, 363 (1994).
39. D. A. Case, H. J. Dyson, and P. E. Wright, *Meth. Enzymol.* **239**, 392 (1994).
40. R. A. Laskowski, J. A. C. Rullmann, M. W. MacArthur, R. Kaptein, and J. M. Thornton, *J. Biomol. NMR* **8**, 477 (1996).
41. W. Kabsch, and C. Sander, *Biopolymers* **22**, 2577 (1983).
42. A. M. Mandel, M. Akke, and A. G. Palmer III, *J. Mol. Biol.* **246**, 144 (1995).
43. K. Kaneko, L. Zulianello, M. Scott, C. M. Cooper, A. C. Wallace, T. L. James, F. E. Cohen, and S. B. Prusiner, *Proc. Natl. Acad. Sci.* **94**, 10069 (1997).
44. H. M. Schätzl, M. Da Costa, L. Taylor, F. E. Cohen, and S. B. Prusiner, *J. Mol. Biol.* **245**, 362 (1995).
45. D. A. Kocisko, S. A. Priola, G. J. Raymond, B. Chesebro, P. T. Lansbury, Jr., and B. Caughey, *Proc. Natl. Acad. Sci. USA* **92**, 3923 (1995).
46. J. A. Mastrianni, M. T. Curtis, J. C. Oberholtzer, M. M. Da Costa, S. DeArmond, S. B. Prusiner, and J. Y. Garbern, *Neurology* **45**, 2042 (1995).