

## Solution Structure of Syrian Hamster Prion Protein rPrP(90–231)<sup>†</sup>

He Liu,<sup>‡</sup> Shauna Farr-Jones,<sup>‡</sup> Nikolai B. Ulyanov,<sup>‡</sup> Manuel Llinas,<sup>§</sup> Susan Marqusee,<sup>§</sup> Darlene Groth,<sup>||</sup>  
Fred E. Cohen,<sup>‡,||,#</sup> Stanley B. Prusiner,<sup>||,△</sup> and Thomas L. James<sup>\*,‡</sup>

Departments of Pharmaceutical Chemistry, Neurology, Medicine, and Biochemistry and Biophysics, University of California, San Francisco, California 94143, and Department of Molecular and Cell Biology, University of California, Berkeley, California 94720

Received December 7, 1998; Revised Manuscript Received February 5, 1999

**ABSTRACT:** NMR has been used to refine the structure of Syrian hamster (SHa) prion protein rPrP(90–231), which is commensurate with the infectious protease-resistant core of the scrapie prion protein PrP<sup>Sc</sup>. The structure of rPrP(90–231), refolded to resemble the normal cellular isoform PrP<sup>C</sup> spectroscopically and immunologically, has been studied using multidimensional NMR; initial results were published [James et al. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 10086–10091]. We now report refinement with better definition revealing important structural and dynamic features which can be related to biological observations pertinent to prion diseases. Structure refinement was based on 2778 unambiguously assigned nuclear Overhauser effect (NOE) connectivities, 297 ambiguous NOE restraints, and 63 scalar coupling constants ( $^3J_{\text{H}^n\text{H}^m}$ ). The structure is represented by an ensemble of 25 best-scoring structures from 100 structures calculated using ARIA/X-PLOR and further refined with restrained molecular dynamics using the AMBER 4.1 force field with an explicit shell of water molecules. The rPrP(90–231) structure features a core domain (residues 125–228), with a backbone atomic root-mean-square deviation (RMSD) of 0.67 Å, consisting of three  $\alpha$ -helices (residues 144–154, 172–193, and 200–227) and two short antiparallel  $\beta$ -strands (residues 129–131 and 161–163). The N-terminus (residues 90–119) is largely unstructured despite some sparse and weak medium-range NOEs implying the existence of bends or turns. The transition region between the core domain and flexible N-terminus, i.e., residues 113–128, consists of hydrophobic residues or glycines and does not adopt any regular secondary structure in aqueous solution. There are about 30 medium- and long-range NOEs within this hydrophobic cluster, so it clearly manifests structure. Multiple discrete conformations are evident, implying the possible existence of one or more metastable states, which may feature in conversion of PrP<sup>C</sup> to PrP<sup>Sc</sup>. To obtain a more comprehensive picture of rPrP(90–231), dynamics have been studied using amide hydrogen–deuterium exchange and <sup>15</sup>N NMR relaxation times ( $T_1$  and  $T_2$ ) and <sup>15</sup>N{<sup>1</sup>H} NOE measurements. Comparison of the structure with previous reports suggests sequence-dependent features that may be reflected in a species barrier to prion disease transmission.

Prions are infectious particles that cause fatal neurodegenerative disorders in humans and animals that can be manifested as sporadic, inherited, and infectious illnesses (1). Prions are devoid of nucleic acid and consist largely, if not entirely, of the scrapie prion protein PrP<sup>Sc</sup> <sup>1</sup> in a conformation that forms insoluble aggregates. A wealth of data supports the concept that prion diseases arise through conformational

change as the normal cellular prion protein PrP<sup>C</sup> is converted into PrP<sup>Sc</sup> through a posttranslational process (1). The conversion of PrP<sup>C</sup> to PrP<sup>Sc</sup> is not accompanied by any apparent chemical modification (2); PrP<sup>C</sup> and PrP<sup>Sc</sup> seem to possess the same covalent structure but differ substantially in conformation. Spectroscopic studies indicate that PrP<sup>C</sup> is high in  $\alpha$ -helix content and nearly devoid of  $\beta$ -sheet, while PrP<sup>Sc</sup> has about 40%  $\beta$ -sheet (3, 4). Earlier studies showed that PrP27–30, which consists of 142 residues and is derived from PrP<sup>Sc</sup> by proteolytic N-terminal truncation under

<sup>†</sup>This work was supported by National Institutes of Health Grants AG10770, GM39247, and RR01081 as well as by gifts from the Leila and Harold G. Mathers, Sherman Fairchild, and Bernard Osher Foundations.

\* To whom correspondence should be addressed: tel, 415-476-1916; fax, 415-502-4690; e-mail, james@picasso.ucsf.edu.

<sup>‡</sup> Department of Pharmaceutical Chemistry, University of California, San Francisco.

<sup>§</sup> Department of Molecular and Cell Biology, University of California, Berkeley.

<sup>||</sup> Department of Neurology, University of California, San Francisco.

<sup>△</sup> Department of Medicine, University of California, San Francisco.

<sup>#</sup> Department of Cellular and Molecular Pharmacology, University of California, San Francisco.

<sup>△</sup> Department of Biochemistry and Biophysics, University of California, San Francisco.

<sup>1</sup> Abbreviations: PrP, prion protein; PrP<sup>Sc</sup>, scrapie PrP isoform; PrP<sup>C</sup>, cellular PrP isoform; PrP 27–30, protease-resistant core of PrP<sup>Sc</sup>; Mo, mouse; Hu, human; SHa, Syrian hamster; GPI, glycosylphosphatidylinositol; CD, circular dichroism; FTIR, Fourier transform infrared; 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; GdnHCl, guanidine hydrochloride; RMSD, atomic root-mean-square deviation; BSE, bovine spongiform encephalopathy; CJD, Creutzfeldt–Jacob disease; GSS, Gerstmann–Sträussler–Scheinker syndrome.

conditions that retain scrapie infectivity (5, 6), polymerizes into amyloid fibrils (7) and has a high  $\beta$ -sheet content (8).

Both PrP<sup>C</sup> and PrP<sup>Sc</sup> possess a disulfide bridge linking residues 179 and 214, a glycosylphosphatidylinositol (GPI) anchor at the C-terminus, and are glycosylated at Asn181 and Asn197 (9). Although the two isoforms of PrP have identical chemical properties, they possess very different biophysical properties. PrP<sup>C</sup> is soluble in nondenaturing detergents, while PrP<sup>Sc</sup> is not. PrP<sup>C</sup> is readily digested by proteases, while PrP<sup>Sc</sup> is partially protease-resistant (5).

Determination of the conformations of PrP<sup>C</sup> and PrP<sup>Sc</sup> is crucial for understanding the molecular basis of prion diseases, and in turn, it may provide the basis for structure-based drug design that prevents the conversion of PrP<sup>C</sup> to PrP<sup>Sc</sup>. Furthermore, observation of the structures that this sequence can adopt should help us to understand how proteins fold. To date, it has not been possible to obtain crystals of PrP<sup>C</sup> for structure determination using X-ray diffraction. Therefore, NMR is the most suitable technique for investigating the structures and dynamics of PrP<sup>C</sup>. The secondary structure and conformational transition of peptides consisting of residues 90–145 and 109–141, including the most conserved hydrophobic stretch 113–128, AGAAAA-GAVVGGGLGGY, were studied by NMR (10). The peptides could be induced to form  $\alpha$ -helical structures in aqueous solution in the presence of some organic solvents or detergents or to form intermolecular  $\beta$ -sheets when physiological salt was added. The aggregate of the  $\beta$  form is partially protease-resistant, while three other shorter peptides, without or with only a portion of the hydrophobic segment, display little or no tendency for  $\alpha \rightarrow \beta$  conversion and protease resistance. This suggests that the hydrophobic core 113–128 may be critical for the structural transition that occurs upon PrP<sup>Sc</sup> formation. On the basis of the peptide studies and the infectivity of the protease-resistant core PrP-27–30, we examined the structure and dynamics of PrP(90–231), which has the same primary sequence but resembles PrP<sup>C</sup> in secondary structure.

Recombinant PrP(90–231), corresponding to residues 90–231 of SHa PrP, was expressed in *Escherichia coli* in quantities sufficient for structural studies (11). The rPrP(90–231) has a well-structured core domain encompassing residues 125–228 (12). It also contains a hydrophobic cluster (residues 113–128) that is partially structured but does not adopt any regular secondary structure. The first 30 N-terminal residues (90–119) are highly flexible. The core domain of SHa rPrP(90–231) has a structure similar to that of the recombinant mouse prion protein MoPrP(121–231) (13, 14), i.e., consisting of three  $\alpha$ -helices and two short antiparallel  $\beta$ -strands. The primary difference between SHa PrP(90–231) and MoPrP(121–231) is that the former exhibits additional structure, e.g., the third helix (C-terminus) in SHa rPrP(90–231) is three turns longer than that in MoPrP(121–231), and residues 167–176, which are not defined in MoPrP(121–231), form a well-defined loop and an additional turn of helical structure at the beginning of the second helix. This region of additional structure corresponds to a discontinuous epitope for binding of a species-specific agent, provisionally designated protein X, which apparently functions as a molecular chaperone in PrP<sup>Sc</sup> formation (15).

Posttranslational cleavage of the 22-residue N-terminal signal peptide and a 23-residue C-terminal signal sequence

yields the functional protein of 209 residues (23–231). The structure and dynamics of nearly full-length, recombinant SHa PrP, designated rPrP(29–231), have also been studied and compared with rPrP(90–231) (16). 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 the octarepeats, displaying substantial flexibility and little structure. A study of full-length mouse rPrP(23–231) came to essentially the same conclusions (17). These studies argue that the N-terminus of PrP<sup>C</sup> may be disordered in vivo, perhaps assuming structure only in interacting with a functional partner. It is plausible that a ligand such as Cu(II) might bind to this region and induce a more well-ordered structure (18). Cu(II) binds to the octarepeat region (19, 20), and PrP-deficient mice were found to have diminished Cu levels in microsomes isolated from the brains of these mice (21). Due to the paramagnetic properties of Cu(II), no NMR data are available on the structure of the metalloprotein.

Here we report NMR studies of SHa rPrP(90–231), including assignment of NMR resonances and NOE cross-peaks, structure refinement with ARIA/X-PLOR and AMBER, backbone dynamics, and hydrogen–deuterium exchange of amide protons. The secondary structure elements and overall folding of the refined structure deviate little from the previously published low-resolution structure (12). However, with more restraints assigned, the backbone atomic root-mean-square deviation (RMSD) is much improved, from 1.64 to 0.67 Å, and structural details which may be pertinent to the PrP<sup>C</sup>  $\rightarrow$  PrP<sup>Sc</sup> transformation are revealed. Both the amide exchange data and the backbone relaxation parameter measurements are consistent with the conclusion that rPrP(90–231) consists of a stable core and a flexible N-terminus. The differences between SHa rPrP(90–231) and MoPrP(121–231) remain and instruct us about relative stability of different structural moieties and structural differences which may pertain to species specificity of prion diseases.

## MATERIALS AND METHODS

**Sample Preparation.** The expression, isotopic labeling, purification, and considerations for proper refolding of recombinant Syrian hamster prion protein rPrP(90–231) have been described previously (11, 22). Purified rPrP(90–231), labeled uniformly with <sup>15</sup>N or <sup>15</sup>N/<sup>13</sup>C, was lyophilized before refolding. The strategy of refolding is to stabilize the oxidized form of rPrP(90–231) at 1 mg/mL in 8 M GdnHCl and rapidly dilute into 20 mM Tris–acetate buffer, pH 8.0/5 mM EDTA to a final concentration of 0.1 mg/mL. The refolded protein was dialyzed against 20 mM sodium acetate, pH 5.2/0.005% sodium azide. Insoluble material was separated by centrifugation using a 0.2-mm filter. Solutions were concentrated by Centriplus 10 (Amicon) to a final protein concentration of 0.7–1.3 mM for NMR studies. All refolding and concentration were done in the presence of 20 mM methionine to prevent sulfoxide oxidation of methionine residues. Deuterated buffer exchange was done simultaneously with the final concentration step. Samples were analyzed by mass spectroscopy, circular dichroism, and Fourier transform infrared spectroscopy to ensure the refolded rPrP(90–231) resembles PrP<sup>C</sup>, i.e., has largely  $\alpha$ -helix content.

**NMR Spectra.** All NMR spectra, except for lower concentration diffusion experiments, were acquired at 30 °C with 1 mM uniformly  $^{15}\text{N}$ -labeled or  $^{15}\text{N}/^{13}\text{C}$ -labeled rPrP(90–231) at pH 5.2 in 10%  $\text{D}_2\text{O}/90\%$   $\text{H}_2\text{O}$ . Diffusion experiments were performed on the 1 mM sample as well as a 0.33 mM sample prepared as described above except that the solvent contained 99%  $\text{D}_2\text{O}$ . The following spectrometers were used: Bruker AMX500, DRX600, and RMX750 and Varian Unity Plus 600, each equipped with a 5-mm  $^{13}\text{C}/^{15}\text{N}/^1\text{H}$  triple resonance probe and pulsed field gradients (PFG). The gradient-enhanced method was used for all the experiments detecting amide protons during the acquisition period, allowing sensitivity-enhanced coherence selection and effective solvent suppression. Transmitter frequencies for  $^1\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}^\alpha$ , aliphatic  $^{13}\text{C}$ , and  $^{13}\text{CO}$  are typically 4.76, 118.4, 57.6, 42.0, and 177.1 ppm, respectively. The data were processed with either NMRPIPE (23) or a locally written graphical interactive NMR processing program, STRIKER (24).

Three-dimensional heteronuclear triple resonance experiments (25), HNCA (26), HNCACB (27), CBCA(CO)NH (28), HNCO (26), and HBHA(CO)NH (29), that correlate resonances from backbone HN,  $^{15}\text{N}$ ,  $\text{H}^\alpha$ ,  $^{13}\text{C}^\alpha$ , and  $^{13}\text{CO}$  and side chain  $\text{H}^\beta$  and  $^{13}\text{C}^\beta$  were acquired for backbone resonance assignments as described previously (12). HNCA, HNCACB, and CBCA(CO)NH spectra were acquired at 500 MHz. HNCO and HBHA(CO)NH spectra were acquired at 600 MHz.

For the assignment of side chain aliphatic  $^{13}\text{C}$  and  $^1\text{H}$  resonances, the 3D heteronuclear experiments C(CO)NH–TOCSY (30) and HCCH–TOCSY (31) were recorded at 600 MHz with isotropic mixing times of 11 and 22 ms. The C(CO)NH–TOCSY spectrum has 256 complex points for  $^1\text{H}$  ( $F_1$ ), 32 for  $^{15}\text{N}$  ( $F_2$ ), and 64 for  $^{13}\text{C}$  ( $F_3$ ), and HCCH–TOCSY has 256 for  $^1\text{H}$  ( $F_1$ ), 150 for  $^{13}\text{C}$  ( $F_2$ ), and 100 for  $^1\text{H}$  ( $F_3$ ). Both spectra were zero-filled to 1024, 256, and 256 in the three respective dimensions. An  $^{15}\text{N}$ -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 sodium 3-(trimethylsilyl)tetra-deuterio-propionate (TSP) at 0 ppm.

To obtain distance restraints for structure calculation, 100 ms  $^{15}\text{N}$ -edited NOESY–HSQC (32) and  $^{13}\text{C}$ -edited NOESY–HSQC (33) were acquired at 750 MHz for protons, with 512, 64, and 16 complex points for the  $^1\text{H}$  ( $F_1$ ),  $^{15}\text{N}/^{13}\text{C}$  ( $F_2$ ), and  $^1\text{H}$  ( $F_3$ ) dimensions; they were each zero-filled to 1024 or 2048, 512, and 512 for the three respective dimensions.  $J$  coupling constants were measured using the 3D HNHA experiment described by Kuboniwa et al. for backbone  $\phi$  torsion angle restraints (34).

**NMR Relaxation and Amide Exchange Experiments.** Spin–lattice relaxation time ( $T_1$ ), spin–spin relaxation time ( $T_2$ ), and heteronuclear  $^{15}\text{N}\{^1\text{H}\}$  values were obtained via pulse sequences previously described (35). Relaxation delays of 10, 20, 80, 160, 320, and 640 ms were used for  $T_1$  measurements and 15, 30, 61, 124, 170, and 216 ms for  $T_2$  measurements. Relaxation parameters were calculated by fitting peak heights measured with the locally written program SPARKY (36).

For the deuterium exchange studies, samples of  $^{15}\text{N}$ -labeled protein were prepared as described above, except that

deuterated buffer exchange was induced by spinning through a Quick-sep spin column (Isolab Inc.) packed with G-25 resin preequilibrated in deuterated buffer. A total of 17  $^{15}\text{N}$ -HSQC spectra were collected at 25 °C over a period of 2 weeks at 500 MHz. For each HSQC spectrum, four scans and 2048 points were collected in the direct dimension with 110 points in the indirect dimension. The data were processed using Felix 97.0. Peak intensities were fit using the Felix macro XPKvolheight written by Michael Akke. At the earliest time point measured, 48 amide proton signals could be detected. The observed exchange rates ( $k_{\text{obs}}$ ) were calculated by fitting peak intensities as a function of exchange time to a single exponential using Kaleidagraph (Synergy Software, Reading, PA). Protection factors ( $P$ ) were calculated as  $P = (k_{\text{rc}}/k_{\text{obs}})$ , where  $k_{\text{rc}}$  are the known intrinsic exchange rates for a residue in a specific tripeptide context (37).

**Self-Diffusion Constant Measurements.** NMR was utilized to measure the self-diffusion constant for rPrP(90–231) in  $\text{D}_2\text{O}$  buffer at 600 MHz using the pulsed field gradient method (38). The NMR probe had a maximum magnetic field gradient strength of 65 G/cm; calibration showed the gradient strength was linear up to 63 G/cm.

We determined the diffusion constant of a 1 mM sample of rPrP(90–231) at pH 5.2 (the same sample for some of the structural studies), a 0.1 mM sample obtained by diluting a 1 mM sample into  $\text{D}_2\text{O}$  buffer, and a sample at 0.33 mM which had never been concentrated to greater than 0.33 mM protein. Self-diffusion experiments were performed at 30 °C, the temperature of the NMR experiments, to determine the structure of rPrP(90–231). A recycle delay of 5 s was used between each transient, the gradient stabilization time was 0.5 ms, the PFG pulse duration was 7.5 ms, and the delay between PFG pulses was 17.5 ms. The peak height of the aromatic envelope of resonances was measured at 15 different gradient strengths from 10 to 45 G/cm. A plot of  $\ln(\text{amplitude})$  vs gradient strength squared was linear over that range for all three samples.

**Assignment of Nuclear Resonances.** The NMR spectral display and analysis program SPARKY, developed at UCSF, was used for assigning the NMR spectra. SPARKY is interfaced with the molecular graphics program MidasPlus (39) to make the display and examination of assignments in the multiple spectra and interim structure convenient. The program also does bookkeeping and data management for the assignments. (See <http://picasso.ucsf.edu/software.html> for information about SPARKY and other NMR-related software.) The backbone  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonances of rPrP(90–231) were assigned by establishing the connectivity of sequential backbone amides using the triple resonance experiments HNCACB and CBCA(CO)NH. The HNCACB experiment correlates the amide proton and nitrogen of a given residue with the intraresidue and preceding residue  $\text{C}^\alpha$  and  $\text{C}^\beta$  resonances, while the CBCA(CO)NH spectrum provides correlations to the  $\text{C}^\alpha$  and  $\text{C}^\beta$  resonances of the preceding residue. HNCA has better sensitivity than HNCACB and is used to build the backbone connectivities when sensitivity loss in HNCACB makes assignments difficult. With the backbone amide protons and nitrogens assigned, the  $^{13}\text{C}$  resonance of CO can be assigned readily using the HNCO experiment, as well the  $\text{H}^\alpha$  and  $\text{H}^\beta$  protons using HBHA(CO)NH. Assignment of these latter two spectra also

provides confirmation for the other backbone resonance assignments.

The C(CO)NH–TOCSY spectrum permits correlation of the backbone amides to  $^{13}\text{C}$ 's along the side chains, allowing the assignment of  $\text{C}'\text{'s}$ . The  $^1\text{H}$  and  $^{13}\text{C}$  resonance assignments of aliphatic side chains were completed with HCCH–TOCSY, which provides correlations among the  $^{13}\text{C}$ 's of the side chains 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 the correlating the side chain amide proton and nitrogen with intraresidue  $\text{C}^\alpha$  and  $\text{C}^\beta$ 's for Asn and  $\text{C}^\beta$ 's and  $\text{C}'\text{'s}$  for Glu. The  $^{15}\text{N}$  NOESY–HSQC spectrum assisted with assignment of partially overlapped amino groups, using cross-peaks to  $\text{H}^\beta$  protons or  $\text{H}'\text{'s}$  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 on the basis of  $^{13}\text{C}$ -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  $\text{H}^\beta$  due to two-bond  $\text{H}^\beta\text{--C}^\beta\text{--C}'^2$  coupling. The  $\delta^1$  methyl has a cross-peak to the  $\gamma^1$  methylene due to two-bond  $\text{H}'^1\text{--C}'^1\text{--C}^{\delta^1}$   $J$  coupling.

*Assignment of NMR Structural Restraints.* Assignments of NOE cross-peaks were also made using SPARKY. The program provides tools that 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 carrying alternative assignments with annotations during the long process of making assignments.

A total of 1895 NOE cross-peaks were unambiguously assigned in  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOESY spectra. Ambiguous assignments with the degeneracy less than 3 were made for 117 additional peaks. These ambiguous restraints can be used in some structure refinement programs such as X-PLOR. Additional NOE restraints were assigned using ARIA during structure refinement (40) (vide infra). The total number of unambiguous NOE restraints after ARIA refinement was 2778, and the number of ambiguous restraints was 297.

NOE intensities (peak heights) were scaled on the basis of the intensities of  $^{15}\text{N}$  or  $^{13}\text{C}$  2D HSQC spectra to take into account the fact that intensity differences in 2D HSQC (due to different backbone or side chain dynamics, different proton exchange rates, or  $^1\text{H}\text{--}^{15}\text{N}/^{13}\text{C}$  couplings) are manifest in the NOE cross-peaks. The properly scaled NOE intensities were classified as strong, medium, weak, or very weak, respectively, corresponding to distance restraints of 1.7–2.8, 2.0–3.5, 3.0–5.0, or 3.5–6.0 Å. The distance calibration is based on NOE peaks that are assigned to the approximately known distances [such as  $d_{\alpha\text{N}}(i,i+3)$ ] in the  $\alpha$ -helices identified by chemical shift indices and preliminary structure calculation; ends of the helices were not included. The average distances and volumes were used in the calibration. 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.

To provide information about the dihedral backbone angle  $\phi$ , three-bond  $^3J_{\text{H}^{\text{N}}\text{H}^\alpha}$  scalar coupling constants were measured using the 3D HNHA spectrum that correlates intraresidue  $\text{H}^{\text{N}}$ ,  $\text{N}$ , and  $\text{H}^\alpha$  (41). Since dephasing of the  $\text{H}^{\text{N}}$  magnetization is dependent on the homonuclear  $\text{H}^{\text{N}}\text{--H}^\alpha$  scalar coupling, the ratio of diagonal to cross-peak intensities in the HNHA spectrum directly relates to the coupling constant  $^3J_{\text{H}^{\text{N}}\text{H}^\alpha}$ . Due to different relaxation effects on the in-phase diagonal and antiphase cross-peaks, the line width and line shape of diagonal peaks and cross-peaks were different. A relatively large error of  $\pm 15\%$  was added to the intensities (peak heights) to compensate for this uncertainty, resulting in  $J$  coupling constants with error bounds ranging from  $\pm 0.44$  to 0.89 Hz. Sixty-three  $J$  coupling constant restraints were used directly in structure refinement. The coefficients used by AMBER in the Karplus equation are 6.40,  $-1.40$ , and 1.90.

Forty-one hydrogen-bonding restraints were identified for the helices. Hydrogen bonding was determined on the basis of NOEs between the amide proton and  $\alpha$  protons at residues  $i-3$  and  $i-4$  [ $d_{\alpha\text{N}}(i,i+3)$  and  $d_{\alpha\text{N}}(i,i+4)$ ], an NOE between an  $\alpha$  proton and residue  $(i+3)$   $\beta$  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  $\beta$ -sheet. These hydrogen bonds were determined on the basis of the observed interstrand NOE connectivity patterns. Some of the more stable 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,  $d_{\text{HN--O}} = 1.5\text{--}2.3$  Å and  $d_{\text{N--O}} = 2.5\text{--}3.3$  Å, were used for each hydrogen bond.

*ARIA/X-PLOR Calculations.* The 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 (40), 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. The same ARIA/X-PLOR procedure and simulated annealing protocol previously described (40) were used to calculate the structure of rPrP(90–231). Ambiguous as well as unambiguous restraints assigned manually (in contrast to ARIA-assigned 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, which utilized additional restraints from ambiguous NOE peaks. The interface of ARIA routines with the X-PLOR simulated annealing calculation makes it possible to assign ambiguous NOE restraints automatically on the basis of 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. All NOE cross-peaks in the spectra (assigned and unassigned) were used with ARIA. NOE cross-peaks in  $^{15}\text{N}$ -NOESY-HSQC and  $^{13}\text{C}$ -NOESY-HSQC spectra were assigned independently. The program merges the restraints assigned for different spectra and removes duplicated assignments. With the decreasing level of ambiguity in the ARIA/X-PLOR iterative procedure, more restraints could be assigned unambiguously, resulting

in rapid convergence of the structure. In all structure calculations, 63 backbone  $J$  coupling constants and 44 hydrogen bond restraints were used.

To test how the ARIA/X-PLOR procedure depended on the initial restraint set, we used two sets of initial NOE restraints for structure calculations. The first set consisted of all 1895 unambiguous and 117 ambiguous restraints. In the second set, 30% of the restraints were randomly removed from the first set, resulting in 1342 unambiguous and 80 ambiguous restraints. The two sets of restraints were used to generate initial structures from an extended polypeptide chain, and eight iterations with parameters similar to those advocated in Nilges' ARIA manual were carried out to complete the ARIA/X-PLOR procedure. The calculation using 100% of the manually assigned restraints generated 10 structures with an RMSD of 0.79 Å, 2789 unambiguous and 324 ambiguous restraints. The calculation using 70% of the manually assigned restraints generated structures with an RMSD of 0.82 Å, 2618 unambiguous and 312 ambiguous restraints. The average RMSD between the two sets of structures calculated with 100% and 70% manually assigned restraints is 1.49 Å. We also tested the possible bias associated with manually assigned restraints by excluding weak NOE peaks assigned manually, resulting in 1533 initial restraints. Eliminating the weak manually assigned restraints generated 10 structures with a 0.75 Å RMSD, a total of 2889 unambiguous and 307 ambiguous restraints. The RMSD between this set of structures and that calculated with 100% manually assigned restraints is 1.69 Å. We found that the results of the ARIA/X-PLOR calculations depend on the initial restraints. It is necessary to assign manually as many peaks as possible to minimize bias and to have a good system to calibrate NOE intensities. Even though the RMSD was improved in ARIA/X-PLOR calculations when weak NOEs were removed, we included them in the final structure calculation because they were unambiguous and agreed with chemical shift data and secondary structure. The structures reported here from ARIA/X-PLOR calculations are based on the initial 1895 unambiguous manually assigned restraints. Following ARIA, 2778 unambiguous and 297 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 on the basis of total energies.

**Restrained Molecular Dynamics Calculations in Water.** 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 (42). Only unambiguous restraints (assigned manually or by ARIA) were used. The restrained molecular dynamics calculations were performed using the SANDER module of Amber 4.1 (43). The unstructured N-terminal residues 90–111 were removed before solvation of the protein with the 6 Å water shell. A distance-dependent dielectric function was used for all AMBER calculations. Five hundred steps of belly water energy minimization and another 500 steps of energy minimization of all atoms were carried out prior to restrained molecular dynamics calculations. The structures were heated to 600 K over 5 ps and kept at this temperature for 10 ps. The system was cooled gradually to 0 K over 10 ps. The total length of the calculation was 25 ps. The restraint force

constant was increased linearly from 0 to 20 kcal/(mol·Å<sup>2</sup>) over 10 ps and maintained at this level until the end of the calculation. Separate scaling factors were used for the protein and solvent molecules. The temperature coupling constant of the solute was increased gradually from 0.04 to the maximum of 0.4 over 15 ps. At the end of the MD run, the temperature was 0 K, and energy was minimized using an NMR restraint force constant of 10 kcal/(mol·Å<sup>2</sup>). Water molecules were stripped off and energy terms were calculated for the protein using AMBER. The nonbonded interaction cutoff was 10 Å for the restrained MD runs and 8 Å for the energy minimizations. The refined structures were deposited in the Protein Data Bank, PDB ID code 1b10, replacing the previously deposited preliminary refined structures (12), PDB ID code 2prp.

## RESULTS

### *Resonance Assignments and Secondary Structure*

**NMR Resonance Assignments.** Complete sequential assignment of the backbone <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N resonances of rPrP(90–231) were obtained by establishing intraresidue and sequential connectivities of N, H<sup>N</sup>, C<sup>α</sup>, and C<sup>β</sup> resonances using the 3D NMR triple-resonance experiments HNCA, HNCACB, and CBCA(CO)HN. The assignments were confirmed with the 100 ms NOESY spectrum of the <sup>15</sup>N-labeled sample, with the exception of the first two N-terminal residues G90 and Q91. Figure 1 shows a 2D <sup>15</sup>N-HSQC spectrum acquired at 750 MHz, showing the chemical shift dispersion of backbone amides as well as side chain amides and amino groups. The side chain arginine amide peaks in the upper right corner are folded with  $\omega_1' = \omega_1 - 24$  ppm, and tryptophan indole peaks in the upper left corner are folded with  $\omega_1' = \omega_1 + 24$  ppm. Signals emanating from backbone amides are all resolved except for overlap in the following nine cases: G90/G94, G93/G114, H96/I184, W99/E146, K101/V189, M109/Y149, D167/Q168, H199/Q186, and I205/Y226. The side chain amino group region is very crowded in the <sup>15</sup>N-HSQC spectrum, but using the same set of triple resonance experiments as for the backbone assignments and the <sup>15</sup>N-edited NOESY spectrum, 11 out of 13 asparagine side chain amino groups can be assigned. The NH<sub>2</sub> resonances of N155 and N170 are overlapped with other NH<sub>2</sub> resonances. The amino group resonances, indicated by horizontal lines, are split due to partial deuteration. The side chain amides of ten glutamines and guanidino groups of eight arginines were all assigned.

Nearly all aliphatic side chain <sup>1</sup>H and <sup>13</sup>C resonances, except for P137  $\gamma$ , P165  $\gamma$  and Q217  $\gamma$ , were assigned using HBHA(CO)NH, C(CO)NH, and HCCH–TOCSY spectra. Aromatic side chain resonance assignments of F141, F198, Y128, Y157, and Y162 were made with TOCSY experimentally optimized for aromatic <sup>1</sup>H–<sup>13</sup>C couplings. Prochiral groups, i.e., methylenes and isopropyl groups, were assigned stereospecifically via a floating chirality method performed during structure calculations with ARIA/X-PLOR (44).

**Secondary Structures and Stability.** The secondary structure of rPrP(90–231) was first identified using the deviation of chemical shifts from random coil values, i.e., the chemical shift index (CSI) (45, 46). The C<sup>α</sup>, CO, and H<sup>α</sup> chemical shift deviations of rPrP(90–231) from random coil values are shown in Figure 2. The first 30 residues of the N-terminus

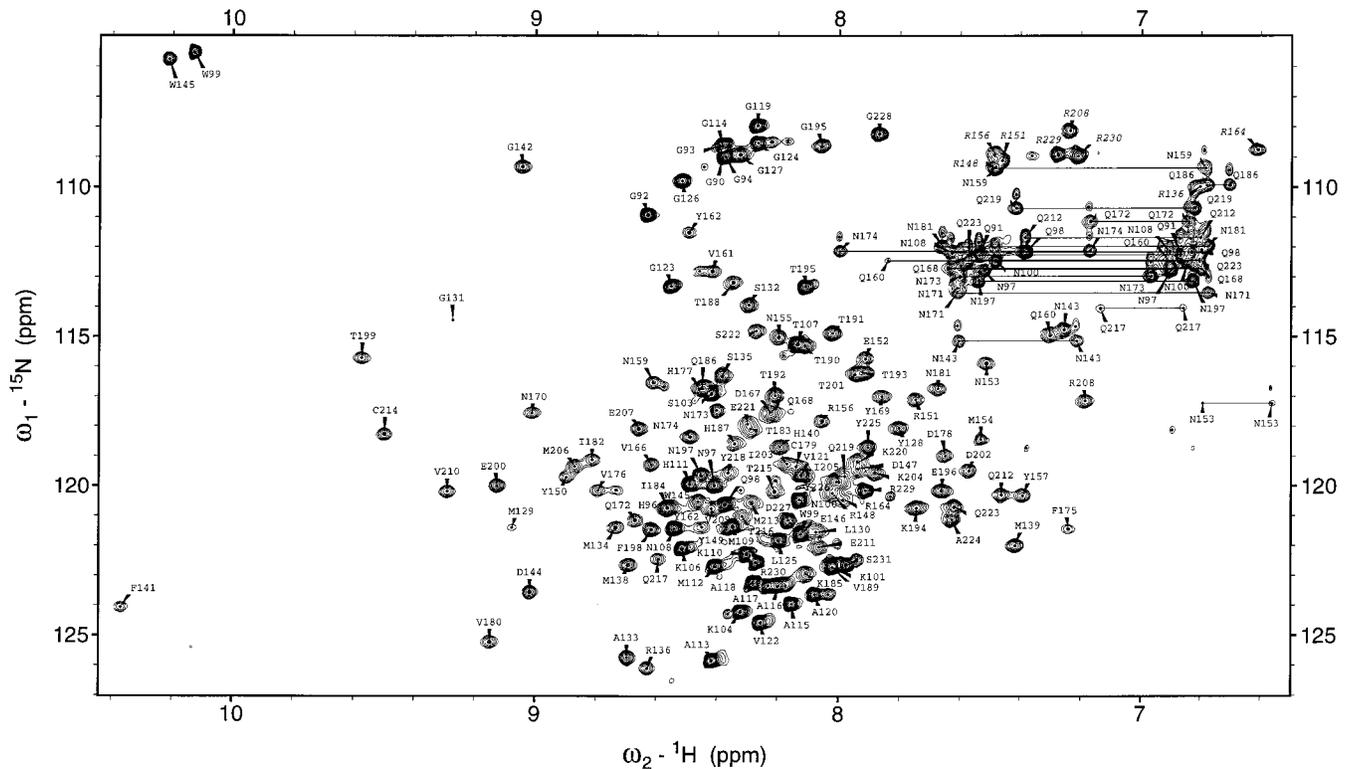


FIGURE 1: Annotated two-dimensional  $^{15}\text{N}$ -HSQC spectrum acquired at 750 MHz. Backbone amide signals are all resolved except for G90/G94, G93/G114, H96/I184, W99/E146, K101/V189, M109/Y149, D167/Q168, H199/Q186, and I205/Y226. The side chain arginine amide peaks in the upper right corner are folded with  $\omega_1' = \omega_1 - 24$  ppm, and tryptophan indole peaks in the upper left corner are folded with  $\omega_1' = \omega_1 + 24$  ppm. The amino group resonances of glutamines and asparagines, indicated by horizontal lines, are split due to partial deuteration. This figure, as well as Figures 4 and 9, was prepared using the program SPARKY (36).

exhibit only small chemical shift deviations from random coil and manifest no systematic shifts except for  $\text{C}^\alpha$  resonances. This region also has weak NOEs with few medium- and long-range NOEs, as well as narrow NMR line widths. This collectively 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. This is indicated by some medium and weak long-range NOEs identified for the regions of residues 96–101 and 112–128.

Three  $\alpha$ -helices were indicated by sequential positive deviations of  $\text{C}^\alpha$  and CO chemical shifts and negative deviations of  $\text{H}^\alpha$  chemical shifts. The three helices span residues 144–156, 172–194, and 200–227, i.e., helices A, B, and C, respectively. The helical structures indicated by CSI are supported by  $d_{\text{NN}}(i, i+1)$ ,  $d_{\alpha\text{N}}(i, i+3)$ ,  $d_{\alpha\beta}(i, i+3)$ , and  $d_{\alpha\text{N}}(i, i+4)$  connectivities. Consensus CSI and NOE connectivities are in agreement (12). The refined structure of rPrP-(90–231) contains three helices that span essentially the same residues as predicted by CSI. Details are discussed in the Structure Description and Analysis section.

On the basis of CSI analysis, two short  $\beta$ -strands may exist. Residues 159–164 show  $\beta$  character consistently in all  $\text{C}^\alpha$ , CO, and  $\text{H}^\alpha$  chemical shift data, but the antiparallel strand, i.e., residues 128–133, manifests much weaker  $\beta$  character with the shift to  $\beta$  being interrupted and inconsistent. According to  $\text{C}^\alpha$  chemical shifts, only residues 129 and 130 shifted to  $\beta$ , but CO chemical shifts suggest that the  $\beta$ -strand may span residues 127–134, while  $\text{H}^\alpha$  CSI values indicate that residues 129–132 are shifted to  $\beta$ . The analysis of NOE cross-peaks shows that hydrogen bonds exist

between M129 O and Y163  $\text{H}^\text{N}$ , between M129  $\text{H}^\text{N}$  and Y163 O, and between V161 O and G131  $\text{H}^\text{N}$ . This indicates that the two  $\beta$ -strands form an antiparallel  $\beta$ -sheet. In the structures obtained using NMR restraints and structure refinement (vide infra),  $\beta$ -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  $\beta$ -sheet characteristics.

Amide hydrogen–deuterium exchange experiments were carried out to determine the regions of rPrP(90–231) most resistant to exchange. All secondary structural elements in rPrP(90–231) contain amide protons which are protected from exchange with deuterium. Figure 3 displays protection factors mapped as colored spheres for protected residues onto a ribbon diagram of the molecule, with the sites divided into three groups according to protection factors. Residues not shown with spheres are those where amide protons exchanged too fast to be measured, i.e., are not detectable in the  $^{15}\text{N}$ -HSQC experiment after 20 min of exchange with deuterium oxide. Helix A, residues 144–154, exhibits intermediate protection. While N155 H is not H-bonded to E152 O in helix A (confirmed by NOESY spectra), the proton exchange data suggest that these residues are protected, which is consistent with the presence of a turn as determined from the hydrogen bonding between N153 O and R156 H and between N153 O/M154 O and Y157 H.

The amide protons of residues 175–186 in helix B are well protected in the proton–deuterium exchange experiment. Lack of protection through to the C-terminal end of this helix (residue 194) suggests that this portion of the helix is more dynamic, permitting amide proton exchange with

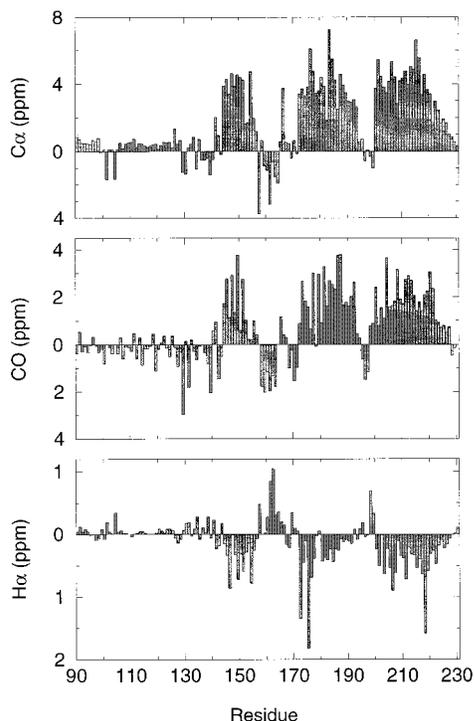


FIGURE 2: Chemical shift deviation of  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}=\text{O}$ , and  $^1\text{H}\alpha$  resonances relative to chemical shifts of random coil polypeptides. The three  $\alpha$ -helices encompassing residues 144–156, 172–194, and 200–227 are indicated by successive positive deviations of  $^{13}\text{C}\alpha$  and  $^{13}\text{C}=\text{O}$  chemical shifts and negative deviations of  $^1\text{H}\alpha$  chemical shifts. Residues 159–164 show  $\beta$  character consistently in all  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}=\text{O}$ , and  $^1\text{H}\alpha$  resonances, but the antiparallel strand, i.e., residues 128–133, exhibits much weaker  $\beta$  character. The first 30 residues of the N-terminus show small chemical shift deviations from random coil.

solvent. Due to repetitive threonines with overlapping  $\text{H}\alpha$ 's, only five  $d_{\alpha\text{N}}(i,i+3)$ ,  $d_{\alpha\beta}(i,i+3)$ , or  $d_{\alpha\text{N}}(i,i+4)$  connectivities in this region support the existence of helical hydrogen bonding for the amides of H187, T188, T191, and T192. Simulated annealing calculations using X-PLOR with these medium-range NOE restraints result in a helix for residues 187–194 in some but not all trials. The limited stability of helical structure for residues 187–194 is also evident in the strong solvent exchange peaks for amides in the  $^{15}\text{N}$ -edited NOESY–HSQC spectrum. However, residues 187–194 do consistently form a helix in structure calculations using ARIA (40), which assigned restraints in this region that are ambiguous due to overlap.

The H–D exchange data indicate that residues 199–221 in helix C are also well protected; 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  $d_{\alpha\text{N}}(i,i+3)$ ,  $d_{\alpha\beta}(i,i+3)$ , and  $d_{\alpha\text{N}}(i,i+4)$  connectivities as well as the chemical shift indices. Long-range NOEs from V166 to Y218, E221, S222, and Y225 (12) and from Q172 to Y215 and Q219 also indicate that the region 222–227 is indeed helical. Figure 4 shows cross-peaks in the  $^{13}\text{C}$ -edited NOESY spectrum supporting the existence of hydrogen bonds which were not found in MoPrP(121–231) (13, 14), i.e., K220 O to A224  $\text{H}^{\text{N}}$  and E221 O to Y225  $\text{H}^{\text{N}}$ . Figure 4a shows the pair of NOE cross-peaks between E221  $\text{H}^{\alpha}$  and A224  $\text{H}^{\beta}$ . In Figure 4b, the cross-peak from D227  $\text{H}^{\beta}$  to A224  $\text{H}^{\alpha}$  and the symmetric peaks from A224  $\text{H}^{\alpha}$  D227  $\text{H}^{\beta}$  are shown.

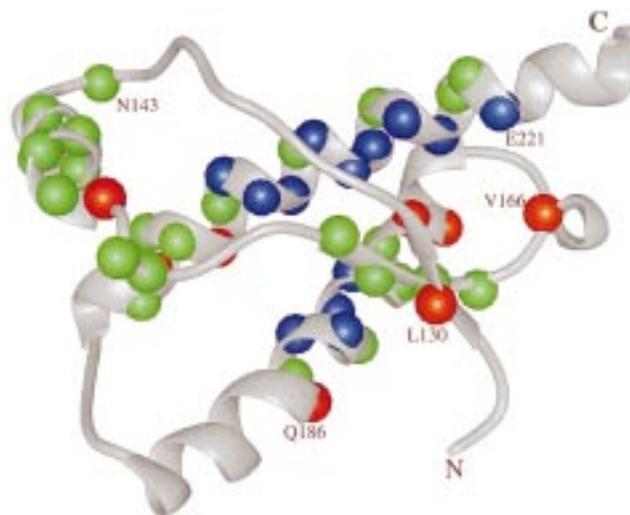


FIGURE 3: Hydrogen–deuterium exchange protection factors shown with the structure of rPrP(90–231). Residues with amide exchange protection factors larger than 4650, meaning well protected, are shown with blue spheres; green spheres represent residues with protection factors less than 4650 but greater than 370; and the red spheres have protection factors below 370. Residues shown without colored spheres are those not detectable in the  $^{15}\text{N}$ -HSQC experiment after 20 min of exchange with deuterium. This figure, as well as Figures 5–9 and 13, was prepared using the Program MidasPlus (39).

In the  $\beta$ -sheet region of rPrP(90–231), the H–D exchange data also show that strand 2 (S2) of the  $\beta$ -sheet is more protected than strand 1 (S1). Residues 161–164 show intermediate protection (green spheres in Figure 3), whereas only the amide of L130 is sufficiently stable in S1 (red sphere in Figure 3). It is interesting that the amide in loop residue V166 is protected; the side chain of V166 interacts extensively with helix C (12), and, in 4 of the 25 refined structures, the amide proton of V166 forms a H-bond with the side chain carboxylate of E221, which could certainly explain why its protection factor is rather good considering it is in a loop (vide infra).

### Three-Dimensional Structure

**NMR Structural Restraints.** NOE restraints for the structure calculation were obtained using 100 ms  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOESY–HSQC spectra. A total of 1895 unambiguous restraints (654 intraresidue, 622 sequential, and 382 medium range and 237 restraints separated by more than four residues) were assigned manually, i.e., using graphical display to make assignments based on resonance positions and secondary structures assumed from chemical shift indices. Another 117 ambiguous restraints, with two or more possible assignments, were also obtained this way. Additional restraints were assigned during structure calculations using ARIA/X-PLOR, an interface of iterative assignment of ambiguous restraints and simulated annealing structure calculations. The total number of unambiguous restraints at the end of the ARIA/X-PLOR iteration was 2778 (1146 intraresidue, 738 sequential, 541 medium range, and 353 long range). The final number of ambiguous restraints was 297. Sixty-three three-bond coupling constants  $^3J_{\text{H}^{\text{N}}\text{H}^{\alpha}}$ , determined from well-resolved diagonal and off-diagonal peaks in the

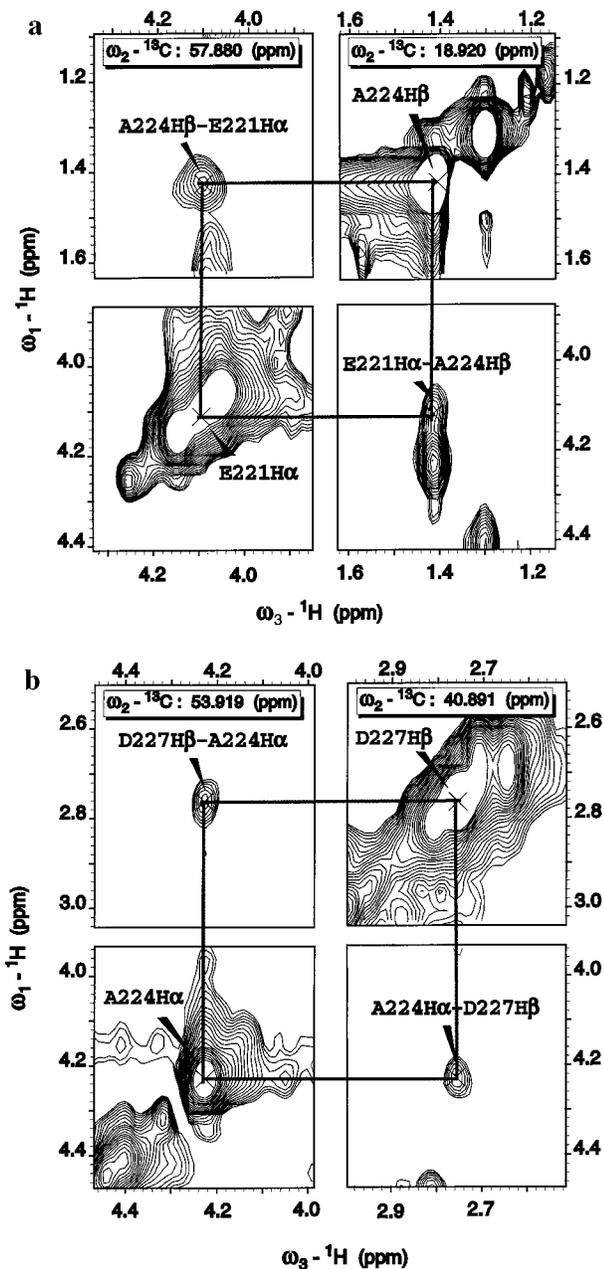


FIGURE 4: NOE cross-peaks in the  $^{13}\text{C}$ -edited NOESY spectrum. (a) Cross-peak between E221 H $\alpha$  and A224 H $\beta$ . (b) Cross-peaks between D227 H $\beta$  and A224 H $\alpha$ . Both are shown with the symmetric mirror peaks in the same  $^{13}\text{C}$ -edited NOESY-HSQC spectrum.

HNHA spectrum, were used directly in ARIA/X-PLOR and AMBER calculations as restraints for backbone torsion angle  $\phi$ . Hydrogen bond restraints were identified on the basis of CSI, medium-range NOEs, and hydrogen-deuterium exchange data. Seven H-bonding restraints of helix structure were used for residues 144–154, 14 for residues 172–189, and 20 for residues 205–225. Three H-bonds were found for the antiparallel  $\beta$ -sheet. A total of 44 H-bonding restraints were used in the structure calculations.

**Structure Description and Analysis.** One hundred structures were generated by ARIA/X-PLOR using a simulated annealing protocol for NMR-restrained structure refinement. The 25 structures having the lowest energy were subjected to 25 ps of restrained molecular dynamics simulation using AMBER with a 6.0 Å water shell around the protein and subsequently energy-minimized. All 25 refined structures

were included to construct an ensemble of rPrP(90–231) structures. The structures were superimposed by minimizing the RMSD of backbone CO, C $\alpha$ , and N atoms of residues 125–227 to the average structure. A stereoview of the superimposed structures, with an RMSD of 0.67 Å for backbone atoms, is shown in Figure 5. The 25 structures were assessed with PROCHECK (47), and the structure with the lowest AMBER energy and good structural parameters is shown in Figure 6.

The rPrP(90–231) structure has a core region consisting of helices A, B, and C, with helices B and C being further stabilized by the disulfide bond between residues C179 and C214 (colored yellow in Figure 6). Helix A includes residues 144–153 with helix boundary residues N143 as the N-terminal cap and M154 as the C-terminal cap (48). At the C-terminus of helix A, a hydrogen-bonded turn extends to residue Y157. This is consistent in all 25 structures; except for one case, the continued turn is identified as a  $3_{10}$  helix using the Kabsch and Sander algorithm (49). Helix B spans residues 172–193. The N-terminus of helix B is capped by N171, but the C-terminus continues either as a helix–turn (7 of 25 structures) or as a hydrogen-bonded turn (18 of 25) to residue K194 and is C-terminally capped by G195. As discussed above, the helical structure for residues 187–194 either is not stable or exists only part of the time. 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, capped N-terminally by T199 and C-terminally by G228. The extent of helix C is consistent among the 25 structures, and it is the most stable helix in rPrP(90–231).

Strands S1 and S2, shown with green and blue arrows, respectively, in Figure 6, comprise a small  $\beta$ -sheet that is three residues long for each strand, i.e., spans residues 129–131 and 161–163. It is likely that extended  $\beta$ -strands may exist at least transiently. Figure 7 shows that in the refined structures residues 159–164 form a good  $\beta$  structure, but the C-terminus of the antiparallel strand 128–134 is distorted, which agrees well with CSI and hydrogen-deuterium exchange data. There are extensive cross-strand NOE connectivities between the extended strands (colored green and blue in Figure 6), but there is no obvious NOE cross-peak pattern to support a longer regular antiparallel  $\beta$ -sheet. In the calculated structures, G131 O and V161 H $\text{N}$  are not involved in hydrogen bonding, and no other hydrogen bond forms for this elongated region except for M134 H $\text{N}$ –Q159 O. In the family of refined structures, residue G131 exists in two conformations. In one conformation C=O points toward V161, and in the other N–H points to V161. The flexibility of S1 is supported by all our data. The G131 amide resonance exhibits an exceptionally broad line width and is very weak in HQSC and other spectra. This implies the existence of multiple conformations with relatively slow exchange. When G131 O is pointing away from V161, S132 O is almost 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

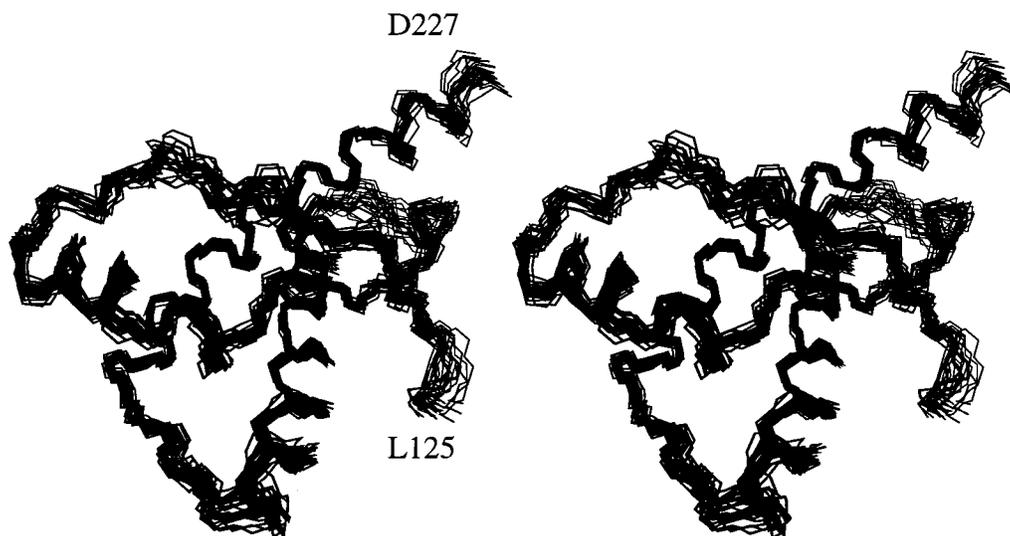


FIGURE 5: Stereoview of 25 structures obtained with AMBER restrained molecular dynamics simulation in a 6.0 Å water shell and energy minimization. The structures were superimposed by minimizing the atomic RMSD of backbone CO, C $\alpha$ , and N atoms of residues 125–227 to the average structure. The average backbone RMSD of the 25 structures is 0.67 Å.

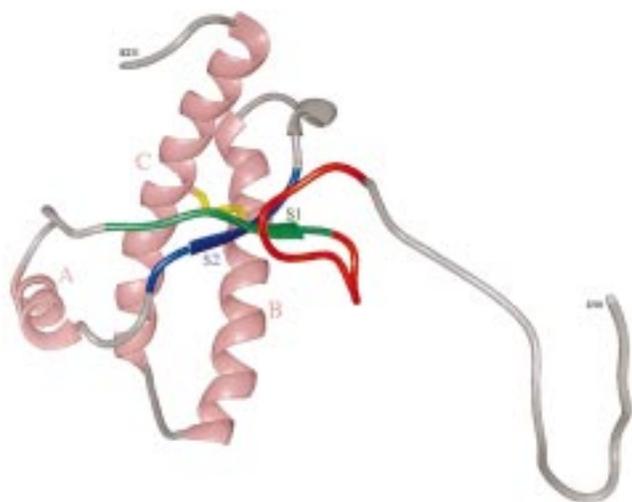


FIGURE 6: Ribbon representation of a rPrP(90–231) structure. The three helices, A, B, and C, are colored magenta. Helices B and C are linked by a disulfide bond (yellow) formed between residues C179 and C214. Strands S1 and S2 of the  $\beta$ -sheet are shown as green and blue arrows, respectively. Residues participating in the extended antiparallel interstrand interactions of the  $\beta$ -strands are 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.

evidence for more than one conformation for V161 in the  $^{15}\text{N}$  HSQC, which has two peaks for this amide (Figure 1).

The  $\beta$ -sheet is located close to the disulfide bond, and S2 interacts with helix B as well as helix C. The amide of Y162 is hydrogen-bonded to O $\gamma$  of T183, and the aromatic ring of Y162 is close to the methyl groups of I182, resulting in an upfield shift of the methyl resonance by 0.4 and 0.6 ppm (relative to random coil values) for  $\delta^1$  and  $\gamma^2$ , respectively. The two methyl groups of L130 are above the plane of the Y162 ring and are also shifted upfield by 0.4 and 0.8 ppm.

The segment between S2 and helix B, i.e., residues 165–169, exhibits a turn which is recognized by the Kabsch and Sander algorithm as either a  $3_{10}$  helix–turn (18 of 25 structures) or a hydrogen-bonded turn (7 of 25 structures).

Four residues, i.e., Q168 in this turn, Q172 at the beginning of helix B, and T215 and Q219 on helix C, were identified in studies using mouse neuroblastoma cells transfected with mutated chimeric Hu/MoPrP genes as forming an epitope which interacts with a molecular chaperone, tentatively named protein X, assisting in conversion of cellular PrP<sup>C</sup> to the scrapie form, PrP<sup>Sc</sup> (15). Q168, as part of the helical turn, is not in close contact with the other three residues that appear to be involved in the protein X binding (Figure 8). In the low-resolution structures published previously (12), fewer restraints for residues 165–171 resulted in this region lacking the definition we have here and led to a conformational envelope permitting Q168 to be close to both Q172 and Q219. However, such a conformation cannot be achieved without severe violation of the NOE restraints between the side chain methyl groups of V166 and residues Y218, E221, S222, and Y225 (purple in Figure 8). In our present refined structure, loop 165–171 is well defined, with all NMR restraints between the loop and helix C satisfied, as well as those within the loop. However, it would be quite feasible for Q168 to be moved into close proximity to Q172, T215, and Q219 upon binding to protein X.

We might speculate that such a conformational change induced by protein X binding may transduce a conformational switch possibly involving some aspects of all structural elements identified in rPrP(90–231) in promoting the PrP<sup>C</sup>  $\rightarrow$  PrP<sup>Sc</sup> transformation.

Segment 113–128 (red in Figure 6) consists largely of glycines and hydrophobic residues and apparently lacks regular secondary structure; however, there is evidence of structure in this region (Figure 9). Weak medium- and long-range NOEs were found for this region, but many of them are ambiguous. About 30 medium- ( $i$  to  $i+2$ , and  $i$  to  $i+3$ ) or long-range NOEs could be assigned and used for this hydrophobic region, in structure calculations. Residues 113–115 are largely disordered; residues 116–118 are also disordered, but weak or transient interactions involving the methyl groups of V121 or V122 are evident in NOESY–HSQC spectra for this region resulting in a pronounced bend or turn. There are two or three bends or turns for this

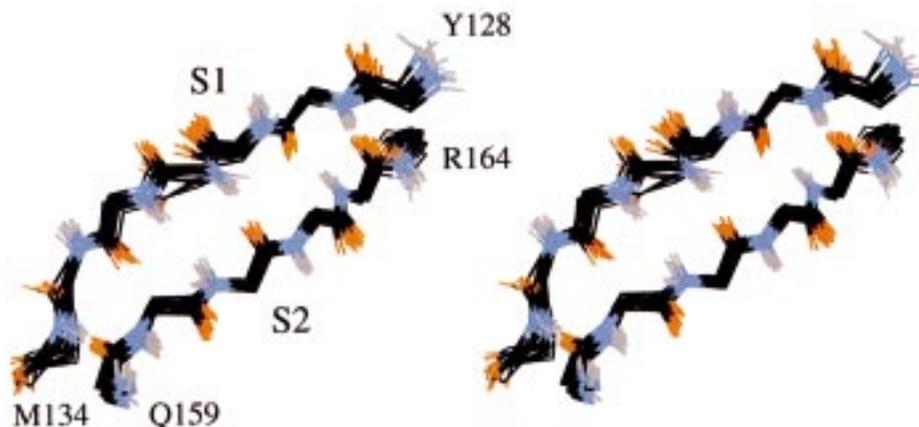


FIGURE 7: Backbone stereoview of the extended  $\beta$ -strands showing that the distortion from standard  $\beta$  structure is largely due to the flexibility of S1. Protons are colored gray, nitrogen is in blue, carbon is in black, and oxygen is in red.

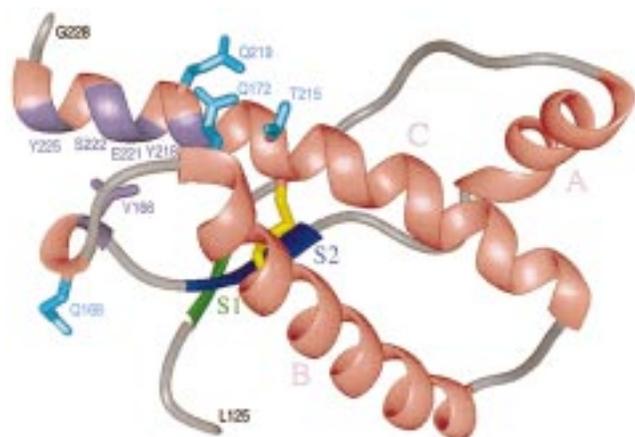


FIGURE 8: View of the interaction between residues 165–172 and helix C. Residues Q168, Q172, T215, and Q219, which may interact with protein X, are shown with side chains colored cyan. The purple moieties of the backbone ribbon indicate residues in close contact with loop residue V166, i.e., residues Y218, E221, S222, and Y225 of helix C.

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 the random coil values. For example, the methyl groups of L125 are shifted upfield by about 0.2 and 0.4 ppm as the result of ring current effects from Y162 and Y128 (*vide infra*). Long-range connectivities indicate that the hydrophobic cluster has interactions with S1 of the short  $\beta$ -sheet, and it also interacts with helix B through contact between the side chains of L125, I182, and Q186. While residues 121–128 are relatively well ordered, A120 apparently adopts more than one conformation. Figure 9a shows superposition of residues 120–131 from the ensemble of 25 refined structures. Residue 120 clearly evinces two different conformations which are consistent with the experimental data. Two amide signals were observed in the  $^{15}\text{N}$ -edited NOESY spectrum for A120, as seen in Figures 9b and 1. Resonances from the minor form of A120  $\text{H}^{\text{N}}$  were not used in determining restraints for the structure refinement.

The flexibility evident in the hydrophobic cluster results in conformational exchange slow on the NMR time scale. This same conformational exchange probably results in the

broadening evident in the weak, broadened  $^{15}\text{N}$ -HSQC spectral signals for S1 residues Met129 (32 Hz), Leu130 (32 Hz), and Gly131 (44 Hz) as well as S2 residues Tyr162 (33 Hz) and Arg164 (33 Hz). The amide line widths for most residues in the protein core are 24–25 Hz. It would appear that conformational heterogeneity in the hydrophobic cluster and the adjacent irregular  $\beta$ -sheet is a hallmark of this protein.

The conformational heterogeneity and dynamics in the hydrophobic cluster and the adjacent  $\beta$ -sheet may well imply that this region is only marginally stable and, under certain conditions, may be susceptible to the  $\beta$  conversion characterizing the  $\text{PrP}^{\text{C}} \rightarrow \text{PrP}^{\text{Sc}}$  transformation. Our earlier studies demonstrated that solution conditions governed secondary structure for various peptides encompassing residues 90–145 (10). X-ray diffraction studies of the 90–145 peptide folded into a  $\beta$ -structure suggest two interstrand spacings which could be accommodated, for example, by having separate  $\beta$ -sheets formed between stretches with short side chains (e.g., residues 113–120) and between stretches with long side chains (e.g., residues 134–141) (50).

**Back-Calculated Chemical Shifts.** To find additional evidence of some persistent structure in the hydrophobic cluster, we looked for resonances emanating from the hydrophobic cluster and adjacent regions which were significantly shifted from random coil values. Resonances shifted from the random coil values were identified from L125, L130, R164, and I182. From analysis of the structures, we noticed that L125 in the hydrophobic cluster and L130, which is adjacent to the hydrophobic cluster, are both close to Y162. Residues I182 and R164 are close to Y128 in the hydrophobic cluster. We back-calculated chemical shifts from the 25 final structures using the method described by Williamson et al. (51). Selected chemical shifts, i.e., those indicative of some interaction between the hydrophobic cluster and the core region, are shown in Table 2. Proton chemical shifts are fairly accurately predicted for I182  $\text{H}^{\text{r}}$ , I182  $\text{H}^{\text{b}}$ , L125  $\text{H}^{\text{b}}$ , and L130  $\text{H}^{\text{d}}$ . We therefore submit that upfield shifts of  $>0.3$  ppm for protons on residues L125, L130, and L182 imply that the predominant population of rPrP(90–231) possesses structure in the hydrophobic cluster region.

The first 23 residues (gray in Figure 6) of the N-terminus of rPrP(90–231) exhibit little structure despite the existence of some structure elements around residues 96–101. The

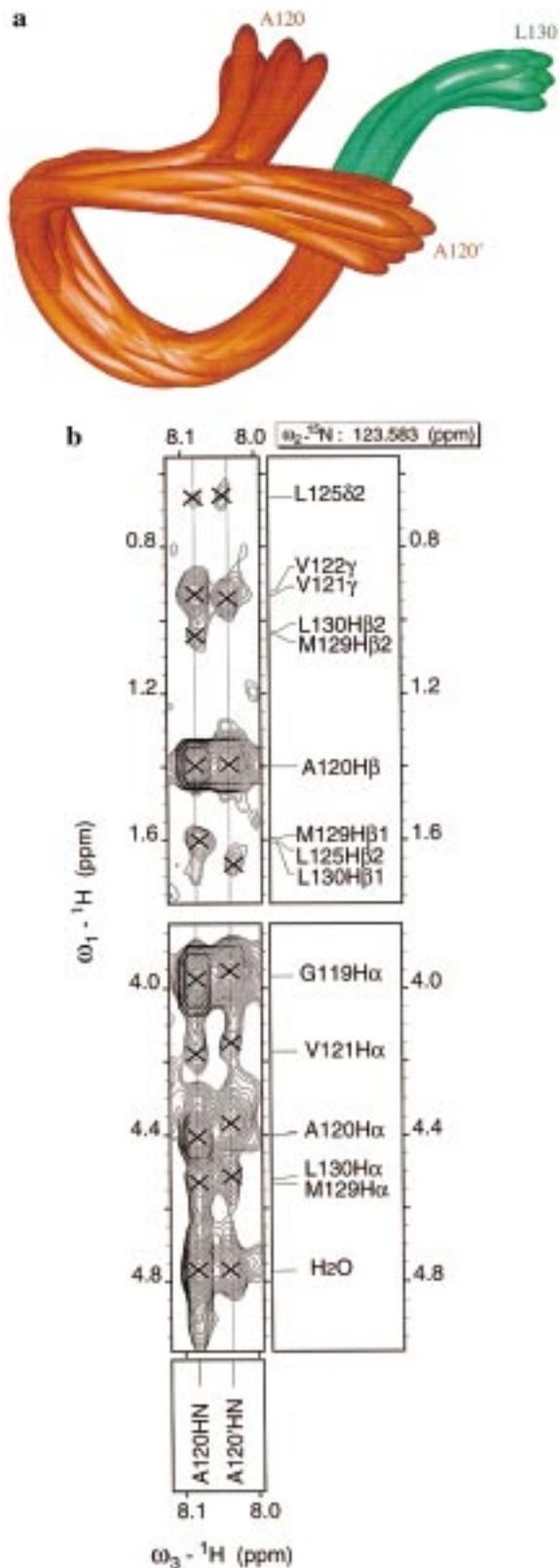


FIGURE 9: (a) Superposition of residues 120–130, showing that A120 has two conformations. Residues 129 and 130 are part of the  $\beta$ -sheet and are thus colored green; residues 120–128 are part of the hydrophobic core and are colored red. (b) Portion of the  $^{15}\text{N}$ -edited NOESY showing two A120 amide resonances and their cross-peaks.

largely unstructured N-terminal region includes epitopes 95–104 and 109–112, which are respectively recognized by

Table 1: Structural Statistics for SHa rPrP(90–231)

no. of distance restraints	
intraresidue	1146
sequential	738
medium range	541
long range	353
unambiguous total	2778
ambiguous	297
no. of torsion angle restraints from $^3J_{\text{H}^{\text{N}}\text{H}^{\text{H}}}$	63
no. of distance restraints	44
energy <sup>a</sup> (kcal/mol)	
$E_{\text{total}}$	$-741 \pm 27$
$E_{\text{vdw}}$	$-761 \pm 20$
$E_{\text{constraint}}$	$430 \pm 34$
$E_{\text{AMBER}}$	$-1171 \pm 61$
residual distance deviation ( $\text{\AA}$ )	$0.032 \pm 0.002$
residual $^3J_{\text{H}^{\text{N}}\text{H}^{\text{H}}}$ deviation (Hz)	$0.059 \pm 0.015$
atomic RMSD of residues 125–227 <sup>b</sup> ( $\text{\AA}$ )	
backbone (N, C $^{\alpha}$ , C)	0.67
all heavy atoms	1.35

<sup>a</sup> 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  $\text{\AA}$  shell of water. Water molecules were removed to calculate the energy terms ( $\pm$  standard deviations) presented. <sup>b</sup> Average parameters for 25 refined structures ( $\pm$  standard deviations).

Table 2: Selected  $^1\text{H}$  Chemical Shifts (ppm) Indicating Interaction between the Hydrophobic Cluster and the Core of SHa rPrP(90–231)

$^1\text{H}$ resonance	random coil shift <sup>a</sup>	exptl shifts	calcd shifts <sup>c</sup>	nearest ring(s)
L125QD	0.94, 0.90	0.75 <sup>b</sup> 0.67 <sup>b</sup>	$0.58 \pm 0.5$ $0.42 \pm 0.2$	Y128, Y162
L130QD	0.94, 0.90	0.66 <sup>b</sup> 0.09 <sup>b</sup>	$0.50 \pm 0.2$ $0.01 \pm 0.5$	Y162
R164QG	1.70	1.14 <sup>b</sup> 1.04 <sup>b</sup>	$1.27 \pm 0.7$ $1.22 \pm 0.6$	Y128
I182G1 (CH $_3$ )	0.95	1.00	$0.52 \pm 0.6$	Y128
I182G2 (CH $_2$ )	1.48, 1.19	0.30	$0.29 \pm 0.4$	Y128
I182D (CH $_3$ )	0.89	0.49	$0.46 \pm 0.2$	Y128

<sup>a</sup> Reference 79, pH 7. <sup>b</sup> Not stereospecifically assigned. <sup>c</sup> From the method of Williamson et al. (51).

recombinant antibody fragment R10 and hybridoma-derived antibody 3F4 (52).

**Structural Statistics.** Structural statistics, covering residues 112–231, of the 25-structure ensemble resulting from refinement are summarized in Table 1. Distance and  $J$  coupling constant restraint violations and energy terms are obtained as the result of restrained energy minimization following restrained molecular dynamics in water with AMBER. Figure 10 shows the distribution of NOE restraints, per-residue atomic RMSD, and backbone angular order parameters. Angular order parameters were calculated for the ensemble of 25 structures as described previously (53). A120 has exceptionally low order parameters for backbone torsion angles  $\phi$  and  $\psi$ , consistent with multiple conformations.

The backbone torsion angles  $\phi$  and  $\psi$  for residues 125–228 of the final 25 structures are shown in Figure 11 as a Ramachandran plot. As calculated by the PROCHECK program, the  $\phi$  and  $\psi$  angles of the structured regions in rPrP(90–231) fall within the areas favorable for  $\alpha$ -helix or  $\beta$ -sheet structure (86.6% in most favorable, 11.6% in allowable, 1.3% in generously allowable regions, and 0.5%

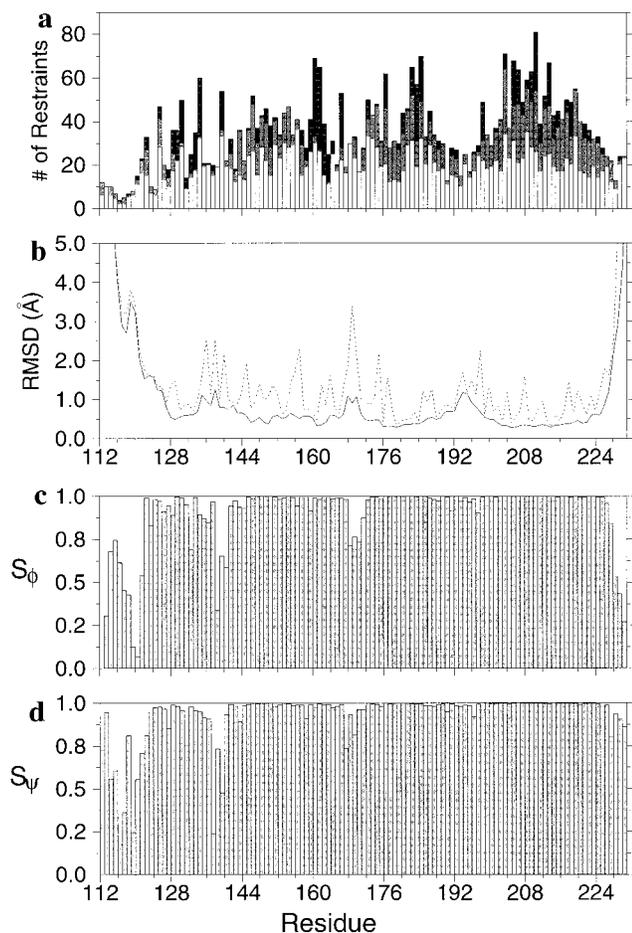


FIGURE 10: (a) Distribution of NOE restraints: white bars for intraresidue and sequential NOEs, light gray for medium-range NOEs, and dark gray for long-range NOEs. (b) The dashed line represents the all-atom RMSD averaged over the 25 structures for each residue, and the solid line represents the average per-residue backbone RMSD. (c) Backbone  $\phi$  angle order parameters calculated for the ensemble of 25 structures. (d) Backbone  $\psi$  angle order parameters calculated for the ensemble of 25 structures.

in disallowed regions), but  $\phi$  and  $\psi$  angles for residues in the 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 angular order parameters are clearly correlated with secondary structural elements, the helices having values of nearly 1. The hydrophobic cluster region is also apparent, with  $\phi$  and  $\psi$  angular order parameters of  $\sim 0.5$ . We did not calculate the order parameters for the region 90–112, since this region was not included in the final stages of the refinement.

We have used the following algorithms to search for possible three-dimensional structure homology of rPrP(90–231) with other known protein structures: DALI (54), CE (55), and CATH (56). We found no matches other than the orientation of two helices in myoglobin and hemoglobin which correspond to the orientation of helices B and C in rPrP(90–231). The sequence alignment and scoring program FASTA (57) has also been employed with the result that for rPrP(90–231) 24% sequence identity was found over 109 residues with two related heat-labile enterotoxins (Protein Data Bank ID codes 1lt3 and 1lt4). However, there is no structural homology between PrP(90–231) and these two proteins.

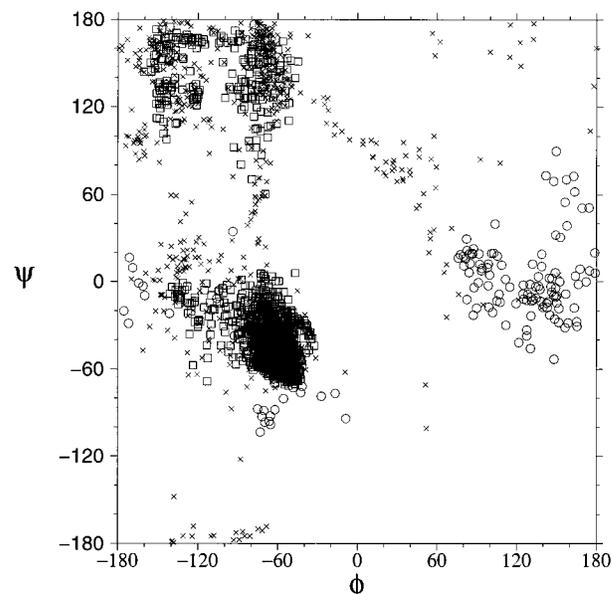


FIGURE 11: Ramachandran plot of the backbone torsion angles  $\phi$  and  $\psi$  of the 25 structures for residues 125–228. The  $\phi$  and  $\psi$  angles are plotted with separate symbols for the structured region (□), less defined loops (×), and glycine residues (○).

### Protein Dynamics

**Molecular Diffusion.** The linearity of the data emanating from the pulse field gradient experiments over a large range of gradient strengths (data not shown) indicates that rPrP(90–231) is homogeneous in terms of oligomeric state. The protein had the same self-diffusion constant at concentrations of 0.1, 0.33, and 1.0 mM at 30 °C. The diffusion constant was measured to be  $0.6 \times 10^{-6} \pm 0.1 \times 10^{-6} \text{ cm}^2/\text{s}$ , assuming a spherical protein. As seen from the structure, the protein is not spherical; furthermore, the unstructured N-terminus would be expected to increase the diffusion constant relative to a compact spherical protein of the same molecular weight. The solvent-accessible surface area of the entire structure (residues 90–231) is calculated to be  $7250 \pm 500 \text{ \AA}^2$  employing the program SURVOL (58) with a probe radius of 1.4 Å and a surface point density of 20.

It is known from sedimentation equilibrium experiments that, at a concentration of 0.025 mM, rPrP(90–231) in the  $\alpha$  form is monomeric (22). 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, given that the same diffusion coefficient is observed for the NMR sample over the range 0.1–1.0 mM and that the oligomeric state is homogeneous, the self-diffusion experiments strongly suggest that the NMR sample of rPrP(90–231) at 1 mM is monomeric.

**Backbone Dynamics.** To characterize the rapid backbone dynamics (i.e., via amides) of rPrP(90–231), the  $^{15}\text{N}$  NMR longitudinal relaxation rates  $R_1$  ( $1/T_1$ ), transverse relaxation rates  $R_2$  ( $1/T_2$ ), and  $^{15}\text{N}\{^1\text{H}\}$  NOE values were measured at 600 MHz (Figure 12). The overall rotational correlation time  $\tau_c$  was estimated as  $11.0 \pm 0.9 \text{ ns}$  using a trimmed mean value of  $R_2/R_1$  (59). It was not possible to obtain a satisfactory description of the backbone dynamics using Lipari–Szabostyle “model-free” analyses (60), even when two molecular motions were assumed. This could be due to difficulty with

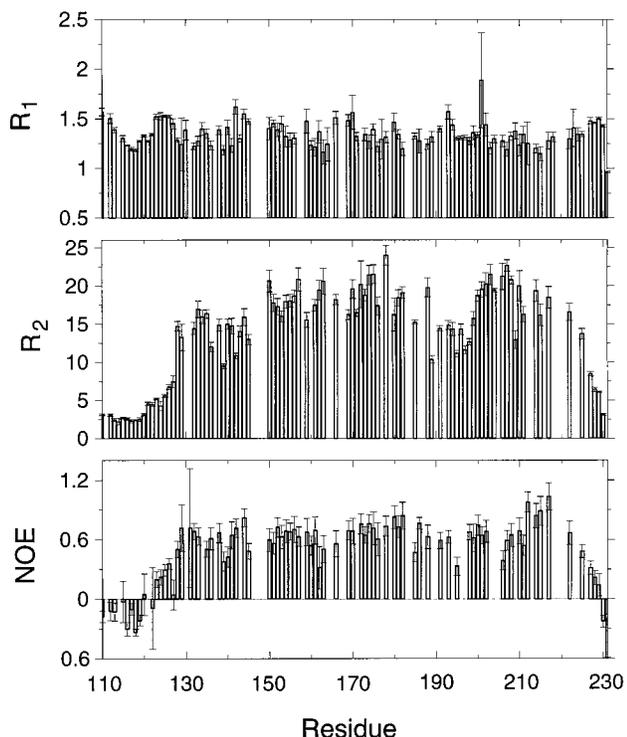


FIGURE 12: Amide  $^{15}\text{N}$  NMR spin-lattice relaxation rate ( $1/T_1$ ), spin-spin relaxation rate ( $1/T_2$ ), and  $^{15}\text{N}\{^1\text{H}\}$  NOE values as a function of residue number reflect backbone flexibility in rPrP(90–231).

the assumption that the protein tumbles isotropically; the pronounced asymmetric shape of the core portion of the protein may be troublesome but, probably more importantly, a structured protein core would certainly not tumble isotropically if it is dragging a long minimally structured tail [the N-terminus of rPrP(90–231)].

Even without mathematical analysis, however,  $^{15}\text{N}$  relaxation parameters indicate that the hydrophobic cluster is more dynamic than the core. The relaxation data also suggest that the C-terminal end of helix B and the subsequent loop residues preceding helix C may possess additional motion. We note that the residue dependence of amide proton exchange correlates with the  $^{15}\text{N}$  relaxation data, especially  $T_2$ .

## DISCUSSION

*Comparison of rPrP(90–231) with Recombinant MoPrP(121–231).* The Syrian hamster PrP sequence is about 94% homologous with MoPrP (61, 62); in fact, there are only six sequence differences in the region 121–231: residues 139 (M  $\rightarrow$  I), 155 (N  $\rightarrow$  Y), 170 (N  $\rightarrow$  S), 203 (I  $\rightarrow$  V), 205 (I  $\rightarrow$  M), and 215 (T  $\rightarrow$  V). Figure 13 shows the superposition of the MoPrP(121–231) structure (Protein Data Bank accession code 1AG2), determined at 20  $^\circ\text{C}$  with 0.8 mM concentration, pH 4.5, with no buffer or salt added, and the corresponding region of the SHa rPrP(90–231) structure, determined at 30  $^\circ\text{C}$  with 0.7–1.3 mM concentration, in 20 mM sodium acetate, pH 5.2, and 0.005% sodium azide. The overall folding of these two proteins is the same, with an RMSD of 1.8  $\text{Å}$  for the backbone atoms of residues 128–214 (i.e., without the hydrophobic cluster or the extension to helix C), and the secondary structure elements are

essentially the same except for the C-terminus of helix C. Residues 167–176 were reported to be disordered in MoPrP(121–231) (13) [later revised to residues 167–171 (14)], but they are well resolved in the spectra and are well defined in structures of SHa rPrP(90–231) with line widths comparable to those from other structured regions of the protein.

Two of the six sequence differences between MoPrP(121–231) and SHa rPrP(90–231), i.e., at residues 203 and 205, would not be expected to have any structural effect and apparently do not; the variation at residue 155 (N  $\rightarrow$  Y) evidently is also without structural effect. However, the change from T215 to V215 and possibly N170 to S170, 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. In particular, substitution of V (in Mo-A) for T (in SHa) may result in the structural difference between MoPrP-A and SHa PrP. The C-terminal helix structure of the mouse protein is broken and bent, with the ring of Y226 involved in a hydrophobic interaction with V215. Residues Q172, V215, and Q219, putatively involved in protein X binding, are also partially shielded by these rings. If that hydrophobic interaction drives the disruption of helix C, it can also serve to destabilize the 165–171 loop. Interaction between V166 and the C-terminus of helix C is also much weaker for MoPrP, with no long-range NOE connectivities reflecting such an interaction being reported.

The loop between helix A and  $\beta$ -strand S1, including the first residue D144 of helix A and part of S1, also differs substantially between MoPrP(121–231) and SHa rPrP(90–231). This region encompasses the species-variant residue 139: I in Mo and Hu sequences and M in the SHa sequence. We find structural differences at two of the three locations (residues 139, 155, and 170) key to the species specificity of prion diseases (63, 64).

*Comparison of rPrP(90–231) with rPrP(29–231).* When the structure and dynamics of SHa rPrP(29–231) were investigated (16), it was found that rPrP(29–231) and rPrP(90–231) have the same structure for residues 125–231 according to observable NOE connectivities and CSI values, and it was also noted that the N-terminal residues 29–124 are highly flexible. The  $\text{C}^\alpha$  chemical shifts of rPrP(29–231) differ from that of rPrP(90–231) by less than 0.2 ppm except for a few residues at the C-terminus of helix B. It was postulated that the chemical shift change might be due to interactions with the octarepeats which are missing in rPrP(90–231) (16). Interestingly, different samples of rPrP(90–231) which had pH differences of up to 0.2 display chemical shift variations for this same region. The protonation state of H187, which is one of the residues with the largest chemical shift change, might vary for the different samples, resulting in the observed resonance shifts. The chemical shift changes could also result from structure instability. The hydrogen-deuterium exchange data and the  $^{15}\text{N}$ -edited NOESY-HSQC spectrum indicate that the helical structure in this region is unstable. Slightly different solution conditions could change the population of conformers with helix B fully intact and thus manifest slight changes in chemical shift.

*Relationship of Structure to Disease.* NMR and other biophysical studies of peptide fragments from PrP indicate that residues 90–145 and 109–141 can form  $\alpha$ -helical

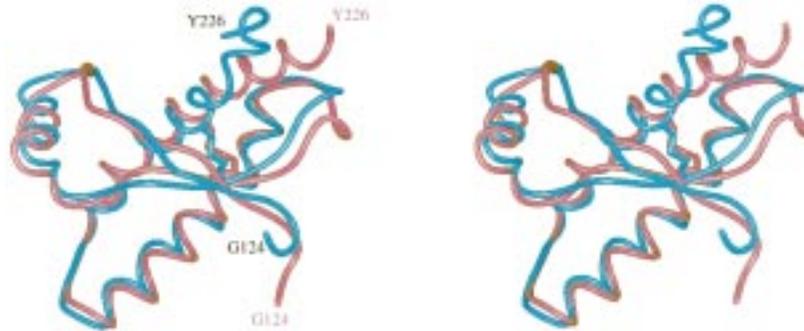


FIGURE 13: Stereoview showing the superposition of the structured core region of SHa rPrP(90–231) (magenta) and MoPrP(121–231) (cyan). Coordinates for the latter structure are available from the Brookhaven Protein Data Bank, accession code 1AG2. For residues 128–214, the backbone RMSD between the two structures is 1.8 Å.

structure or  $\beta$  aggregates depending on the solution conditions, whereas the three shorter peptides 109–122, 104–122, and 129–141, containing either part or none of the hydrophobic core residues, show little propensity for the  $\alpha \rightarrow \beta$  conversion (10, 65). This hydrophobic cluster seems to be a necessary component for PrP to fold into the PrP<sup>Sc</sup> isoform. This is supported by studies with transgenic mice where removal of N-terminal residues beyond position 93 did not support PrP<sup>Sc</sup> formation. Interestingly, Tg mice expressing PrP(121–231) develop a cerebellar disorder in the neonatal period, suggesting that the PrP(121–231) fragment is neurotoxic and that the N-terminal moiety of PrP is necessary for normal function (66).

While the deposition of amyloid plaques containing N-terminal truncated protease-resistant fragments of PrP is diagnostic of prion disease when it occurs, such plaques are not an obligatory feature of these diseases (67). The analysis of brain homogenates from GSS patients showed two major protease-resistant fragments with molecular masses of 21 and 8 kDa (68–70). Both fragments have multiple N-terminal starting points: at residues 78, 80, and 82. The 21 kDa fragment is truncated only at the N-terminus, while the 8 kDa fragment ends with multiple C-terminal truncation spanning residues 147–153. Whether these peptides acquire a high  $\beta$ -sheet content before or after being cleaved is unknown. Presumably, the high  $\beta$ -sheet content of these PrP peptides causes them to aggregate and in some cases form amyloid plaques. Obviously, the  $\beta$  aggregates and plaques can form without the C-terminal residues, i.e., most of the structured core of rPrP(90–231) and MoPrP(121–231). This notion does not conflict with the fact that many of the point mutations leading to inherited prion diseases are in the C-terminal region (1, 14). Protein X is postulated to interact with helix C and loop 165–172 (15). Consistent with these observations, the two domains, structured core and hydrophobic cluster, interact with each other as we demonstrate in the structure presented here.

So far, we have very little information on the N-terminus, comprised substantially of the four octapeptide repeat sequences, and how it interacts with the structured core, but there is evidence that the N-terminus and the hydrophobic cluster play an important role in the  $\alpha \rightarrow \beta$  conversion and the molecular pathology of prion diseases. Region 113–131 is the most conserved sequence among mammals (71). Protease-resistant fragments associating with different pathological phenotypes of prion diseases always contain this hydrophobic cluster, albeit with a mutation inducing  $\beta$ -sheet

formation in some cases (e.g., A117V can lead to GSS) (72, 73). In our structure, M129 is at the border of the  $\beta$ -strand S1. Methionine is a strong  $\alpha$ -helix-forming residue, while valine is the strongest  $\beta$ -forming residue. The presence of valine at 129 should further stabilize S1, and perhaps this is important in propagation of the different disease phenotypes. For example, GSS patients with different point mutations in this region present with different clinical phenotypes. The mutation A117V leads to GSS patients with dementia when residue 129 is methionine, but a family of GSS patients with a valine substitution at both residue 117 and 129 presented with ataxia (74). We previously noted that the polymorphism of residue 129 is correlated with disease phenotype expressed by the D178N mutation (12, 75). These clinical observations are consistent with the notion that the different prion disease manifestations relate to somewhat different conformations of PrP<sup>Sc</sup> being present. Residue 129 is quite close to D178; indeed, for 8 of the 25 refined structures presented here, the carboxylate of D178 is H-bonded to the ring hydroxyl of Y128. The D178N mutation could disrupt that H-bond, resulting in different conformations for the M129 and V129 variants.

The characteristics of the hydrophobic cluster and  $\beta$ -strand S1 are reminiscent of the ensemble of conformations which have been postulated to coexist at an intermediate stage of protein folding or unfolding (76). Compact denatured state conformations can be principally driven by hydrophobic interactions with addition of H-bonding to form helices or  $\beta$ -sheets. In the case of PrP(90–231), the hydrophobic cluster and  $\beta$ -strand S1 appear to have short-lived conformations which need additional environmental impetus to form a single, more stable structure. We hypothesize that this structural vacillation is key to the PrP<sup>C</sup>  $\rightarrow$  PrP<sup>Sc</sup> transformation, as different members of the ensemble of conformations characterizing the hydrophobic cluster in the naked protein may serve to facilitate arrangement into different conformations of PrP<sup>C</sup> or PrP<sup>Sc</sup>.

A hydrophobic cluster has also been identified in the  $\beta$ -amyloid peptide ( $A\beta$ ) that forms amyloid fibrils in Alzheimer's diseases (77). A point mutation in the hydrophobic cluster of the  $A\beta$  amyloid peptide abolishes its ability to form amyloid plaques (78). The importance of the hydrophobic cluster in other neurodegenerative diseases associated with protein aggregates suggests that conformational promiscuity of a hydrophobic cluster may be the central theme of many neurodegenerative diseases.

## ACKNOWLEDGMENT

We are 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. We gratefully acknowledge Thomas Goddard for responsive development of features useful for structure analysis in the SPARKY software and Dr. Peter J. Dommey for suggesting use of ARIA.

## REFERENCES

- Prusiner, S. B. (1997) *Science* 278, 245–251.
- Stahl, N., Baldwin, M. A., Teplow, D. B., Hood, L., Gibson, B. W., Burlingame, A. L., and Prusiner, S. B. (1993) *Biochemistry* 32, 1991–2002.
- Pan, K.-M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R. J., Cohen, F. E., and Prusiner, S. B. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90, 10962–10966.
- Pergami, P., Jaffe, H., and Safar, J. (1996) *Anal. Biochem.* 236, 63–73.
- Meyer, R. K., McKinley, M. P., Bowman, K. A., Braunfeld, M. B., Barry, R. A., and Prusiner, S. B. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 2310–2314.
- Prusiner, S. B., Groth, D. F., Bolton, D. C., Kent, S. B., and Hood, L. E. (1984) *Cell* 38, 127–134.
- Prusiner, S. B., McKinley, M. P., Bowman, K. A., Bolton, D. C., Bendheim, P. E., Groth, D. F., and Glenner, G. G. (1983) *Cell* 35, 349–358.
- Caughey, B., and Raymond, G. J. (1991) *J. Biol. Chem.* 266, 18217–18223.
- Stahl, N., Baldwin, M. A., Teplow, D. B., Hood, L., Gibson, B. W., Burlingame, A. L., and Prusiner, S. B. (1993) *Biochemistry* 32, 1991–2002.
- Zhang, H., Kaneko, K., Nguyen, J. T., Livshits, T. L., Baldwin, M. A., Cohen, F. E., James, T. L., and Prusiner, S. B. (1995) *J. Mol. Biol.* 250, 514–526.
- Mehlhorn, I., Groth, D., Stöckel, J., Moffat, B., Reilly, D., Yansura, D., Willett, W. S., Baldwin, M., Fletterick, R., Cohen, F. E., Vandlen, R., Henner, D., and Prusiner, S. B. (1996) *Biochemistry* 35, 5528–5537.
- James, T. L., Liu, H., Ulyanov, N. B., Farr-Jones, S., Zhang, H., Donne, D. G., Kaneko, K., Groth, D., Mehlhorn, I., Prusiner, S. B., and Cohen, F. E. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 10086–10091.
- Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R., and Wüthrich, K. (1996) *Nature* 382, 180–182.
- Riek, R., Wider, G., Billeter, M., Hornemann, S., Glockshuber, R., and Wüthrich, K. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 11667–11672.
- Kaneko, K., Zulianello, L., Scott, M., Cooper, C. M., Wallace, A. C., James, T. L., Cohen, F. E., and Prusiner, S. B. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 10069–10074.
- Donne, D. G., Viles, J. H., Groth, D., Mehlhorn, I., James, T. L., Cohen, F. E., Prusiner, S. B., Wright, P. E., and Dyson, H. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 13452–13457.
- Riek, R., Hornemann, S., Wider, G., Glockshuber, R., and Wüthrich, K. (1997) *FEBS Lett.* 413, 282–288.
- Stöckel, J., Safar, J., Wallace, A. C., Cohen, F. E., and Prusiner, S. B. (1998) *Biochemistry* 37, 7185–7193.
- Hornshaw, M. P., McDermott, J. R., Candy, J. M., and Lakey, J. H. (1995) *Biochem. Biophys. Res. Commun.* 214, 993–999.
- Pan, K.-M., Stahl, N., and Prusiner, S. B. (1992) *Protein Sci.* 1, 1343–1352.
- Brown, D. R., Qin, K., Herms, J. W., Madlung, A., Manson, J., Strome, R., Fraser, P. E., Kruck, T., von Bohlen, A., Schulz-Schaeffer, W., Giese, A., Westaway, D., and Kretschmar, H. (1997) *Nature* 390, 684–687.
- Zhang, H., Stöckel, J., Mehlhorn, I., Groth, D., Baldwin, M. A., Prusiner, S. B., James, T. L., and Cohen, F. E. (1997) *Biochemistry* 36, 3543–3553.
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* 6, 277–293.
- Day, M. (1994) *STRIKER (UCSF)*, Copyright University of California.
- Muhandiram, D. R., and Kay, L. E. (1994) *J. Magn. Reson. B* 103, 203–216.
- Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson.* 96, 432–440.
- Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson. B101*, 201–205.
- Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc.* 114, 6291–6293.
- Grzesiek, S., and Bax, A. (1993) *J. Biomol. NMR* 3, 185–204.
- Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1993) *J. Biomol. NMR* 3, 225–231.
- Kay, L. E., Xu, G.-Y., Singer, A. U., Muhandiram, D. R., and Forman-Kay, J. D. (1993) *J. Magn. Reson. B101*, 333–337.
- Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry* 28, 6150–6156.
- Muhandiram, D. R., Farrow, N. A., Xu, G. Y., Smallcombe, S. H., and Kay, L. E. (1993) *J. Magn. Reson. B102*, 317–321.
- Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) *J. Biomol. NMR* 4, 871–878.
- Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shelton, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry* 33, 5984–6003.
- Goddard, T. (1998) *SPARKY (UCSF)*, Copyright University of California.
- Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins* 17, 75–86.
- Altieri, A. S., Hinton, D. P., and Byrd, R. A. (1995) *J. Am. Chem. Soc.* 117, 7566–7567.
- Gallo, K., Huang, C., Ferrin, T. E., and Langridge, R. (1989) *MIDASPlus*, University of California, San Francisco.
- Nilges, M., Macias, M. J., O'Donoghue, S. I., and Oschkinat, H. (1997) *J. Mol. Biol.* 269, 408–422.
- Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc.* 115, 7772–7777.
- Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc.* 117, 5179–5197.
- Pearlman, D. A., Case, D. A., Caldwell, J., Ross, W. S., Cheatham, T. E., III, Ferguson, D. N., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1995) *AMBER 4.1 (UCSF)*, Copyright University of California.
- Folmer, R. H. A., Hilbers, C. W., Konings, R. N. H., and Nilges, M. (1997) *J. Biomol. NMR* 9, 245–258.
- Wishart, D. S., and Sykes, B. D. (1994) *Methods Enzymol.* 239, 363–392.
- Case, D. A., Dyson, H. J., and Wright, P. E. (1994) *Methods Enzymol.* 239, 392–416.
- Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* 8, 477–486.
- Richardson, J. S., and Richardson, D. C. (1988) *Science* 240, 1648–1652.
- Kabsch, W., and Sander, C. (1983) *Biopolymers* 22, 2577–2637.
- Inouye, H., and Kirschner, D. A. (1998) *J. Struct. Biol.* 122, 247–255.
- Williamson, M. P., Kikuchi, J., and Asakura, T. (1995) *J. Mol. Biol.* 247, 541–546.
- Peretz, D., Williamson, R. A., Matsunaga, Y., Serban, H., Pinilla, C., Bastidas, R., Rozenshteyn, R., James, T. L., Houghten, R. A., Cohen, F. E., Prusiner, S. B., and Burton, D. R. (1997) *J. Mol. Biol.* 273, 614–622.
- Hyberts, S. G., Goldberg, M. S., Havel, T. F., and Wagner, G. (1992) *Protein Sci.* 1, 736–751.
- Holm, L., and Sander, C. (1993) *J. Mol. Biol.* 233, 123–138.

55. Shindyalov, I. N., and Bourne, P. E. (1998) *Protein Eng.* 11, 739–747.
56. Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B., and Thornton, J. M. (1997) *Structure* 5, 1093–1108.
57. Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 2444–2448.
58. Alard, P. (1991) Calcul de surface d'énergie dans le domaine des molécules, Ph.D. Thesis Dissertation, Université libre de Bruxelles.
59. Kay, L. E., Torchia, D. A., and Bax, A. (1989) *Biochemistry* 28, 8972–8979.
60. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc.* 104, 4546–4559.
61. Basler, K., Oesch, B., Scott, M., Westaway, D., Wälchli, M., Groth, D. F., McKinley, M. P., Prusiner, S. B., and Weissmann, C. (1986) *Cell* 46, 417–428.
62. Locht, C., Chesebro, B., Race, R., and Keith, J. M. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 6372–6376.
63. Scott, M., Groth, D., Foster, D., Torchia, M., Yang, S.-L., DeArmond, S. J., and Prusiner, S. B. (1993) *Cell* 73, 979–988.
64. Kocisko, D. A., Priola, S. A., Raymond, G. J., Chesebro, B., Lansbury, P. T., Jr., and Caughey, B. (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92, 3923–3927.
65. Gasset, M., Baldwin, M. A., Lloyd, D. H., Gabriel, J.-M., Holtzman, D. M., Cohen, F., Fletterick, R., and Prusiner, S. B. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 10940–10944.
66. Shmerling, D., Hegyi, I., Fischer, M., Blättler, T., Brandner, S., Götz, J., Rüllicke, T., Flechsig, E., Cozzio, A., von Mering, C., Hangartner, C., Aguzzi, A., and Weissmann, W. (1998) *Cell* 93, 203–214.
67. DeArmond, S. J., and Prusiner, S. B. (1997) in *Greenfield's Neuropathology* (Lantos, P., and Graham, D., Eds.) pp 235–280, Edward Arnold, London.
68. Tagliavini, F., Prelli, F., Ghiso, J., Bugiani, O., Serban, D., Prusiner, S. B., Farlow, M. R., Ghetti, B., and Frangione, B. (1991) *EMBO J.* 10, 513–519.
69. Tagliavini, F., Prelli, F., Porro, M., Rossi, G., Giaccone, G., Farlow, M. R., Dlouhy, S. R., Ghetti, B., Bugiani, O., and Frangione, B. (1994) *Cell* 79, 695–703.
70. Parchi, P., Chen, S. G., Brown, P., Zou, W., Capellari, S., Budka, H., Hainfellner, J., Reyes, P. F., Golden, G. T., Hauw, J. J., Gajdusek, D. C., and Gambetti, P. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 8322–8327.
71. Bamforth, P., Wille, H., Telling, G. C., Yehiely, F., Prusiner, S. B., and Cohen, F. E. (1996) *Cold Spring Harbor Symp. Quant. Biol.* 61, 495–509.
72. Chen, S. G., Teplow, D. B., Parchi, P., Teller, J. K., Gambetti, P., and Autilio-Gambetti, L. (1995) *J. Biol. Chem.* 270, 19173–19180.
73. Telling, G. C., Parchi, P., DeArmond, S. J., Cortelli, P., Montagna, P., Gabizon, R., Mastrianni, J., Lugaresi, E., Gambetti, P., and Prusiner, S. B. (1996) *Science* 274, 2079–2082.
74. Mastrianni, J. A., Curtis, M. T., Oberholtzer, J. C., Da Costa, M. M., DeArmond, S., Prusiner, S. B., and Garbern, J. Y. (1995) *Neurology* 45, 2042–2050.
75. Goldfarb, L. G., Petersen, R. B., Tabaton, M., Brown, P., LeBlanc, A. C., Montagna, P., Cortelli, P., Julien, J., Vital, C., Pendelbury, W. W., et al. (1992) *Science* 258, 806–808.
76. Dill, K. A., and Chan, H. S. (1997) *Nat. Struct. Biol.* 4, 10–19.
77. Glenner, G. G., and Wong, C. W. (1984) *Biochem. Biophys. Res. Commun.* 120, 885–90.
78. Esler, W. P., Stimson, E. R., Ghilardi, J. R., Vinters, H. V., Lee, J. P., Mantyh, P. W., and Maggio, J. E. (1996) *Biochemistry* 35, 749–757.
79. Bundi, A., and Wüthrich, K. (1979) *Biopolymers* 18, 285–298.

BI982878X