High-Level Expression and Characterization of a Purified 142-Residue Polypeptide of the Prion Protein†

Ingrid Mehlhorn,‡ Darlene Groth,‡ Johannes Stöckel,‡ Barbara Moffat,§ Dorothea Reilly,§ Daniel Yansura,§ W. Scott Willett,|| Michael Baldwin,‡ Robert Fletterick,|| Fred E. Cohen,||,⊥ Richard Vandlen,§ Dennis Henner,§ and Stanley B. Prusiner*,‡,||

Departments of Neurology, of Biochemistry and Biophysics, and of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94143, and Genentech Inc., South San Francisco, California 94080

Received December 15, 1995; Revised Manuscript Received February 21, 1996⊗

ABSTRACT: The major, and possibly only, component of the infectious prion is the scrapie prion protein (PrP^{Sc}); the protease resistant core of PrP^{Sc} is PrP 27−30, a protein of ∼142 amino acids. PrP^{Sc} is derived from the cellular PrP isoform (PrP^{C}) by a post-transliatonal process in which a profound conformational change occurs. Syrian hamster (SHa) PrP genes of varying length ranging from the N- and C-terminally truncated 90−228 up to the full-length mature protein 23−231 were inserted into various secretion and intracellular expression vectors that were transformed into *Escherichia coli* deficient for proteases. Maximum expression was obtained for a truncated SHaPrP containing residues 90−231, which correspond to the sequence of PrP 27−30; disruption of the bacteria using a microfluidizer produced the highest yields of this protein designated rPrP. After solubilization of rPrP in 8 M GdnHCl, it was purified by size exclusion chromatography and reversed phase chromatography. During purification the recovery was ∼50%, and from each liter of *E. coli* culture, ∼50 mg of purified rPrP was obtained. Expression of the longer species containing the basic N-terminal region was less successful and was not pursued further. The primary structure of rPrP was verified by Edman sequencing and mass spectrometry, and secondary structure determined by circular dichroism and Fourier transform infrared spectroscopy. When rPrP was purified under reducing conditions, it had a high β-sheet content and relatively low solubility similar to PrP^{Sc}, particularly at pH values >7. Refolding of rPrP by oxidation to form a disulfide bond between the two Cys residues of this polypeptide produced a soluble protein with a high α-helical content similar to PrP^{C}. These multiple conformations of rPrP are reminiscent of the structural plurality that characterizes the naturally occuring PrP isoforms. The high levels of purified rPrP which can now be obtained should facilitate determination of the multiple tertiary structures that PrP can adopt.

Prion diseases seem to be disorders of protein conformation that produce fatal CNS degeneration (Cohen et al., 1994; Prusiner, 1994). The prion diseases of animals include scrapie of sheep and goats as well as bovine spongiform encephalopathy. Those of humans include Creutzfeldt−Jakob disease (CJD), Gerstmann−Sträussler−Scheinker syndrome (GSS), and fatal familial insomnia (FFI) (Gajdusek, 1977; Masters et al., 1981; Medori et al., 1992). The human prion diseases are unique in that they can present as sporadic, inherited, or infectious illnesses (Prusiner, 1989). In many cases, the human diseases have been transmitted to experimental animals (Brown et al., 1994; Tateishi et al., 1992; Telling et al., 1995). The transmissible pathogens causing the prion diseases are composed largely, if not entirely, of an abnormal isoform of the prion protein (PrP) designated PrP^{Sc} (Prusiner, 1991).

Both PrP^{Sc} and the normal, cellular PrP isoform (PrP^{C}) are encoded within a single exon of a chromosomal gene as a protein of ∼250 amino acids (Basler et al., 1986). Many mammalian PrPs have a 22 amino acid N-terminal signal sequence (Hope et al., 1986; Turk et al., 1988) and a 23 amino acid C-terminal signal sequence encoding for attachment of a glycosylphosphatidylinositol (GPI) anchor (Stahl et al., 1987, 1990). The mature protein of 209 amino acids contains one disulfide bond (Turk et al., 1988) and has two sites of asparagine-linked glycosylation (Endo et al., 1989; Oesch et al., 1985). By SDS−PAGE, both PrP isoforms exhibit heterogeneity with an M_r of 33−35 kDa. PrP^{C} is ubiquitous in neuronal cells and may be converted into the abnormal pathogenic isoform PrP^{Sc} through a profound conformational change (Pan et al., 1993). PrP^{C} is soluble in non-denaturing detergents (Meyer et al., 1986; Pan et al., 1992) and is readily degraded by proteolysis (Oesch et al., 1985), whereas PrP^{Sc} is insoluble and is resistant to proteolysis. Limited digestion of PrP^{Sc} results in truncation of

† This work was supported by grants from the National Institutes of Health (NS14069, AG08967, AG02132, NS22786, and AG10770) and the American Health Assistance Foundation, as well as by gifts from the Sherman Fairchild Foundation and Bernard Osher Foundation. I.M. is supported by a Senior Research Fellowship from the Canadian Medical Research Council and the French Foundation for Alzheimer's Research. J.S. is supported by the Stipendienprogramm Infektionsforschung of the German Cancer Research Center, Heidelberg. ESI-MS was carried out in the UCSF Mass Spectrometry Facility (A. L. Burlingame, Director) supported by NIH NCRR BRTP RR01614. MALDI-MS was carried out with equipment kindly provided by Perseptive Biosystems, Inc., Cambridge, MA.

* All correspondence to Dr. Stanley Prusiner, Department of Neurology, HSE-781, University of California, San Francisco, CA 94143-0518. Tel: (415) 476-4482. FAX: (415) 476-8386.

‡ Department of Neurology.

§ Genentech Inc.

|| Department of Biochemistry and Biophysics.

⊥ Department of Cellular and Molecular Pharmacology.

⊗ Abstract published in *Advance ACS Abstracts,* April 1, 1996.

S0006-2960(95)02965-5 CCC: $12.00

the N-terminus to produce PrP 27−30 of ∼142 amino acids. Attempts to identify a post-translational covalent modification that initiates the conformational change during the formation of PrP^{Sc} have been unsuccessful (Stahl et al., 1992, 1993). PrP^{C} has a high α-helical content and is devoid of β-sheet while PrP^{Sc} has a high β-sheet content (Pan et al., 1993; Safar et al., 1993). PrP 27−30 polymerizes in rod-shaped polymers that exhibit properties similar to those of many purified amyloids including a high β-sheet content (Caughey et al., 1991; Gasset et al., 1993; Glenner et al., 1972; Pan et al., 1993; Prusiner et al., 1983; Safar et al., 1993).

To study the stucture of the PrP isoforms, it has been desirable to express high levels of recombinant PrP reflecting the structure of the naturally occurring PrP^{C} isoform. A number of recombinant expression systems have been explored including mammalian, baculovirus, vaccinia virus, bacteria, and yeast, and a variety of constructs were expressed in these systems, including fusion proteins, GPI-minus proteins, and different portions of PrP (Caughey et al., 1988; Rogers et al., 1992; Scott et al., 1988). However, significant difficulties were experienced in almost all cases, such as low levels of expression, insolubility, and protein instability during expression or purification. A glutathione S-transferase fusion protein has recently been described, but the cleaved PrP product was unstable and could not be isolated (Weiss et al., 1995). We report here on the expression of SHaPrP genes of varying length ranging from the N- and C-terminally truncated 90−228 up to the full-length mature protein 23−231 that were inserted into expression vectors and transformed into *Escherichia coli* deficient for proteases. Maximum expression was obtained with SHaPrP containing residues 90−231 corresponding to the sequence of PrP 27−30. After purification, ∼50 mg of this recombinant SHaPrP-(90−231), designated rPrP, was obtained from each liter of *E. coli* culture. When rPrP was purified under reducing conditions, it had a high β-sheet content and relatively low solubility similar to PrP^{Sc}, particularly at pH values >7. Refolding of rPrP by oxidation to form a disulfide bond between the two Cys residues of this polypeptide produced a soluble protein with a high α-helical content similar to PrP^{C}. These multiple conformations of rPrP recollect the structural plasticity exhibited by the naturally occurring PrP isoforms. The availability of large quantities of purified rPrP should facilitate studies designed to determine the multiple tertiary structures PrP is capable of adopting.

MATERIALS AND METHODS

Molecular biology reagents were purchased from New England Biolabs, Boerhinger Mannheim, and BRL. Media for small-scale expression (AP5) were described previously (Chang et al., 1987). Other reagents were obtained from the following suppliers: 8 M Gdn-HCl solution, DTT (Pierce); Gdn-HCl salt, acetonitrile (Fisher); ampicillin, EDTA, iodoacetic acid, MES, Trizma base (Sigma); TFA (Perkin Elmer); $^{2}H_2O$ (Aldrich); proteinase K, protease inhibitors (Boerhinger Mannheim). Media for other expression systems were from Gibco and Sigma.

Expression Constructs. PCR (Perkin-Elmer) was used to amplify the DNA corresponding to different portions of the Syrian hamster prion protein in order to ligate it into *E. coli* secretion vectors. Several 5′ oligonucleotide primers were synthesized with an *Mlu*I restriction site within the C-terminal coding sequence of the STII signal peptide (Lee et al., 1983; Picken et al., 1983) and the initial amino acids of the appropriate PrP sequence. One 3′ oligonucleotide primer matching the 3′ end of PrP, a stop codon, and a *Bam*HI restriction site was used with each of the 5′ oligonucleotides. The PCR-amplified products were purified, ligated into the vectors previously digested with *Mlu*I/*Bam*HI and transformed into DH5α. Clones containing the PrP insert were sequenced and transformed into the protease deficient expression strain 27C7 (ATCC 55244).

Expression. Preliminary small-scale expression was carried out in shaken tubes; 5 mL cultures were grown at 37 °C for 3 h in LB medium supplemented with ampicillin (50 μg/mL) followed by a 100-fold dilution into AP5 medium. These cultures were harvested 21 h later, and the total cell protein was isolated by an acetone precipitation. In all cases, the equivalent of 0.1 OD_{600} unit of cell pellets was loaded onto an SDS−PAGE gel and analyzed by Coomassie staining or Western blots (Barry & Prusiner, 1986; Towbin et al., 1979). Large-scale expression was carried out as described previously for other proteins using a different medium (Carter et al., 1992); 500 mL of an overnight culture grown in LB medium supplemented with ampicillin was inoculated into 7 L of fermentation medium in an aerated 10 L fermentor (Braun, model E10). Cells were grown at 37 °C at a high agitation rate and expression was induced by phosphate starvation. After 4 h, a 50% glucose solution was added at a rate of 1 mL/min; glucose levels were monitored using a glucose dipstick (Diastix, Miles Inc.). A pH of 7.4 was maintained throughout the run by the automated addition of 10% H_2SO_4 or 24% NH_4OH. The final volume was 10 L in which an OD_{600} of ≥100 was achieved after 36 h. The *E. coli* was harvested by centrifugation at 10 000*g* for 30 min, and the resulting paste was stored at −20 °C.

Purification. Typically, 100 g of *E. coli* paste was resuspended in 1 L of 25 mM Tris-HCl, pH 8.0/5 mM EDTA (buffer A). This was centrifuged at 10 000*g* for 20 min, and the supernatant containing soluble periplasmic proteins was discarded. The pellet was resuspended in 1 L of buffer A, passed through a cell disrupter twice (Microfluidics International, model MF110), and centrifuged at 30 000*g* for 1 h, after which the supernatant was discarded and the pellet was washed once in buffer A and centrifuged again at 30 000*g* for 1 h. At this stage the pellet could be stored at −20 °C prior to further separation. It was subsequently solubilized in 8 M Gdn-HCl/25 mM Tris-HCl, pH 8.0/100 mM DTT (buffer B) and centrifuged at 14 000*g* for 20 min to remove the remaining insoluble matter. Aliquots of 6 mL of the supernatant containing ∼200 mg of total protein were separated by size exclusion chromatography (SEC) using a 26 mm × 60 cm HiLoad Superdex 200 column (Pharmacia), eluting with 6 M Gdn-HCl/12.5 mM Tris-HCl, pH 8.0/5 mM DTT/1 mM EDTA (buffer C) at a flow rate of 2 mL/min. Fractions enriched for the recombinant prion protein as identified by SDS−PAGE were pooled and further purified by reversed phase high-performance liquid chromatography (RP-HPLC) employing a 25 mm × 25 cm C-4 column (Vydac): buffer 1, H_2O/0.1% TFA; buffer 2, acetonitrile/0.09% TFA; flow rate, 5 mL/min. The recombinant protein rPrP was found in fractions containing 40% acetonitrile. If the SEC eluate was stored at 4 °C for several days prior to RP-HPLC, then the recombinant protein was eluted in earlier fractions containing only 35% acetonitrile.

ELISA Assay. The anti-PrP monoclonal antibody 3F4 IgG was purified using a Pierce ImmunoPure (A) purification

kit followed by Fab fragmentation with a Pierce ImmunoPure Fab kit. The Fab was used at a concentration of 20 μg/mL in Tris-buffered saline (TBST). Costar 96-well plates were coated with 50 μL of 3F4 Fab fragment solution incubated overnight at 4 °C, washed twice with TBST, blocked with 3% BSA in TBST for 1 h at room temperature, and washed twice more. To each well was added 50 μL of antigen in TBST containing 0.05% TWEEN incubated at 37 °C for 1 h. After five washes, a 50 μL aliquot of a second monoclonal antibody (13A5) at a 1:5000 dilution in TBST was added to each well. After incubation for 1 h at 37 °C, the plate was washed five times and then treated for 1 h at 37 °C with goat anti-mouse Fc IgG (Pierce) diluted 1:500 that was alkaline phosphatase-conjugated in TBST, and then washed 9 times before being developed with 100 μL of a phosphate substrate solution prepared from tablets (Sigma). Light absorbance was monitored at 405 nm after 15, 30, 45, and 60 min. The readings were calibrated against HPLC-purified rPrP that had been amino acid analyzed.

Protein Determination, Amino Acid Analysis, Sequencing, and Protein Chemistry. Total protein was determined by Coomassie protein assay reagent. Concentration of rPrP was determined by ELISA assay or estimated by densitometry on crude samples. Carefully measured aliquots of the protein samples were dried in a Speed-Vac with norleucine as internal standard, 300 μL 6 N HCl containing 0.1% phenol was added, and the samples were hydrolyzed under vacuum at 110 °C for 24 h using a Millipore Picotag workstation. The hydrolysates were dried, reconstituted with Beckman sample buffer, and analyzed on a Beckman model 6300 amino acid analyzer. Automated Edman analysis was carried out using a Hewlett-Packard model G1005A protein sequencer; sequence interpretation was performed on a Digital Equipment VAX 8750 as described (Henzel et al., 1987). Cyanogen bromide digestion was carried out by standard methods (Lundblad, 1995). Carboxymethylation was carried out by incubating with 6 mM iodoacetic acid in 6 M Gdn-HCl at 4 °C for 1.5 h in the dark. Digestion with endoproteinase Lys-C was carried out in 50 mM Tris-HCl, pH 8.5/1 mM EDTA at 37 °C for 15 h; peptides were separated by RP-HPLC as described previously (Stahl et al., 1993) and analyzed by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (see below).

Mass Spectrometry. Three alternative protocols were employed for mass spectrometric analysis of the recombinant protein. After the initial survey by SDS−PAGE, the presence and purity of rPrP in RP-HPLC fractions was confirmed by MALDI-MS using a Perseptive Biosystems Voyager mass spectrometer. A 2 μL sample from each selected fraction was mixed with a large excess of α-cyano-2-hydroxycinnamic acid, dried onto a stainless steel target, and analyzed for the recombinant protein. The resolution of the MALDI technique was sufficient to unambiguously identify the presence of the desired product, but the mass measurement accuracy was not sufficient to confirm the presence or absence of the disulfide bond. After confirming the fractions containing the recombinant protein, 10 μL aliquots were analyzed by electrospray ionization (ESI) mass spectrometry using either a VG Instruments Platform or a Perkin Elmer Sciex APIII. Careful calibration of the ESI mass spectrometer allowed the molecular masses to be measured with sufficient accuracy and precision to identify the presence or absence of the disulfide bond. M_r values were calculated for the predicted amino acid sequence using the MacBioSpec program (Perkin-Elmer Sciex) using average rather than monoisotopic atomic masses.

Spectroscopy. A fraction of the reduced protein containing two free cysteines was refolded by dissolving in 8 M Gdn-HCl and diluting 100-fold into 20 mM Tris-HCl, pH 8.0, giving high yields of soluble protein. The final concentration of rPrP in the refolding buffer was 100 μg/mL. ESIMS showed it had become oxidized by formation of a disulfide bond although the solution contained no oxidizing agent other than dissolved air. Samples of the reduced protein and the refolded oxidized form were concentrated using a Centricon (Amicon) with a molecular weight cutoff of 10 000 Da. The buffer for the reduced protein was 10 mM MES, pH 6.5, whereas the oxidized form was concentrated in the refolding buffer described above.

CD spectra were recorded using solutions of concentration 0.5 mg/mL in 0.01 cm path length quartz cells at room temperature in a spectropolarimeter (Jasco, model 720) purged with dry nitrogen gas. Spectra were generated from accumulations of 10 individual scans, from which a corresponding spectrum of the buffer was subtracted. For FTIR spectroscopy, buffer exchange was carried out twice using ^2H_2O buffer, reducing the H_2O content to ∼0.25%. Samples of ∼3 μL were dried as thin films on a barium fluoride window. Spectra were recorded using an IR microscope (Perkin Elmer System 2000) continuously purged with dry nitrogen. Approximately 100 scans were accumulated for each spectrum until any signal due to water vapor was minimized relative to the background scan. Spectra were analyzed in the amide I region (1700−1600 cm^{-1}) for secondary structure (Byler & Susi, 1986; Goormaghtigh et al., 1990; Nguyen et al., 1995).

RESULTS

Expression of Recombinant Prion Protein. We chose the SHaPrP sequence for most of our expression studies for two reasons: (1) $SHaPrP^C$ and PrP^{Sc} are the most extensively studied prion proteins, and (2) the SHaPrP sequence differs substantially from human (Hu) PrP, implying a substantial barrier for the transmission of SHa prions to humans (Basler et al., 1986; Schätzl et al., 1995; Telling et al., 1995). A variety of secretion and intracellular expression vectors was examined, and SHaPrP sequences of various lengths were expressed. Full-length mature SHa PrP extends from residues 23 to 231 (Oesch et al., 1985), and this was our original goal for recombinant expression. Other constructs were truncated at the N-terminus to eliminate a highly basic region of the protein. N-terminally truncated species starting at residue 90 were also investigated as this is the start site for the infectious polypeptide PrP 27−30 (Prusiner et al., 1984). C-terminal truncation at 228 was another variable as termination at this site was identified as a post-translational modification in ∼10% of PrP 27−30 molecules (Stahl et al. 1993). Considerable effort has been devoted to expressing SHaPrP in bacterial systems since the highest levels of expression are frequently achieved with such systems (Table 1). These experiments were not all carried out under identical conditions and were not necessarily optimized. In most cases, we could detect recombinant protein on a Western blot, but in general, expression levels were too low to pursue and were lowest when the full-length molecule (amino acids 23−231) was expressed either alone or as a fusion protein.

Table 1: Summary of the Bacterial Expression Systems That Were Investigated To Express Various SHa PrP Sequences

vector system	hamster sequence expressed	*E. coli* strain	level of expression
pTAC-secretion	90−228	X90	very low expression
pTACTAC-secretion	90−228	X90	0.4 mg/mL soluble 4 mg/mL insoluble extremely unstable
pTACTAC-secretion	90−231	X90	same as pTACTAC 90−228
OMP A-secretion	23−231	SB221	not detectable; vectors were unstable
	28−231	SB221	
pET 11a-secretion	90−228	DE3 and PLYS	low expression
pMAL P2-secretion	fusion with *malE* and 23−231	TB1	low expression
pMAL C2-intracellular	fusion with *malE* and 23−231	TB1	low expression
STII TIR variant 4	23−231	27C7	very low expression
	27−231	27C7	low expression
	29−231	27C7	low expression
	90−231	27C7	high expression

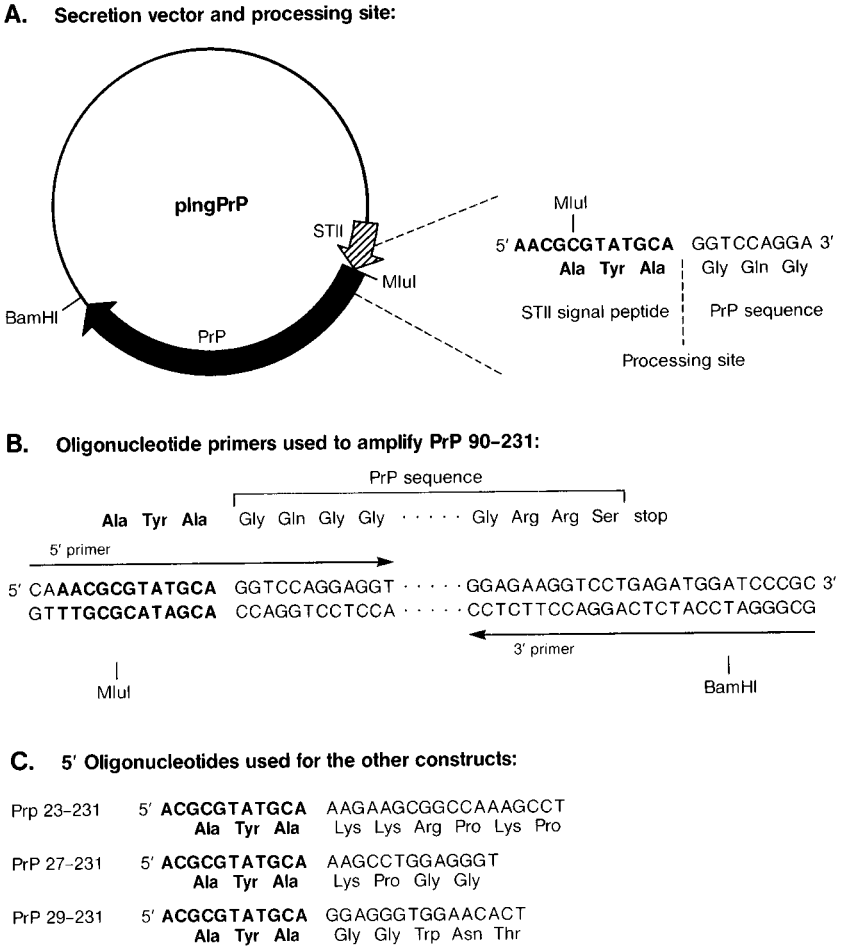

FIGURE 1: A. Secretion vector showing the PrP sequence inserted between the *Mlu*I and *Bam*HI sites and processing site for the removal of the STII signal peptide. B. Oligonucleotide primers employed to amplify Syrian hamster PrP 90−231. C. The 5′ oligonucleotide primers used to express other PrP constructs.

Expression of PrP residues 90−231 in the secretion vector pTAC (Higaki et al., 1989) was not detectable on a Coomassie-stained gel and was not significantly induced by the addition of IPTG. A similar problem was encountered with the pET11a vector (Novagen) in the strain PLSYS. Expression levels improved for the pET11a vector transformed into the strain DE3, but the level of expression was lower than that found in the pTACTAC vector (Muchmore et al., 1989). The least satisfactory expression was observed in the OmpA system (Ambion), from which no recombinant PrP could be detected by Western blot analysis, possibly due to a rearrangement event in the promoter region; this system was later discontinued by its suppliers. The fusion vector pMal E (New England Biolabs) also produced poor results; recombinant PrP fused to the *malE* gene could be detected in the intracellular vector only, but again the level of protein expression was significantly lower than that achieved with pTACTAC.

Until recently pTACTAC was the most promising secretion vector, using the bacterial signal peptide, his J (Higaki et al., 1989). Measures taken to optimize production from this vector included using different bacterial strains and media, varying the temperature and the time of induction, and even modifying the ribosomal binding site, but these failed to increase expression levels. Despite this, a purification protocol was developed, but the protein was extremely unstable, even with the addition of a cocktail of protease inhibitors. As determined by CD, the purified recombinant

Table 2: Enrichment of rPrP at Various Points in the Purification

sample	volume (mL)	total protein (mg)[a]	total rPrP (mg)[a]	rPrP as % of total protein
resuspended paste[b]	1000	5090[c]	75[d]	1.5
washed pellet after cell disruption[e]	1000	1780[c]	82[d]	4.6
guanidine supernatant [f,g]	200	1610[c]	75[d]	4.7
pooled SEC fractions[g,h]	667	300[i]	89[d]	30
RP-HPLC fractions[j]	5−10	54[k]	54[k]	≥95[l]

[a] All quantities are with reference to 100 g of crude protein paste from fermentation. [b] 100 g of crude protein paste resuspended in 1000 mL of buffer A. [c] Determined with Coomassie protein assay reagent (Pierce). [d] Determined by ELISA (see text). [e] After cell disruption, a 30 000*g* spin, washing with buffer A and an additional 30 000*g* spin. [f] After solubilization in buffer B and centrifugation at 14 000*g*. [g] Prior to assay the Gdn-HCl was removed with MeOH/$CHCl_3$/H_2O and an additional wash with MeOH. [h] After separation by SEC. Fractions were selected that contained predominantly rPrP. [i] Determined by densitometry on Coomassie-stained SDS−PAGE gels. [j] After separation by RP-HPLC. Fractions were selected that contained ≥95% pure rPrP as determined by SDS−PAGE, Western blots, and mass spectrometry. This represented ∼50% of the total rPrP eluted from the HPLC column. [k] Determined by amino acid analysis. [l] Confirmed by densitometry of SDS−PAGE gels, Edman sequencing, and mass spectrometry.

protein had a secondary structure similar to that of PrP^C (data not shown).

We next examined expression vectors based on the STII signal sequence. Initially the coding sequence for PrP amino acids 29−231 was inserted into the vector STII TIR variant 4, just downstream of the STII signal sequence. This vector is a derivative of phGH1 (Chang et al., 1987) with a modified translation initiation region (L. Simmons and D. Yansura, unpublished results). Low-level expression of the expected protein was detected by Coomassie-stained gel (data not shown) of whole cell pellets of induced cultures. Edman sequencing of the recombinant protein confirmed that the signal peptide was correctly removed. PCR was subsequently used to clone various DNA sequences corresponding to SHaPrP amino acids 23−231 (full-length mature), 27−231, 29−231, and 90−231 (Figure 1). The constructs were sequenced to confirm that mutations had not been introduced by PCR. Expression, judged by Coomassie-stained SDS−PAGE gels, was highest for 90−231 (Table 1) and decreased dramatically as the N-terminally charged amino acids were added, although expression could be detected on a Western blot in every case. N-terminal sequencing of the expressed proteins verified that the bacterial STII signal peptide was being processed correctly.

On the basis of the previous experiments, the recombinant construct corresponding to SHaPrP residues 90−231, designated rPrP, was selected as being most likely to yield high-level expression on a large scale (Figure 1). The secretion vector STII TIR variant 4 containing the truncated SHaPrP sequence was transformed into the protease deficient *E. coli* strain, 27C7. An overnight culture was used to inoculate the medium, in which expression of rPrP could be detected after 28 h. In cultures harvested by centrifugation at 36 h after inoculation, rPrP represented ∼1.5% of total protein and was expressed at a level of ∼75 mg/L of culture as determined by ELISA (Table 2). However, suppression of the ELISA may have reduced the apparent yield in the crude fractions; densitometry of SDS−PAGE gels suggested the rPrP constituted ∼3% of total protein. During expression it aggregated into insoluble particles in the periplasmic space. The insolubility of rPrP was exploited in developing a purification protocol.

Purification of rPrP. The purification protocol is outlined in Figure 2. The *E. coli* were collected by centrifugation at 10 000*g* for 30 min. The paste was resuspended in buffer A and centrifuged to remove soluble periplasmic proteins. The rPrP remained in the pellet, which was resuspended in the same buffer and passed through a microfluidizer in order to disrupt most of the cells. This disruption by microfluidization proved to be an essential step in order to release high yields of rPrP from the *E. coli* cells. Alternatives that were investigated released less PrP, and their use was not pursued. These included high salt concentrations, sonication in the presence of urea, and the use of a French press; repeated 1 min sonication bursts in the presence of 6 M urea was the most effective alternative as it released ∼30% of the rPrP, whereas all other methods released less. After cell disruption in the microfluidizer, the recombinant protein remained in the insoluble pellet. Centrifugation and washing

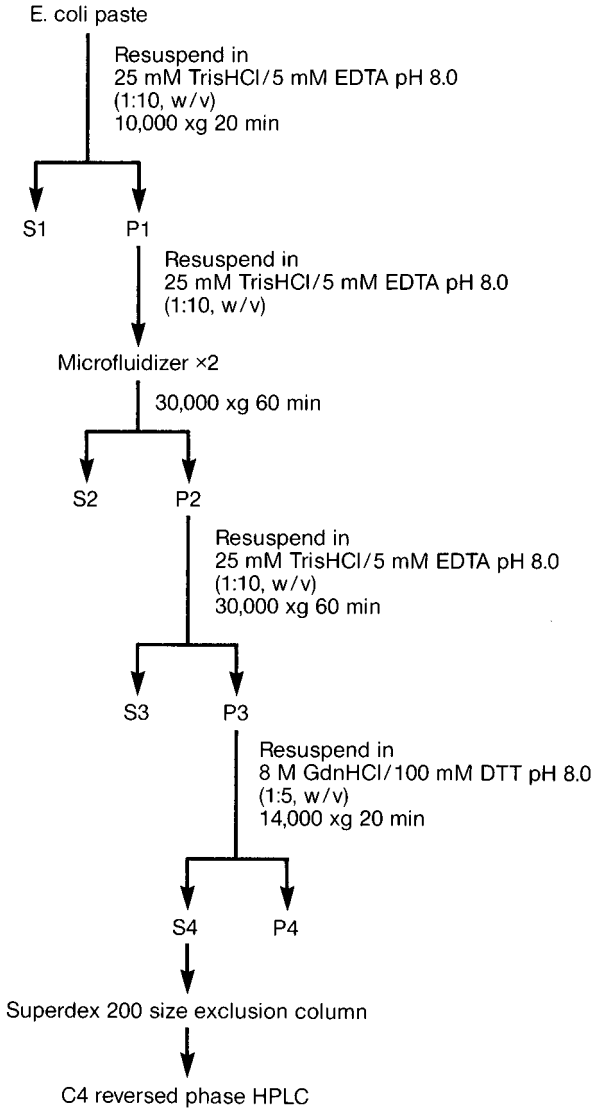

FIGURE 2: Purification protocol for rPrP.

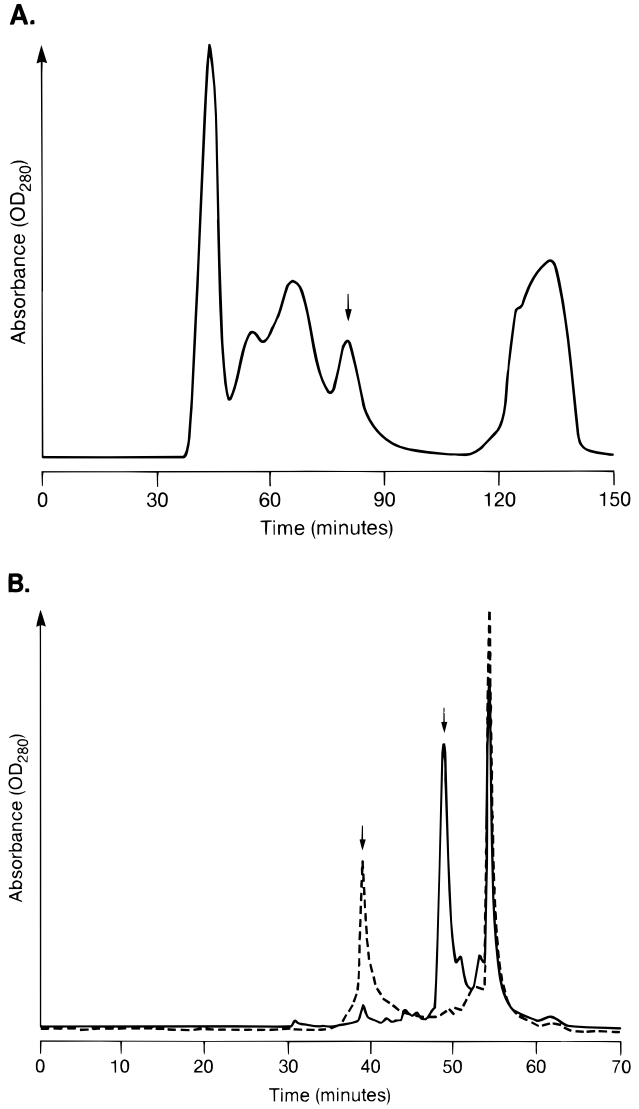

FIGURE 3: Chromatograms showing the purification of rPrP. A. SEC of the supernatant from the Gdn-HCl solubilization; arrow denotes peak that was collected. B. RP-HPLC of the pooled SEC fractions; solid trace shows the chromatogram for freshly isolated SEC fractions, whereas dashed trace corresponds to material that was stored for several days prior to RP-HPLC. In each case the arrow indicates the peak that was collected.

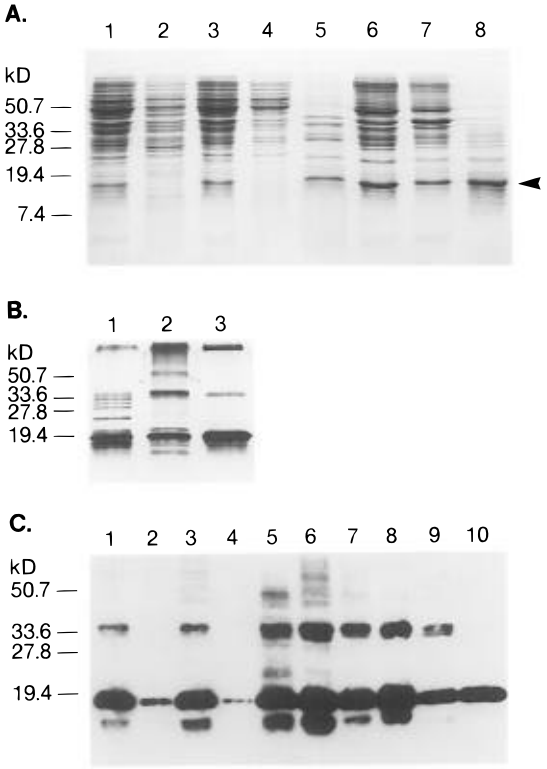

FIGURE 4: 15% polyacrylamide gels following the purification profile for rPrP. A. Coomassie-stained gel of early purification steps. 20 μg of total protein was loaded for each lane. Arrowhead indicates rPrP migration. Lane 1, starting material; lane 2, soluble periplasmic proteins; lane 3, pellet after releasing periplasmic proteins; lane 4, supernatant after microfluidizer; lane 5, pellet after microfluidizer; lane 6, supernatant after solubilization in buffer B; lane 7, pellet after solubilization in buffer B; lane 8, pooled fractions after SEC. B. Silver stain gel of the final purification step and refolded rPrP. Approximately 1 μg of total protein was loaded for each lane. Lane 1, pooled fractions from SEC; lane 2, reduced rPrP eluted from RP HPLC, the upper bands are impurities as described in the text; lane 3, refolded and oxidized rPrP. The band at ~33 kDa is due to a dimer. C. Western blot analysis using the monoclonal antibody 13A5 following the purification of rPrP. 200 ng of total protein was loaded for each lane. Lanes 1−8, as for A; lane 9, reduced rPrP eluted from RP HPLC; lane 10, refolded and oxidized rPrP; the band at ~33 kDa is due to a minor amount of dimer.

of the pellet removed more of the contaminating soluble cytoplasmic *E. coli* proteins. At this stage assays of the total pellet showed it contained ~5% rPrP. The pellet was then solubilized in buffer B, 50%−80% of the rPrP going into solution.

rPrP was separated from the high molecular weight *E. coli* proteins by SEC with elution in buffer C (Figure 3A) which by SDS−PAGE gave fractions in which the desired product constituted ~30% of the total protein. Buffers B and C were able to keep rPrP in solution since each contained Gdn-HCl; furthermore, both buffers were compatible with the two chromatographic procedures. Alternative buffers containing urea were employed, but these were found to cause carbamylation of rPrP. Urea-containing buffers also required the presence of ≥1 M sodium chloride to prevent aggregation; if the SEC was performed at ≤6 °C with urea, aggregation could not be prevented.

The eluate from the SEC column was further purified by RP-HPLC. Electrophoresis (Figure 4A,B) and immunoblots (Figure 4C) confirmed the presence and purity of rPrP which appeared as a single band in fractions eluted from the RP-HPLC column. MALDI-MS confirmed the purity and approximate M_r of these fractions. When the SEC eluate was immediately subjected to further separation, the rPrP was eluted at 49 min in fractions containing 40% acetonitrile (Figure 3B). By ESI-MS this species had a molecular mass of 16 243.9 Da (Figure 5A), corresponding to the reduced form without a disulfide bond (calculated M_r, 16 243.2 Da). This was confirmed by derivatization with iodoacetic acid which induced a shift in the molecular mass to 16 360.0 Da (Figure 5C) due to the addition of two carboxymethyl groups (calculated M_r, 16 360.3). In contrast, rPrP in SEC fractions that had been stored at 4 °C for ~1 week eluted from the RP-HPLC at 39 min in fractions containing ~35% acetonitrile. This species had a molecular mass of 16 241.6 (Figure 5B), corresponding to oxidized rPrP containing a disulfide bond (calculated M_r, 16 241.2 Da), and proved to be resistant to carboxymethylation. Evidence for the disulfide bond was also provided by the MALDI-MS observation of two linked peptides containing the two cysteines, residues 110−185 and 205−220, in an endoproteinase Lys-C digest (data not shown). Thus, the different RP-HPLC retention times corresponded to the reduced and oxidized forms of

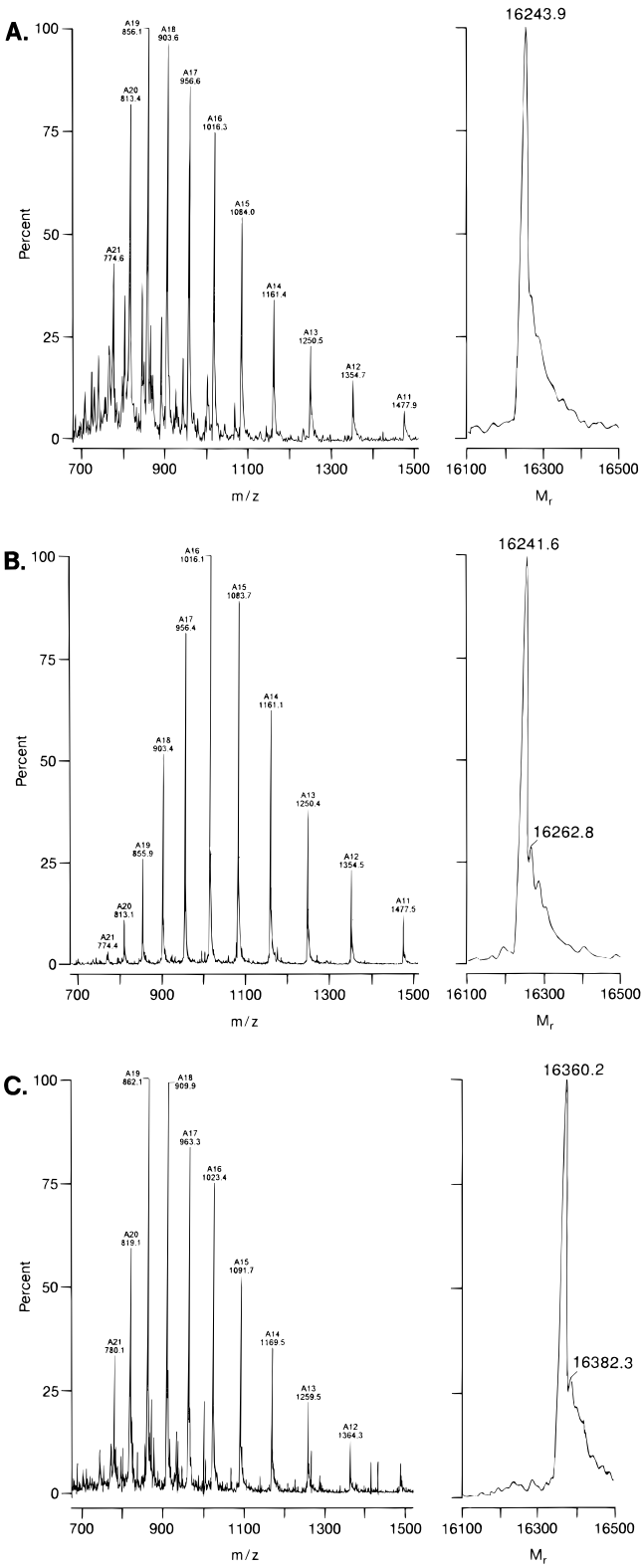

FIGURE 5: ESI mass spectrometry of RP-HPLC purified rPrP. In each case, the raw data is shown in the left hand panel and a computer-generated transform of the same data is shown on the right. The abscissa for the raw data is mass/charge (atomic mass units), whereas for the transforms it is M_r (Da). A. Mass spectrum of the late eluting peak indicated by the arrow in Figure 3B, solid trace. B. Mass spectrum of the early eluting peak indicated by the arrow in Figure 3B, dashed trace. C. Mass spectrum of the same material as in panel A after carboxymethylation with iodoacetic acid.

the protein; it was shown that these two forms could be interchanged by either reduction with fresh DTT or oxidation with air.

Table 3: Amino Acid Analysis of rPrP[a]

amino acid	amount (pmol)	mol %	residues	
			sample	theoretical
Asp	1328	13.73	19	19
Thr	805	8.33	11.9	12
Ser	322	3.33	5	5
Glu	1271	13.15	18	17
Gly	897	9.28	13	14
Ala	503	5.2	7.1	8
Val	586	6.06	8.3	9
Met	577	5.97	8.2	9
Ile	306	3.16	4	4
Leu	148	1.53	2	2
Nle	204	0	0	0
Tyr	684	7.07	10	10
Phe	227	2.34	3	3
His	377	3.89	5	5
Lys	577	5.96	8	8
Arg	724	7.49	10	8
Pro	339	3.5	5	5
Cys	0	0	0	2

[a] As described in the text. The detection and analysis of cysteine by this protocol frequently gives lower yields than determined by other methods. Norleucine (Nle) was added as an internal standard.

RP-HPLC fractions containing reduced rPrP that eluted late were sometimes contaminated with co-eluting proteins, as depicted in the ESI spectrum (Figure 5A, left-hand panel), although these contaminants could be avoided by selective collection of the fractions. Contaminants were infrequent in the fractions containing oxidized rPrP (Figure 5B); thus, RP-HPLC of the oxidized rPrP typically gave superior purification, although minor differences in the timing of the collection of fractions produced differences in the purity of the fractions. In order to ensure that the final product was >95% pure, the fractions that were collected generally represented only ~50% of the rPrP.

Determination of the yield of rPrP proved to be difficult due to the high levels of contaminating protein in the early steps and insolubility of rPrP (Table 2). ELISA assays suggested the concentration of rPrP in the starting material to be 75 mg/L of culture, whereas densitometry indicated that this should be ~3% of the total protein, or 150 mg. As the purification proceeded, the total protein declined as measured by BCA or Coomassie assays while rPrP increased as determined by ELISA. This was attributed to suppression of the rPrP signal in the ELISA of the crude fractions which became less significant in the later stages of purification. After the final RP-HPLC step, the rPrP was determined by amino acid analysis; typically, 1 L of fermentation medium yielded ~50 mg of purified recombinant protein. We estimate the yield of purified rPrP to be ~60%, but it may be as low as ~30% if our determinations of rPrP in homogenates are in error as noted above.

Primary Structure of rPrP. DNA sequencing confirmed the correct oligonucleotide sequence in the DNA construct. Amino acid analysis of rPrP was consistent with the predicted composition of this polypeptide (Table 3). Edman sequencing of the intact rPrP identified the first 79 residues, which together with sequencing of peptides derived from cyanogen bromide cleavage confirmed the sequence of more than 70% of the protein (Figure 6). No modifications of residues were observed other than a small degree of Trp oxidation. An aliquot of rPrP was digested with endoproteinase Lys-C, and the resulting peptides were separated by RP-HPLC. When these peptides were analyzed by MALDI-MS, all of the anticipated peptides terminating in Lys were observed with

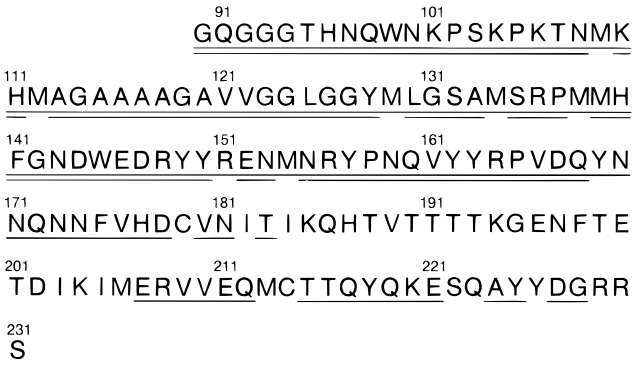

FIGURE 6: Amino acid residues of PrP 90−231 that were confirmed by Edman sequencing of the intact protein and of peptides generated from cyanogen bromide digests. Double underlining indicates residues that were identified in both the protein and the peptides; single underlining indicates residues that were identified in either the protein or the peptides.

the exception of the hydrophilic tetrapeptide TNMK which was not retained on the column (data not shown). However, the residues of this tetrapeptide were identified by Edman sequencing. The foregoing results combined with the M_r values obtained by ESI-MS argue that the primary structure of rPrP is correct.

Secondary and Higher Order Structure. Differences in the higher order structures of the reduced and oxidized forms of rPrP were suggested by the ESI mass spectral profiles. The distribution of charge states for the reduced form maximized at the attachment of 19 protons (Figure 5A), whereas the most abundant species in the ESI spectrum of the oxidized form corresponded to 16 protons (Figure 5B). These findings are consistent with the reduced form having a more open structure that was more accessible to protonation than the more compact oxidized form.

Both forms of rPrP remained soluble in the acidic medium resulting from concentration of the RP-HPLC buffer by vacuum centrifugation, which removed most of the acetonitrile and concentrated the TFA. Buffer exchange for the reduced rPrP into 10 mM MES, pH 6.5, resulted in higher solubility than in Tris buffers at pH 8 or higher. The CD spectrum of the reduced protein revealed a mixture of random coil and β-sheet, the latter yielding a minimum at 217 nm (Figure 7A), whereas the FTIR spectrum of the reduced rPrP revealed the presence of both α-helix (maximum at 1660 cm^{-1}) and β-sheet (maxima at 1626 and 1695 cm^{-1}) (Byler & Susi, 1986; Goormaghtigh et al., 1990) (Figure 7B). Although some properties of reduced rPrP are similar to PrP^{Sc}, rPrP was sensitive to proteinase K digestion (data not shown). In contrast to reduced rPrP, the oxidized protein was soluble in Tris buffers at pH 8.0. Refolding the reduced protein under oxidizing conditions gave a protein even more soluble than that derived from the early-eluting HPLC fractions. The CD spectrum of oxidized rPrP showed minima at 208 and 222 nm, indicative of an α-helical structure (Figure 7A). This was confirmed by the FTIR spectrum that showed only the α-helical component and no β-sheet. A more detailed study of the refolding of rPrP and an analysis of the relative amounts of the various secondary structural elements based on CD and FTIR spectra will be reported later (H. Zhang et al., unpublished data).

DISCUSSION

The profound conformational change that PrP^{C} undergoes when it is converted into PrP^{Sc} in the absence of any known chemical modification presents a structural conundrum (Pan et al., 1993; Stahl et al., 1993). Although existing PrP^{Sc} appears to act as ligand in the production of nascent PrP^{Sc} from PrP^{C}, the conformational transition seems to be unprecedented. To learn more about this process, it is important to establish the secondary and tertiary structures of the two PrP isoforms.

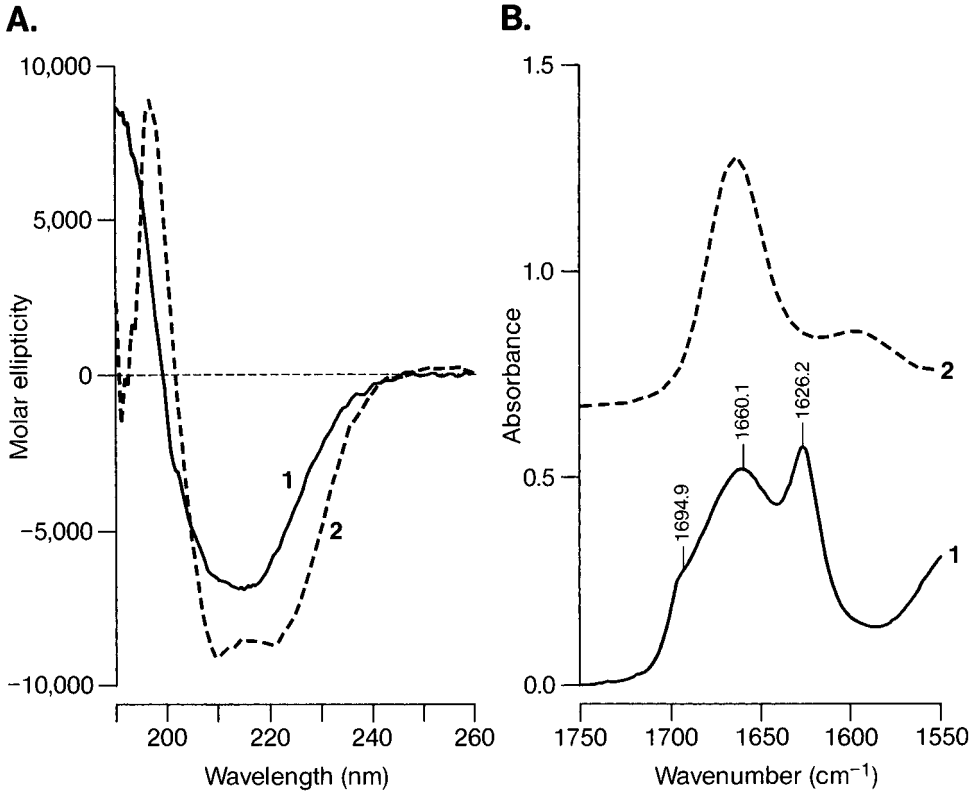

FIGURE 7: A. CD spectra: solid trace, the reduced protein; dashed trace, the refolded and oxidized form. B. Amide I band of the FTIR spectra: solid trace, the reduced protein; dashed trace, the refolded and oxidized form.

Because of the difficulties associated with obtaining substantial amounts of PrP^C and the insolubility of PrP^{Sc}, we explored several approaches to investigating the structures of these PrP isoforms. First, we used molecular modeling to predict the tertiary structures of both PrP^C and PrP^{Sc} in conjunction with spectroscopic and biological data (Huang et al., 1994). However, such predictions require physical data from either X-ray crystallography or NMR to assess their validity. Second, we constructed synthetic peptides corresponding to regions of putative secondary structure and determined their structures by spectroscopy and NMR (Gasset et al., 1992; Zhang et al., 1995). Data from a peptide of 56 amino acids in length corresponding to PrP residues 90−145 suggest that the assignments of the first two helices in the four-helix bundle model of PrP^C are likely to be correct (Huang et al., 1994; Zhang et al., 1995). Third, we examined numerous expression systems in order to produce large quantities of a PrP polypeptide as reported here (Table 1).

Initially we attempted to obtain an expression system that would allow purification of a native protein under non-denaturing conditions. Early attempts to purify the protein in this state were not fruitful, and we eventually resorted to extraction and purification under strongly denaturing conditions. The product required subsequent refolding in order to carry out any determination of the secondary or tertiary structure. If apparent at the outset that denaturation and refolding would be necessary, a different strategy most likely would have been adopted, with intracellular expression and extraction from inclusion bodies.

The difficulties in expressing PrP at high levels appear to be associated with the amino acids found at the N-terminus of mature PrP (Table 1). Using the bacterial secretion vector, STII TIR variant 4, only low levels of expression of PrP were found when the charged or basic amino acids at positions 22−29 were included in the construct. This finding may account for earlier difficulties associated with expressing PrP in bacterial systems. Expression of rPrP containing residues 90−231, which corresponds to the sequence of PrP 27−30, gave the highest levels of expression. On the basis of N-terminal sequencing, the bacterial signal peptide was removed, indicating that rPrP was processed correctly during this translocation into the periplasm where it formed insoluble aggregates. Mass spectrometry and Edman sequencing established that the primary structure of rPrP was as expected.

During purification, we were initially puzzled by the changes in the time of elution of rPrP from RP-HPLC. This unexpected behavior of rPrP was found to result from different oxidation states of the disulfide bond generating two different conformational states, one of which is characterized by a high α-helical content and the other by a high β-sheet content not unlike PrP^C, and PrP^{Sc}, respectively. While the two conformations of rPrP seem to be dependent on the oxidation state of the disulfide bond, it seems doubtful that this is the case for PrP^{Sc} formation. Although both PrP^C and PrP^{Sc} possess a disulfide bond, high concentrations of neither DTT nor β-mercaptoethanol altered infectivity of prion rods (Prusiner et al., 1980; Turk et al., 1988). Our finding that reduced rPrP has a high β-sheet content raises the possibility that redox processes might participate transiently in the formation of PrP^{Sc}. It is certainly possible that a reduced intermediate that represents a transition state in the conversion of PrP^C into PrP^{Sc} exists. Whether reduction of the disulfide bond in PrP^C will induce it to adopt the conformation of PrP^{Sc} remains to be determined.

Preliminary studies have shown that protease resistance can be induced in rPrP by incubation with either PrP^{Sc} or the synthetic peptide corresponding to PrP residues 90−145 (K. Kaneko et al., unpublished data). Bioassays in Syrian hamsters are in progress to determine if either the oxidized or reduced forms of rPrP will induce scrapie. The availability of a large segment of PrP which corresponds to the amino acid sequence of PrP 27−30 should facilitate the search for conditions under which this recombinant protein can be folded to induce prion formation and neurodegeneration.

REFERENCES

Barry, R. A., & Prusiner, S. B. (1986) *J. Infect. Dis. 154*, 518−521.

Basler, K., Oesch, B., Scott, M., Westaway, D., Wälchli, M., Groth, D. F., McKinley, M. P., Prusiner, S. B., & Weissmann, C. (1986) *Cell 46*, 417−428.

Brown, P., Gibbs, C. J., Jr., Rodgers-Johnson, P., Asher, D. M., Sulima, M. P., Bacote, A., Goldfarb, L. G., & Gajdusek, D. C. (1994) *Ann. Neurol. 35*, 513−529.

Byler, D. M., & Susi, H. (1986) *Biopolymers 25*, 469−487.

Carter, P., Kelley, R. F., Rodrigues, M. L., Snedecor, B., Covarrubias, M., Velligan, M. D., Wong, W. L., Rowland, A. M., Kotts, C. E., Carver, M. E., Yang, M., Bourell, J. H., Shepard, H. M., & Henner, D. (1992) *Bio/Technology 10*, 163−167.

Caughey, B., Race, R. E., Vogel, M., Buchmeier, M. J., & Chesebro, B. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4657−4661.

Caughey, B. W., Dong, A., Bhat, K. S., Ernst, D., Hayes, S. F., & Caughey, W. S. (1991) *Biochemistry 30*, 7672−7680.

Chang, C. N., Rey, M., Bochner, B., Heyneker, H., & Gray, G. (1987) *Gene 55*, 189−196.

Cohen, F. E., Pan, K.-M., Huang, Z., Baldwin, M., Fletterick, R. J., & Prusiner, S. B. (1994) *Science 264*, 530−531.

Endo, T., Groth, D., Prusiner, S. B., & Kobata, A. (1989) *Biochemistry 28*, 8380−8388.

Gajdusek, D. C. (1977) *Science 197*, 943−960.

Gasset, M., Baldwin, M. A., Lloyd, D., Gabriel, J.-M., Holtzman, D. M., Cohen, F., Fletterick, R., & Prusiner, S. B. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10940−10944.

Gasset, M., Baldwin, M. A., Fletterick, R. J., & Prusiner, S. B. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1−5.

Glenner, G. G., Eanes, E. D., & Page, D. L. (1972) *J. Histochem. Cytochem. 20*, 821−826.

Goormaghtigh, E., Cabiaux, V., & Ruysschaert, J.-M. (1990) *Eur. J. Biochem. 193*, 409−420.

Henzel, W. J., Rodriguez, H., & Watanabe, C. (1987) *J. Chromatogr. 404*, 41−52.

Higaki, J. N., Evnin, L. B., & Craik, C. S. (1989) *Biochemistry 28*, 9256−9263.

Hope, J., Morton, L. J. D., Farquhar, C. F., Multhaup, G., Beyreuther, K., & Kimberlin, R. H. (1986) *EMBO J. 5*, 2591−2597.

Huang, Z., Gabriel, J.-M., Baldwin, M. A., Fletterick, R. J., Prusiner, S. B., & Cohen, F. E. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 7139−7143.

Lee, C. H., Moseley, S. L., Moon, H. W., Whipp, S. C., Gyles, C. L., & So, M. (1983) *Infect. Immun. 42*, 264−268.

Lundblad, R. L. (1995) *Techniques of Protein Modification*, CRC Press, Boca Raton, FL.

Masters, C. L., Gajdusek, D. C., & Gibbs, C. J., Jr. (1981) *Brain 104*, 535−558.

Medori, R., Tritschler, H.-J., LeBlanc, A., Villare, F., Manetto, V., Chen, H. Y., Xue, R., Leal, S., Montagna, P., Cortelli, P., Tinuper, P., Avoni, P., Mochi, M., Baruzzi, A., Hauw, J. J., Ott, J., Lugaresi, E., Autilio-Gambetti, L., & Gambetti, P. (1992) *N. Engl. J. Med. 326*, 444−449.

Meyer, R. K., McKinley, M.P., Bowman, K. A., Braunfeld, M. B., Barry, R. A., & Prusiner, S. B. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2310−2314.

Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., & Dahlquist, F. W. (1989) *Methods Enzymol. 177*, 44−73.

Nguyen, J., Baldwin, M. A., Cohen, F. E., & Prusiner, S. B. (1995) *Biochemistry 34*, 4186−4192.

Oesch, B., Westaway, D., Wälchli, M., McKinley, M. P., Kent, S. B. H., Aebersold, R., Barry, R. A., Tempst, P., Teplow, D. B., Hood, L. E., Prusiner, S. B., & Weissman, C. (1995) *Cell 40*, 735−746.

Pan, K.-M., Stahl, N., & Prusiner, S. B. (1992) *Protein Sci. 1*, 1343−1352.

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. U.S.A. 90*, 10962−10966.

Picken, R. N., Mazaitis, A. J., Mass, W. K., Rey, M., & Heyneker, H. (1983) *Infect. Immun. 42*, 269−275.

Prusiner, S. B. (1989) *Annu. Rev. Microbiol. 43*, 345−374.

Prusiner, S. B. (1991) *Science 252*, 1515−1522.

Prusiner, S. B. (1994) *Annu. Rev. Microbiol. 48*, 655−685.

Prusiner, S. B., Groth, D. F., Cochran, S. P., Masiarz, F. R., McKinley, M. P., & Martinez, H. M. (1980) *Biochemistry 19*, 4883−4891.

Prusiner, S. B., McKinley, M. P., Bowman, K. A., Bolton, D. C., Bendheim, P. E., Groth, D. F., & Glenner, G. G. (1983) *Cell 35*, 349−358.

Prusiner, S. B., Groth, D. F., Bolton, D. C., Kent, S. B., & Hood, L. E. (1984) *Cell 38*, 127−134.

Rogers, M., Taraboulos, A., Scott, M., Borchelt, D., Serban, D., Gyuris, T., & Prusiner, S. B. (1992) in *Prion Diseases of Humans and Animals* (Prusiner, S. B., Collinge, J., Powell, J., & Anderton, B., Eds.) pp 457−469, Ellis Horwood, London.

Safar, J., Roller, P. P., Gajdusek, D. C., & Gibbs, C. J., Jr. (1993) *J. Biol. Chem. 268*, 20276−20284.

Schätzl, H. M., Da Costa, M., Taylor, L., Cohen, F. E., & Prusiner, S. B. (1995) *J. Mol. Biol. 245*, 362−374.

Scott, M., Butler, D., Bredesen, D., Wälchli, M., Hsiao, K., & Prusiner, S. B. (1988) *Protein Eng. 2*, 69−76.

Stahl, N., Borchelt, D. R., Hsiao, K., & Prusiner, S. B. (1987) *Cell 51*, 229−240.

Stahl, N., Baldwin, M. A., Burlingame, A. L., & Prusiner, S. B. (1990) *Biochemistry 29*, 8879−8884.

Stahl, N., Baldwin, M. A., Teplow, D., Hood, L. E., Beavis, R., Chait, B., Gibson, B., Burlingame, A. L., & Prusiner, S. B. (1992) in *Prion Diseases and Animals* (Prusiner, S. B., Collinge, J., Powell, J., & Anderton, B., Eds.) pp 361−379, Ellis Horwood, London.

Stahl, N., Baldwin, M. A., Teplow, D. B., Hood, L., Gibson, B. W., Burlingame, A. L., & Prusiner, S. B. (1993) *Biochemistry 32*, 1991−2002.

Tateishi, J., Doh-ura, K., Kitamoto, T., Tranchant, C., Steinmetz, G., Warter, J. M., & Boellaard, J. W. (1992) in *Prion Diseases of Humans and Animals* (Prusiner, S. B., Collinge, J., Powell, J., & Anderton, B., Eds.) pp 129−134, Ellis Horwood, London.

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.

Towbin, H., Staehelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350−4354.

Turk, E., Teplow, D. B., Hood, L. E., & Prusiner, S. B. (1988) *Eur. J. Biochem. 176*, 21−30.

Weiss, S., Famulok, M., Edenhofer, F., Wang, Y.-H., Jones, I. M., Groschup, M., & Winnacker, E.-L. (1995) *J. Virol. 69*, 4776−4783.

Zhang, H., Kaneko, K., Nguyen, J. T., Livshits, T. L., Baldwin, M. A., Cohen, F. E., James, T. L., & Prusiner, S. B. (1995) *J. Mol. Biol. 250*, 514−526.

BI952965E