Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 13452–13457, December 1997
Biochemistry

Structure of the recombinant full-length hamster prion protein PrP(29–231): The N terminus is highly flexible

(NMR structure/conformational change/backbone dynamics)

DAVID G. DONNE*†‡, JOHN H. VILES*†‡, DARLENE GROTH§, INGRID MEHLHORN§, THOMAS L. JAMES¶‖, FRED E. COHEN¶**††, STANLEY B. PRUSINER§‡‡, PETER E. WRIGHT*†§§, AND H. JANE DYSON*§§

*Department of Molecular Biology and †Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 90237; and Departments of §Neurology, ¶Pharmaceutical Chemistry, ‖Radiology, **Cellular and Molecular Pharmacology, ††Medicine, and ‡‡Biochemistry and Biophysics, The University of California, San Francisco, CA, 94143

Contributed by Stanley B. Prusiner, September 16, 1997

ABSTRACT **The prion diseases seem to be caused by a conformational change of the prion protein (PrP) from the benign cellular form PrP^{C} to the infectious scrapie form PrP^{Sc}; thus, detailed information about PrP structure may provide essential insights into the mechanism by which these diseases develop. In this study, the secondary structure of the recombinant Syrian hamster PrP of residues 29–231 [PrP(29–231)] is investigated by multidimensional heteronuclear NMR. Chemical shift index analysis and nuclear Overhauser effect data show that PrP(29–231) contains three helices and possibly one short β-strand. Most striking is the random-coil nature of chemical shifts for residues 30–124 in the full-length PrP. Although the secondary structure elements are similar to those found in mouse PrP fragment PrP(121–231), the secondary structure boundaries of PrP(29–231) are different from those in mouse PrP(121–231) but similar to those found in the structure of Syrian hamster PrP(90–231). Comparison of resonance assignments of PrP(29–231) and PrP(90–231) indicates that there may be transient interactions between the additional residues and the structured core. Backbone dynamics studies done by using the heteronuclear $[^{1}H]$-^{15}N nuclear Overhauser effect indicate that almost half of PrP(29–231), residues 29–124, is highly flexible. This plastic region could feature in the conversion of PrP^{C} to PrP^{Sc} by template-assisted formation of β-structure.**

Prion diseases are a novel class of neurodegenerative diseases, including scrapie in sheep; bovine spongiform encephalopathy in cattle; and kuru, Creutzfeldt–Jacob disease, Gerstmann–Sträussler–Scheinker syndrome, and fatal familiar insomnia in humans (1, 2). A substantial body of evidence has shown that the scrapie isoform of the prion protein (PrP^{Sc}) is the major component of prions (3, 4). The prion protein (PrP) is composed of ≈250 amino acids and is encoded by a single chromosomal gene. Cellular PrP (PrP^{C}) is processed to remove a 22-amino acid N-terminal signal peptide and 23 C-terminal residues upon the addition of the glycosylphosphatidylinositol anchor to Ser-231. Residues 179 and 214 are linked by a disulfide bridge, and Asn-181 and Asn-197 are glycosylation sites in the native protein. The fundamental event in prion disease pathology appears to be the profound conformational change during conversion of PrP^{C} to PrP^{Sc} (5, 6).

To understand the conformational transition that accompanies PrP^{Sc} formation, a knowledge of the three-dimensional structure of the PrP is essential. CD and Fourier transform infrared studies suggest that conversion of α-helices into β-sheets features in the formation of PrP^{Sc} (5). Recently, the solution structure of a mouse PrP fragment, MoPrP(121–231), was reported to contain three α-helices and two short antiparallel β-strands (7, 8). Although MoPrP(121–231) contains a significant portion of a minimal region of the protease-resistant core (residues 90–231) that can mediate disease, the fragment lacks the biologically important N-terminal octarepeat region ([P(Q/H)GGG(G/-)WGQ]x5) and does not contain all of the sites of pathologic point mutations (9). The recent NMR structure of the Syrian hamster (SHa) PrP(90–231) (10) corresponding to the 27–30-kDa infectious core of PrP^{Sc} (designated PrP 27–30) indicates that the longer fragment also contains three helices and a short, irregular two-strand β-sheet. All three helices are longer in PrP(90–231) than those in MoPrP(121–231). In addition, the disordered loop between the last two helices in MoPrP(121–231) shows a greater degree of order in the structure of PrP(90–231), and residues 113–125 show some interactions with the folded core of the protein. The octarepeat region is not present in either MoPrP(121–231) or PrP(90–231), but an increased number of octarepeats is known to cause inherited prion disease (11–13). Because PrP can adopt two or more conformations, it is unclear whether the residues corresponding to MoPrP(121–231) will display context-dependent structural behavior in the full protein.

We report here a structural study by NMR of a recombinant SHa PrP(29–231) that represents the entire protein, including the octarepeats. The largely conserved octarepeat region may play an important role in the conversion of PrP^{C} to PrP^{Sc}, either by participation in the profound conformational change, or in nonspecific binding to membranes (14) in the subcellular compartment where the conversion takes place (15). An earlier study of an octarepeat peptide by CD spectroscopy suggested that the repeat region adopts a nonrandom, extended conformation with the properties of a poly-L-proline left-handed helix (14). In the present paper the secondary structure of full-length SHa PrP(29–231) has been determined, and the dynamic properties of the protein backbone have been measured. We postulate a plausible role for the octarepeat region in prion disease pathogenesis. During the preparation of this paper, a report describing a preliminary analysis of the NMR spectrum of MoPrP(23–231) appeared (16). The present paper describes an extensive analysis of the full-length SHa PrP(29–231), including complete backbone

Abbreviations: PrP, prion protein; PrP^{Sc}, scrapie PrP isoform; PrP^{C}, cellular PrP isoform; Mo, mouse; SHa, Syrian hamster; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; TOCSY, total correlation spectroscopy; CSI, chemical shift index.

‡D.G.D. and J.H.V. contributed equally to this work.

§§To whom reprint requests should be addressed at: Department of Molecular Biology, The Scripps Research Institute, MB-2, 10550 North Torrey Pines Road, La Jolla, CA 92037. e-mail: wright@scripps.edu or dyson@scripps.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

resonance assignments obtained from triple-resonance experiments and measurements of polypeptide chain dynamics.

MATERIALS AND METHODS

Protein Expression and Purification. Expression of both recombinant SHa PrP(29–231) and PrP(90–231) was accomplished by using a secreted ph041 vector driven by an alkaline phosphatase promoter in a protease-deficient strain of *Escherichia coli* (27C7). Expression of the proteins was induced about 9 hr postinoculation by phosphate depleting of the growth medium. Uniform isotopic labeling of the protein was done by minimal modifications of the protocol described earlier (17). The bacterium was grown overnight in a shaker flask containing 100 ml of Celtone microbial growth media labeled with either ^{15}N or $^{15}N/^{13}C$. The culture then was used for inoculation of a 2-liter fermentation where the following components were isotopically labeled: isoleucine, ammonium sulfate, ammonium hydroxide, glucose, and 10% reconstituted Celtone powdered media for the 20% yeast extract and NZ amines. The fermentations were allowed to proceed for either 24 hr for ^{15}N labeling or 12–14 hr for $^{15}N/^{13}C$ labeling. Wet cell paste yield was between 40–50 g, ultimately giving a minimum of 100 mg of purified protein. This yield is about half of the amount normally obtained from unlabeled fermentations.

Purification was as described except that 20 mM methionine was added to the size exclusion material before HPLC separation to reduce methionine oxidation. Purified PrP(29–231) or PrP(90–231) was lyophilized before refolding. Lyophilized PrP(29–231) or PrP(90–231) was solubilized at 1 mg/ml in 6–8 M guanidinium hydrochloride and then rapidly diluted to 100 μg/ml in 25 mM Tris·HCl at pH 8.0 and 5 mM EDTA concentration. The refolded protein then was dialyzed overnight against 20 mM sodium acetate at pH 5.0 and 0.005% sodium azide. Protease inhibitors are essential for maintaining full-length protein. Insoluble material was removed by filtration through a 0.2 μm filter. Samples were concentrated first by Centriprep 10 (Amicon) and then further by Centriplus 10 (Amicon). Deuterated buffer exchange was done simultaneously with the final concentrating step. Samples were analyzed by mass spectroscopy, Fourier transform infrared, and CD spectroscopy.

NMR Spectroscopy. NMR samples were prepared in 20 mM sodium acetate-*d*5 buffer at a final pH of 5.2 in 90% ^1H_2O/10% ^2H_2O. All solutions contained 0.005% sodium azide to inhibit bacterial growth. All deuterated reagents were purchased from Cambridge Isotope Laboratories (Cambridge, MA). NMR data were acquired at 30°C calibrated with neat methanol. The protein concentration for the ^{15}N-labeled sample was about 1 mM and the $^{15}N/^{13}C$-labeled sample about 0.6 mM for PrP(29–231). The final concentration for ^{15}N-labeled PrP(90–231) used in the heteronuclear nuclear Overhauser effect (NOE) experiments was about 1 mM. The spectrometers used in the study were Bruker DMX 750 for the ^{15}N and ^{13}C separated NOE spectroscopy (NOESY) experiments and Bruker DRX 600 and AMX 600 instruments for all other triple-resonance experiments. All spectrometers are equipped with 5-mm triple-axis gradient triple-resonance probes and an 8-mm actively shielded Z-gradient probe for the DRX 600 instrument. For backbone and side-chain resonance assignments, the following experiments were carried out: two-dimensional ^{15}N and ^{13}C heteronuclear single-quantum coherence (HSQC) (18, 19), three-dimensional HNCA (20), HNCACB (21), CBCA(CO)NH (22), ^{15}N-total correlation spectroscopy (TOCSY)-HSQC (23), C(CO)NH-TOCSY (24), HCCH-TOCSY (25), ^{15}N-NOESY-HSQC (23) with 50-ms and 80-ms mixing times, ^{13}C-NOESY-HSQC (26) with an 80-ms mixing time, HNCO (20), (HCA)CO(CA)NH (27), and HBHA(CBCACO)NH (28). All experiments were modified for samples dissolved in 90% H_2O with sensitivity-enhanced gradient coherence selection (29, 30) and minimal water saturation (31). The heteronuclear [^1H]-^{15}N NOE experiments (32) for both PrP(29–231) and PrP(90–231) were carried out on both Bruker DRX 600 and AMX-II 500 spectrometers. Two data sets were acquired for both PrP(29–231) and PrP(90–231). In each data set, two experiments with and without proton saturation were carried out. The final NOE values were calculated from the ratio of the saturated intensities divided by the unsaturated intensities from each data set. The mean values from the two data sets then were calculated, and the error bars were taken to be the difference between the values from the two data sets. All NMR data were processed and analyzed by using FELIX95 (Molecular Simulations, San Diego, CA) that was run on Silicon Graphics (Indigo and O2) computers. All spectra were referenced relative to 2,2-dimethyl-2-silapentane-5-sulfonate, directly for ^1H and indirectly for ^{13}C and ^{15}N dimensions as described by Wishart *et al.* (33).

RESULTS

Backbone and Side-Chain Resonance Assignments of PrP(29–231). ^{15}N-labeled and $^{13}C/^{15}N$-labeled protein samples of PrP(29–231) were used for the multidimensional heteronuclear NMR studies. The protein was found to be monomeric and stable for up to 3 months at pH 5.2 and 30°C in 20 mM sodium acetate buffer with 10% D_2O and 0.005% sodium azide, as monitored by CD spectroscopy and NMR. Residue types were identified by using aliphatic ^{13}C chemical shifts from a combination of HNCA, CBCA(CO)NH, and C(CO)NH-TOCSY experiments (28). Sequential assignments of ^1H, ^{13}C, and ^{15}N resonances of the protein were obtained by using a combination of C^α- and CO-based triple-resonance experiments. In the C^α-based approach, combined use of HNCACB and CBCA(CO)NH experiments allowed about 80% of the amide proton and nitrogen resonances of residue *i* to be correlated with the C^α and C^β chemical shifts of residues *i* and *i*-1. Sequential connectivities not identified in the above combination because of low sensitivity of the HNCACB experiment were made by using HNCA and CBCA(CO)NH experiments. Ambiguities caused by C^α degeneracy were resolved by the CO-based triple-resonance experiments [HNCO and (HCA)CO(CA)NH] that take advantage of the large intrinsic dispersion of carbonyl chemical shifts (34). In these experiments, the amide proton and nitrogen resonances of residue *i* were correlated with the carbonyl chemical shifts of residues *i* and *i*-1. Side-chain H^α and H^β resonances were obtained by using the HBHA(CO)NH experiment in conjunction with the ^{15}N-TOCSY-HSQC and ^{15}N-NOESY-HSQC experiments. Resonance assignments were verified by using sequential through-space connectivities in ^{15}N-resolved NOESY spectra. The combined use of these experiments yielded complete backbone assignments except for the N-terminal Gly-29. Complete C^α assignments were achieved except for Leu-130, and 97% of C^β and 98% of CO resonances were assigned, as were 93% of H^α resonances. Assignments of the aromatic ring protons have not been made at this point. The assignments for PrP(90–231) were obtained as described (10).

Resonances of corresponding residues in four of five octarepeats (PHGGGWGQ) essentially were unresolved. The first octarepeat has a slightly different sequence (PQGGGTWGQ) from the other four, and the GTW portion could be assigned sequence-specifically. Residues in the same position in the remaining four octarepeats were degenerate in all frequencies and gave one intense cross peak in C^α and H^α, C^β, and H^β (if any), N, NH, and even CO resonances. Comparison of the intensity of amide proton/nitrogen cross peaks in the ^{15}N HSQC spectrum with residues on either side of the octarepeat region gave the expected 1:4 intensity ratio.

Secondary Structure of PrP(29–231). The chemical shift index (CSI) has been used to identify secondary structure elements (35–37). The deviation of C^{α}, H^{α}, H^{β}, and ^{13}CO chemical shifts from their random coil values (38, 39) suggests whether the residues are involved in α-helices or β-sheets in proteins. The presence of secondary structure elements then can be verified from short-range and medium-range NOEs that are indicative of different types of secondary structures (40). Fig. 1 shows the C^{α}, CO, and H^{α} chemical shift deviation of PrP(29–231) from random coil values. The most striking feature is that almost half of the protein is indistinguishable in chemical shift from random coil. In particular no deviations are greater than 0.5 ppm from random-coil C^{α} chemical shifts for residue 29–124. In contrast the C-terminal portion of PrP(29–231) displays large positive and negative deviations (typically 4 ppm) from random-coil values for C^{α} resonances. The same patterns are seen in the chemical shifts of H^{α}, CO, and C^{β} resonances.

CSI analysis indicates three α-helices and two short β-strands. Helix A spans residues 144–156, helix B contains residues 172–193, and helix C includes residues 200–227. The locations of the helices are confirmed by the presence of medium-range $\alpha N(i, i+3)$ and $\alpha N(i, i+4)$ NOEs and supported by continuous sequential NN NOEs in the helical regions. As shown in Fig. 2, medium-range NOEs are observed between residues 144–157, 174–192, and 200–226. A short piece of β-strand spanning residues 160–163 is identified by C^{α}, H^{α}, CO, and C^{β} chemical shifts. The accompanying β-strand to pair with residues 160–163 is less clear, although residues 137–140 are suggested to contain β-structure by C^{α} and CO chemical shifts. There is no indication from the CSI that the β-strand identified in the solution structures of MoPrP(121–231) (7) and PrP(90–231) (10) between residues 129–131 is present in a β-conformation (Fig. 1), although interpretation of the data is complicated by missing assignments caused by exchange-broadened resonances in this region. On the other hand, there is evidence in the form of a single long-range NOE between the NH protons of residues 129 and 163 for a close contact between the 129–131 and 160–163 regions of the polypeptide; there is no such contact evident between 160–163 and 137–140, the β-region identified by the chemical shifts. For the octarepeat region, the NOEs remain ambiguous because of degeneracy in proton chemical shifts and are not shown in Fig. 2.

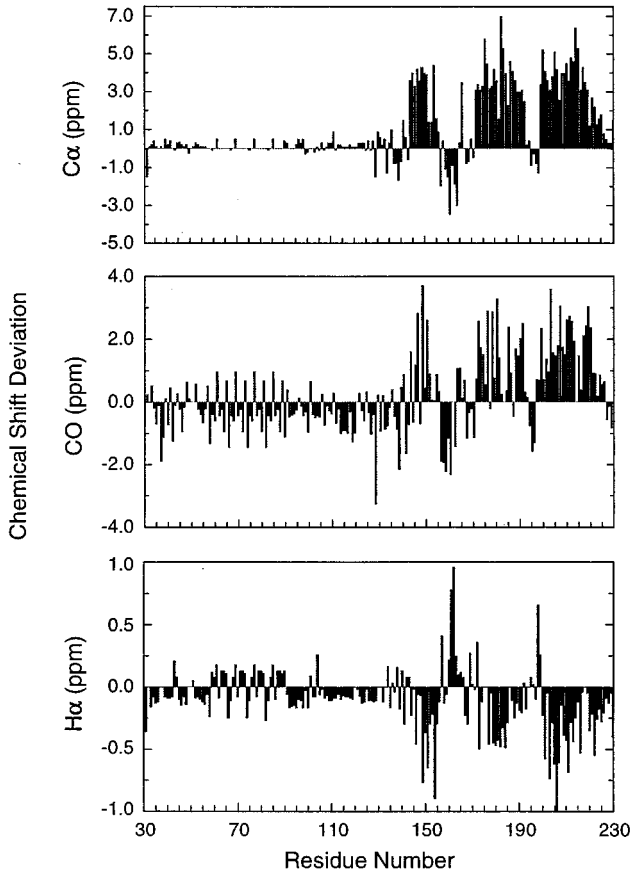

FIG. 1. Summary of C^{α}, H^{α}, and ^{13}CO chemical shift deviations from random-coil values (38, 39) used to identify the secondary structure of PrP(29–231). The chemical shift deviations in ppm are plotted against residue numbers in PrP(29–231). Note the repeating nature of shift values in the octarepeat region.

To investigate the possible influence of the additional 61 residues in PrP(29–231) upon the structure of PrP(90–231), a comparison was made of the chemical shift assignments from the two proteins. The C^{α} chemical shift differences are shown in Fig. 3. In addition to the expected chemical shift difference for the residues at the N terminus of PrP(90–231), residues 187–193 display consistent and significant differences (greater than 0.3 ppm) between PrP(29–231) and PrP(90–231). Amide protons and nitrogens also give the most significant changes in chemical shifts for residues 187–193, with amide nitrogen shift differences greater than 0.2 ppm in this region. Most of the C^{α} and amide nitrogen chemical shifts show differences less than $\pm$ 0.1 ppm between the two proteins.

Backbone Dynamics of PrP(29–231) and PrP(90–231). Heteronuclear $[^1H]$-^{15}N NOE experiments were carried out to assess the flexibility of PrP(29–231) and PrP(90–231). PrP(29–231) contains 13 prolines and an N-terminal amino group not available for heteronuclear NOE measurements. Amide peaks from 14 residues are severely overlapped, and the intensities could not be reliably measured. In addition, residues L130 and G131 give very weak signals and also were excluded from the NOE measurements. Of the 188 amide peaks, 172 heteronuclear NOEs are included in the analysis. For PrP(90–231), apart from prolines, residues 146, 162, 184, and 189 were excluded because of severe overlap of the peaks.

The heteronuclear NOE data are shown for PrP(29–231) and PrP(90–231) in Fig. 4 *A* and *B*, respectively. As can be seen from Fig. 4*A*, the 500-MHz heteronuclear NOE values for residues 31–124 of PrP(29–231) are all negative, typically -0.5. Amides in the octarepeat region give similar NOE values to those on both sides of the region. The periodicity shown for residues 59–91 arises because the NOE values were approximated by dividing (by 4) the volumes of the degenerate cross peaks corresponding to residues in analogous positions in the four octarepeats. By contrast, strong positive NOEs (typically over 0.65) are observed for residues 134–223. The C-terminal residues 228–231 give very small positive or negative NOEs. In addition, residues 190–197, corresponding to the loop between helix B and helix C have relatively lower positive NOE values. The heteronuclear NOEs obtained at 600 MHz gave comparable magnitudes, and the change in the sign of the NOEs occurs again at residue Gly-124. The linewidth of the amide peaks in the HSQC spectrum for the N-terminal region is about 6 Hz narrower than those in the core region. In addition, residues 129, 130, and 131 give very weak signals in all two-dimensional and three-dimensional experiments, presumably caused by exchange broadening of these amide peaks. The $[^1H]$-^{15}N NOEs for the residues in the structured core region of PrP(90–231) have similar values to those of corresponding residues of PrP(29–231) (Fig. 4). For both proteins, the dividing point for negative and positive NOEs occurs between residues 124 and 125. The heteronuclear NOEs of the loop residues (190–197) also are lowered in PrP(90–231), as in PrP(29–231).

FIG. 2. Sequential and medium-range NOEs used to verify the secondary structure elements derived from CSI analysis. NOEs between amide protons of consecutive residues and between H^{α} and the amide proton of subsequent residues are represented by bars connecting the residues. For NN(i, $i+1$) and αN(i, $i+1$) NOEs, the thickness of the bar qualitatively represents the relative intensity (weak, medium, or strong) of the NOE. The NOEs for four octarepeats are not shown because they were assigned only to residue types in the repeat region because of overlap. The three unambiguously identified α-helices also are shown.

DISCUSSION

For the identification of secondary structure elements, the H^{α}-based and ^{13}C-based CSI analyses were used. By using the majority rule (two of three or three of four) (37) and considering that $^{13}C^{\alpha}$ chemical shift has the highest predicting power, we unambiguously have identified three α-helices, with helix A spanning residues 144–156, helix B including residues 172–193, and helix C containing residues 200–227. By using a consensus from C^{α}, CO, and H^{α} CSI a 92% confidence in prediction has been reported (37). Furthermore, the helical residues as indicated by CSI analysis are almost identical to those derived from medium range [αN(i, $i+3$) and αN(i, $i+4$)] NOEs (Fig. 2). The lengths of the three helices derived for SHa PrP(29–231) are similar to those reported for PrP(90–231) (10) but are different from those of MoPrP(121–231) (7, 8). In MoPrP(121–231), the three helices A, B, and C span residues 144–154, 179–193, and 200–217, respectively (7). The differences in the lengths of the three helices presumably are caused by the absence of critical residues in MoPrP(121–231) (10). In addition, the differences in buffer conditions and the more acidic pH for the study of MoPrP(121–231) also may have contributed to the variations in secondary structure.

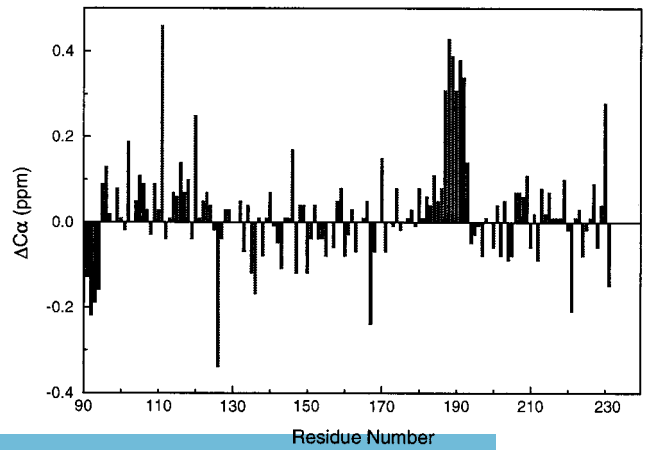

FIG. 3. Difference in the C^{α} chemical shift between values measured for PrP(29–231) and PrP(90–231) plotted against residue number.

The chemical shift and NOE-derived helices of PrP(29–231) are very similar to those in the recently determined NMR structure of PrP(90–231) (10). A short antiparallel β-sheet has been reported for MoPrP(121–231) (7) and an irregular β-sheet for SHa PrP(90–231) (10). All of the chemical shift indices suggest a β-strand between residues 160–163 of PrP(29–231). However, the location of the other strand is not clear from the chemical shift data, with the segment between residues 137–140 being the only region where both the C^{α} and CO chemical shifts suggest a possibility of a β-strand. In the reported NMR structures of MoPrP(121–231) and PrP(90–231), the β-sheet consists of residues 161–164 and 128–131; the presence of this latter strand in PrP(29–231) cannot be inferred from the CSIs. Whether this discrepancy reflects poor predictive power of the CSI index method in this case, where the β-strands are short and some assignments are missing, or whether it shows the presence of small populations of alternative conformations for the prion proteins in the presence of the 29–129 region must await further structural work on PrP(29–231). Residues 137–140 might simply be in an extended conformation that fulfills the geometric characteristics of a β-strand. It has been speculated that the short β-sheet in the MoPrP(121–231) serves as a nucleation site for the conversion of PrP^{C} to PrP^{Sc} (7). It is attractive to consider the possibility that a β-strand consisting of residues 137–140 in PrP(29–231) could form part of a β-sheet during the conversion of PrP^{C} to PrP^{Sc}. This is consistent with computational studies of PrP^{Sc}, in which one of the four plausible β-strands spans residues 138–144 (41).

The heteronuclear $[^{1}H]$-^{15}N NOE is an excellent measure of flexibility in proteins (42). Most of the residues between 130–227 of PrP(29–231) give large positive NOEs, indicating

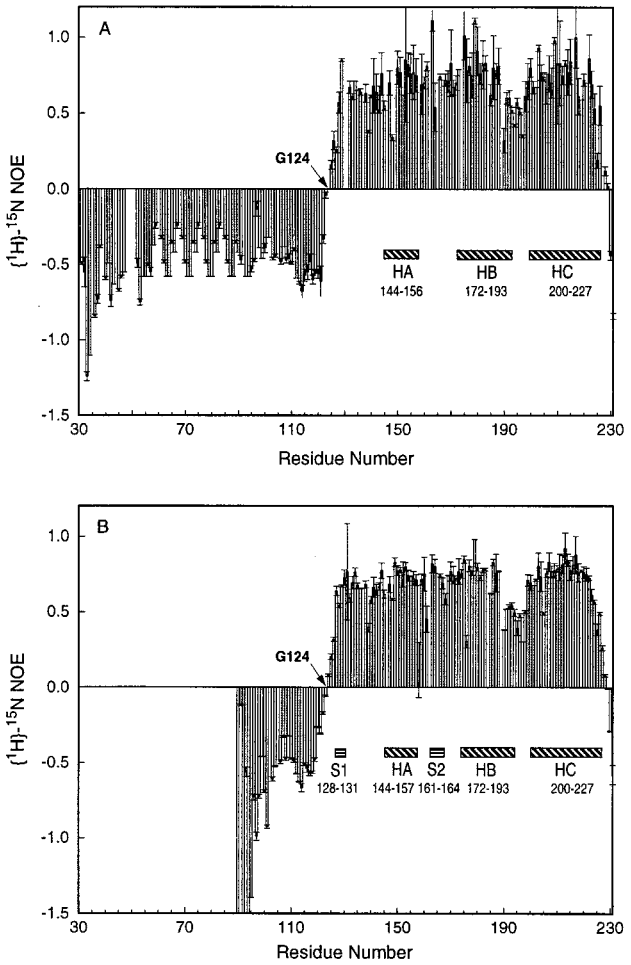

FIG. 4. Summary of heteronuclear [^{1}H]-^{15}N NOE data used to assess the backbone flexibility of the PrP(29–231) and PrP(90–231). (*A*) The mean values of heteronuclear NOE from two sets of data are plotted against residue number in the full-length SHa PrP(29–231). The error bars represent the difference between the two data sets. The NOEs in four of the octarepeats (residues 59–91) are approximate values estimated by dividing the degenerate peak volumes by four. (*B*) Heteronuclear NOE values for PrP(90–231). All spectroscopic parameters and processing procedures are the same as those of PrP(29–231).

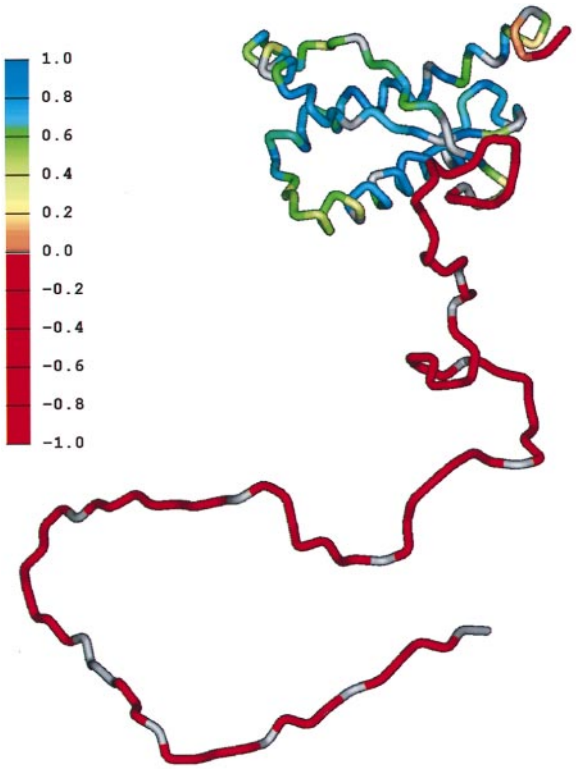

FIG. 5. Schematic diagram showing the flexibility of the polypeptide chain for PrP(29–231). The structure of the portion of the protein representing residues 90–231 was taken from the coordinates of PrP(90–231) (10). The remainder of the sequence was hand-built for illustration purposes only. The color scale shows the heteronuclear [^{1}H]-^{15}N NOE data from Fig. 4*A*, from red for the lowest (most negative) values, where the polypeptide is most flexible, to blue for the highest (most positive) values in the most structured and rigid regions of the protein.

these residues are within the structured core. Residues 190–197, corresponding to the loop between helix B and helix C, have smaller, but still positive, NOE values, indicating these residues are relatively more flexible than those in regular secondary structure. Five residues at the C terminus of both PrP(29–231) and PrP(90–231) are flexible, a property common to the termini of many proteins. The region from 90–124 is clearly flexible in both PrP(29–231) and PrP(90–231), with large negative NOEs and local correlation times less than 1 ns, indicating that these residues do not form part of any stable secondary or tertiary fold and are highly flexible (43). The flexibility of the various parts of the polypeptide chain in PrP(29–231) is shown schematically in Fig. 5. The presence of a folded C-terminal domain of residues 124–231, with disorder in the remainder of the chain, also has been noted in MoPrP(23–231) (16).

The octarepeat region (five repeats in most species) is a functionally important section of the PrP (9). Addition of two, four, five, six, seven, eight, nine, and 18 octarepeats in humans results in an inherited prion disease (11–13). One additional octarepeat has been reported in some cattle but it does not seem to alter the susceptibility of those animals to bovine spongiform encephalopathy (44, 45). Deletion of one repeat in humans creates a polymorphism that apparently does not cause disease (9). From a structural viewpoint, it is clear that the presence of the 61 additional residues in PrP(29–231), relative to PrP(90–231), does not have a significant effect on the secondary structure elements. In particular, residues 90–124 remain highly flexible in the full-length protein. This is consistent with recent studies of the antigenic structures of PrP^C and PrP^{Sc}, which document that the structure of the region between 90–124 varies between the cellular and scrapie isoforms whereas the C-terminal region is unaltered (46).

The only major difference between PrP(29–231) and PrP(90–231) is an apparent consistent difference in $^{13}C^{\alpha}$ chemical shifts for residues 187–193 (Fig. 3), suggesting that stabilization of helical structure has occurred in helix B in PrP(29–231). This effect may well be caused by transient tertiary interactions between the N-terminal residues containing the octarepeats and helix B in the longer fragment.

Although the function of the octarepeat region remains uncertain, several possibilities can be re-evaluated in light of our structural observations. First, the octarepeats may participate in binding to the membrane or another protein. In either case, the flexibility of this region, as documented by heteronuclear NOE data, would facilitate binding to a partner macromolecule. Alternatively, the octarepeats may be in an apoprotein conformation that will adopt a more organized structure in the presence of a cofactor. Candidates for this cofactor would include Cu(II) (J. Stöckel, F.E.C., and S.B.P., unpublished observations). Finally, in spite of the flexibility of this region, chemical shift data indicate that the 29–125 region stabilizes the conformational ensemble of the helix between residues 187 and 193. This region is important in PrP^C/PrP^{Sc} recognition but also must serve a normal physiologic function.

Although the function of PrP remains unknown, the large unstructured N terminus of PrP(29–231) might provide a clue to its function. Although it is unusual for almost half of any protein to be unstructured, the phenomenon is not unprecedented. Whether the discovery of the unstructured N terminus of full-length PrP will provide insight into the function of PrP^C remains to be established, but this highly flexible region probably provides the plasticity required for the conformational transition of PrP^C into PrP^{Sc}. Studies with recombinant mAbs support this postulate in that epitopes located at the N terminus, but not the C terminus, of PrP^{Sc} are buried whereas epitopes throughout native PrP^C are accessible (46).

We conclude from an analysis of four relevant NMR studies of PrP^C molecules (refs. 7, 10, and 16, and this work) that 29–124 forms a flexible region, and that 125–231 forms a disulfide-bridged three-helix bundle. A small amount of context-dependent β-structure may exist in PrP^C, but its relevance is uncertain. The flexible regions are clearly biologically important, if only in the process of conversion from PrP^C to PrP^{Sc}. It is not clear whether they participate in the normal cellular function of the PrP in the same unstructured state. Comparison of biologically inactive (121–231) and active (90–231) fragments of PrP^C highlights the role of the 165–175 loop in PrP^{Sc} replication, which may be mediated by another factor, termed protein X (47). The surface of PrP^C contains distinct PrP^{Sc} and protein X recognition domains that are constructed from spatially proximal, but sequentially distant regions of the polypeptide chain. PrP^{Sc} recognizes a surface that includes residues from the regions 90–144 and 180–205, whereas protein X recognizes an antipodal surface that includes residues from the regions 160–180 and 205–231 (48). Thus, the biological and structural data presently available on the PrPs indicate that conversion of PrP^C to PrP^{Sc} is a complex process that involves context-dependent conformational changes in the N-terminal region of the protein (46). The discovery that this region is largely unstructured in solution is important for our understanding of the conversion process, because the energy barrier to the formation of the β-sheet containing PrP^{Sc} will be much lower from an initial "random coil" than from a structure that already contains stable secondary structure.

We thank Dr. John Chung for assistance with NMR experiments and Mike Pique for preparing Fig. 5. This work was supported by grants from the National Institutes of Health.

1. Gajdusek, D. C. (1977) *Science* **197,** 943–960.
2. Prusiner, S. B. (1997) *Science* **278,** 245–251.
3. Prusiner, S. B. (1982) *Science* **216,** 136–144.
4. Prusiner, S. B. (1991) *Science* **252,** 1515–1522.
5. Pan, K.-M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R. J., Cohen, F. E. & Prusiner, S. B. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 10962–10966.
6. Horwich, A. L. & Weissman, J. S. (1997) *Cell* **89,** 499–510.
7. Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R. & Wüthrich, K. (1996) *Nature (London)* **382,** 180–182.
8. Billeter, M., Riek, R., Wider, G., Hornemann, S., Glockshuber, R. & Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 7281–7285.
9. Bamborough, P., Wille, H., Telling, G. C., Yehiely, F., Prusiner, S. B. & Cohen, F. E. (1996) *Cold Spring Harbor Symp. Quant. Biol.* **61,** 495–509.
10. 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. & Cohen, F. E. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 10086–10091.
11. Goldfarb, L. G., Brown, P., McCombie, W. R., Goldgaber, D., Swergold, G. D., Wils, P. R., Cervenakova, L., Baron, H., Gibbs, C. J., Jr. & Gajdusek, D. C. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 10926–10930.
12. Poulter, M., Baker, H. F., Frith, C. D., Leach, M., Lofthouse, R., Ridley, R. M., Shah, T., Owen, F., Collinge, J., Brown, J., Hardy, J., Mullan, M. J., Harding, A. E., Bennett, C., Doshi, R. & Crow, T. J. (1992) *Brain* **115,** 675–685.
13. Capellari, S., Vital, D., Parchi, P., Petersen, R. B., Ferrer, X., Jarnier, D., Pegoraro, E., Gambetti, P. & Julien, J. (1997) *Neurology* **49,** 133–141.
14. Smith, C. J., Drake, A. F., Banfield, B. A., Bloomberg, G. B., Palmer, M. S., Clarke, A. R. & Collinge, J. (1997) *FEBS Lett.* **405,** 378–384.
15. Kaneko, K., Vey, M., Scott, M., Pilkuhn, S., Cohen, F. E. & Prusiner, S. B. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 2333–2338.
16. Riek, R., Hornemann, S., Wider, G., Glockshuber, R. & Wüthrich, K. (1997) *FEBS Lett.* **413,** 282–288.
17. Mehlhorn, I., Groth, D., Stockel, J., Moffat, B., Reilly, D., Yansura, D., Willett, W. S., Baldwin, M., Fletterick, R., Cohen, F. E., Vandlen, R., Henner, D. & Prusiner, S. B. (1996) *Biochemistry* **35,** 5528–5537.
18. Bodenhausen, G. & Ruben, D. J. (1980) *Chem. Phys. Lett.* **69,** 185–189.
19. Vuister, G. W. & Bax, A. (1992) *J. Magn. Reson.* **98,** 428–435.
20. Grzesiek, S. & Bax, A. (1992) *J. Magn. Reson.* **96,** 432–440.
21. Wittekind, M. & Mueller, L. (1993) *J. Magn. Reson.* **101,** 201–205.
22. Grzesiek, S. & Bax, A. (1992) *J. Am. Chem. Soc.* **114,** 6291–6293.
23. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M. & Clore, G. M. (1989) *Biochemistry* **28,** 6150–6156.
24. Logan, T. M., Olejniczak, E. T., Xu, R. X. & Fesik, S. W. (1993) *J. Biomol. NMR* **3,** 225–231.
25. Bax, A., Clore, G. M. & Gronenborn, A. M. (1990) *J. Magn. Reson.* **88,** 425–431.
26. Muhandiram, D. R., Farrow, N. A., Xu, G.-Y., Smallcombe, S. H. & Kay, L. E. (1993) *J. Magn. Reson. Series B* **102,** 317–321.
27. Löhr, F. & Rüterjans, H. (1995) *J. Biomol. NMR* **6,** 189–197.
28. Grzesiek, S. & Bax, A. (1993) *J. Biomol. NMR* **3,** 185–204.
29. Palmer, A. G., III, Cavanagh, J., Wright, P. E. & Rance, M. (1991) *J. Magn. Reson.* **93,** 151–170.
30. Kay, L. E., Keifer, P. & Saarinen, T. (1992) *J. Am. Chem. Soc.* **114,** 10663–10665.
31. Grzesiek, S. & Bax, A. (1993) *J. Am. Chem. Soc.* **115,** 12593–12594.
32. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D. & Kay, L. E. (1994) *Biochemistry* **33,** 5984–6003.
33. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L. & Sykes, B. D. (1995) *J. Biomol. NMR* **6,** 135–140.
34. Yao, J., Dyson, H. J. & Wright, P. E. (1997) *FEBS Lett.* in press.
35. Spera, S. & Bax, A. (1991) *J. Am. Chem. Soc.* **113,** 5490–5492.
36. Wishart, D. S., Sykes, B. D. & Richards, F. M. (1992) *Biochemistry* **31,** 1647–1651.
37. Wishart, D. S. & Sykes, B. D. (1994) *J. Biomol. NMR* **4,** 171–180.
38. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. (1995) *J. Biomol. NMR* **5,** 67–81.
39. Merutka, G., Dyson, H. J. & Wright, P. E. (1995) *J. Biomol. NMR* **5,** 14–24.
40. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
41. Huang, Z., Prusiner, S. B. & Cohen, F. E. (1995) *Fold. Design* **1,** 13–19.
42. Kay, L. E., Torchia, D. A. & Bax, A. (1989) *Biochemistry* **28,** 8972–8979.
43. Fushman, D., Cahill, S. & Cowburn, D. (1997) *J. Mol. Biol.* **266,** 173–194.
44. Goldmann, W., Hunter, N., Martin, T., Dawson, M. & Hope, J. (1991) *J. Gen. Virol.* **72,** 201–204.
45. Prusiner, S. B., Fuzi, M., Scott, M., Serban, D., Serban, H., Taraboulos, A., Gabriel, J. M., Wells, G. A., Wilesmith, J. W., Bradley, R., DeArmond, S. J. & Kristensson, K. (1993) *J. Infect. Dis.* **167,** 602–613.
46. Peretz, D., Williamson, R. A., Matsunaga, Y., Serban, H., Pinilla, C., Bastidas, R. B., Rozenshteyn, R., James, T. L., Houghten, R. A., Cohen, F. E., Prusiner, S. B. & Burton, D. R. (1997) *J. Mol. Biol.*, in press.
47. Telling, G. C., Scott, M., Mastrianni, J., Gabizon, R., Torchia, M., Cohen, F. E., DeArmond, S. J. & Prusiner, S. B. (1995) *Cell* **83,** 79–90.
48. Kaneko, K., Zulianello, L., Scott, M., Cooper, C. M., Wallace, A. C., James, T. L., Cohen, F. E. & Prusiner, S. B. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 10069–10074.