REFINEMENT OF UNDER-DETERMINED LOOPS OF HUMAN PRION PROTEIN BY DATABASE-DERIVED DISTANCE CONSTRAINTS

By

Feng Cui

Kriti Mukhopadhyay

Won-Bin Young

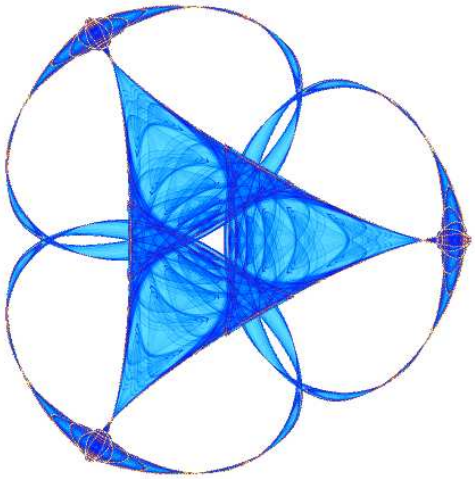
Robert L. Jernigan

and

Zhijun Wu

IMA Preprint Series # 2178

(November 2007)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

400 Lind Hall

207 Church Street S.E.

Minneapolis, Minnesota 55455–0436

Phone: 612-624-6066 Fax: 612-626-7370

URL: http://www.ima.umn.edu

Refinement of Under-Determined Loops of Human Prion Protein by Database-Derived Distance Constraints

Feng Cui[1*], Kriti Mukhopadhyay[2*], Won-Bin Young[4*], Robert L. Jernigan[1,3], Zhijun Wu[1,2]

[1]*Program on Bioinformatics and Computational Biology*
[2]*Department of Mathematics*
[3]*Department of Biochemistry, Biophysics, and Molecular Biology*
Iowa State University
[4]*NewLink Genetics Inc., University Research Park*
Ames, Iowa 50010 U.S.A

Email: fcui@nci.nih.gov, zhijun@iastate.edu

Abstract

Computational simulations of the conversion from the normal cellular prion (PrP^{c}) to the scrapie prion (PrP^{Sc}) are usually based on the structures determined by NMR because of the difficulties in crystallizing prion protein. Due to insufficient experimental restraints, a biologically critical loop region in PrP^{c} (residues 167-171), which is a potential binding site for Protein X, is under-determined in most mammalian species. Here, we show that by adding information about distance constraints derived from a database of high-resolution protein structures, this under-determined loop as well as other secondary structural elements of the E200K variant of human prion protein ($hPrP^{c}$), a disease-related isoform, can be refined into more realistic structures in the structural ensemble with improved quality and increased accuracy. In particular, the ensemble becomes more compact after the refinement and the percentage of residues in the most favourable region of the Ramachandran diagram is increased to about 90% in the refined structures from the 80 to 85% range in the previously reported structures. Our results not only provide significantly improved structures of the prion protein and hence would facilitate insights into its conversion in the spongiform encephalopathies, but also demonstrate the strong potential for using databases of known protein structures for structure determination and refinement.

1. Introduction

Spongiform encephalopathies, or prion diseases, are a group of neurodegenerative diseases in mammalian species characterized by a progressive vacuolation of brain tissue, amyloid protein deposits, and astrogliosis . Specific examples of the diseases include scrapie in sheep, transmissible mink encephalopathy in mink (TME), chronic wasting disease in mule deer (CWD), bovine spongiform encephalopathy (BSE) in cows, and Gerstmann-Sträussler-Scheinker disease (GSS), fatal familial insolmnia (FFI), kuru, and Alpers syndrome in humans [2-3]. The pathogenesis of the diseases is associated with accumulation of the infectious "scrapie" form of prion protein (PrP^{Sc}) in brain tissue, which is transformed from its normal cellular form (PrP^{c}) [2-5]. The two forms of prion protein are distinct in many aspects. For example, PrP^{c} is α-helix-rich, soluble, and highly sensitive to proteinase K digestion, whereas PrP^{Sc} is β-sheet-rich, insoluble, and resistant to proteinase K digestion [6-11]. The striking difference in secondary structures implies a major conformational transition from PrP^{c} to PrP^{Sc}, which has been considered as a key process involved in the pathogenesis of prion diseases. However, this so-called 'prion-protein-only' hypothesis is still not fully validated, and the mechanism of the conformational conversion is still unclear. One of the key obstacles for understanding the details of the prion conformational conversion is that the PrP^{Sc} sample is hard to purify for biochemical and structural characterization. Moreover, the normal, cellular isoform of PrP (PrP^{c}) is also difficult for high-resolution spectroscopic or crystallographic study [12], especially for certain

[1*] These authors contributed equally to the work

biologically critical loop regions (see below). Thus, high-quality prion structures are urgently needed to provide better insights into its transition process.

The E200K variant of the human prion protein (denoted as E200K hereafter) contains a point mutation from negative-charged glutamic acid (E) to positive-charged lysine (K) at the 200th codon. This mutation is linked to the human familiar prion diseases [13]. However, this charge-switching mutation that may alter surface electrostatic potential seems not bring dramatic changes in 3-D structures. Comparison between the E200K NMR structures [14] and other wild-type prion NMR [15-16] and X-ray [17-18] structures from different species shows that they agree very well in several stable secondary elements: Helix 1 (residues 144-153), Helix 2 (residues 172-194), and Helix 3 (residues 200-227), and two β-sheets (residues 129-131 and 161-163). However, they differ significantly at two loop regions, Loop 1 (residues 167-171) and Loop 2 (residue 195-199). The Loop 1 region is biologically critical because it is shown to act as a binding site for "Protein X," which might function as a mediator for the transition from PrP^{c} to PrP^{Sc} [19-23]. Whether the E200K mutation can alter the conformations of this loop region still remains unclear.

Nevertheless, the Loop 1 is under-determined in E200K. There are only a few experimental NOE (Nuclear Overhauser Effect) constraints available in this region (see Table 1), much fewer than other regions of the same protein. The analysis of PROCHECK [24-25] also showed that none of the residues in Loop 1 fall within the most favored regions of the Ramachandran plot (see Results). Thus, alternative constraints that are derived from structural analysis of high resolution X-ray structure may be well suited for this particular case.

The enhancement of NMR structures can be achieved by adding theoretical conformational constraints such as dihedral angles and inter-atomic distances based on statistical analysis of databases of high-resolution protein structures. In particular, it has been shown that inter-atomic distance constraints can improve NMR structures with increased precision and accuracy , and can function in an equivalent way as some experimental NMR restraints such as torsion angle restraints, without compromising the quality of the NMR structures . Moreover, these distance constraints impose literally no extra cost on NMR structure refinement [28-29].

In this work, we used a selected set of inter-atomic distance constraints between heavy atoms as additional constraints to refine the E200K variant of human prion protein . Our results showed that the critical Loop 1 region as well as some other secondary elements of E200K can be significantly improved in terms of the precision and accuracy and the Ramachandran plots of the structures by using database-derived distance constraints. The improved structures provide additional structural information that can be taken as a better starting point for the studies of prion protein conversion in the spongiform encephalopathies. To the best of our knowledge, this is the first evidence that the distance constraints derived from a structural database can be used to optimize the under-determined regions of biologically important proteins.

2. Methods

Inter-atomic distances can be categorized according to the specific atom pairs, residue pairs, and sequential residual separations. Different types of distances are subject to different statistical distributions in the structural databases, which have been employed to construct various statistical potentials for contact determination, inverse folding, structure alignment, X-ray structure refinement, etc [30-35]. A set of X-ray structures (~3000) with similarity of 90% or less and resolutions of 2.0 Å or better have been collected from the Protein Data Bank (PDB) and used to obtain the statistical distributions of distances of these different types. To generate a minimal amount of required data, the distances between two heavy atoms N, C^{α}, C', C^{β}, and O in two residues either sequentially adjacent or separated by one residue have been considered, although they could be extended to other more general types. The distribution of each type of distances was calculated by counting the number of occurrences of the distances within a set of distance intervals. Two distance distributions are given as examples in Fig. 1, one for the distances between the C atom of ARG at position *i* and the O of ILE at neighboring position *i* + 1 (Fig. 1(a)) and the other for the distances between the C^{β} atom of ALA at position *i* and the N atom of LEU at the next-nearest neighbor position *i* + 2 (Fig. 1(b)). Two hundred distance intervals are specified in the horizontal direction. The length of each interval is equal to 0.1 Å. The ordinate values show the frequencies of the distances in the corresponding distance intervals. The means μ and standard deviations σ of the distributions have been used to specify the range constraints for the corresponding distances to be between $\mu - 2\sigma$ and $\mu + 2\sigma$.

The NMR experimental data of E200K (residues 125-228) (Table 1) was downloaded from BioMagResBank (BMRB). The structure was then refined using NMR experimental constraints plus additional database-derived distance constraints. The standard torsion angle dynamic simulated annealing protocol implemented in CNS was used for the refinement.

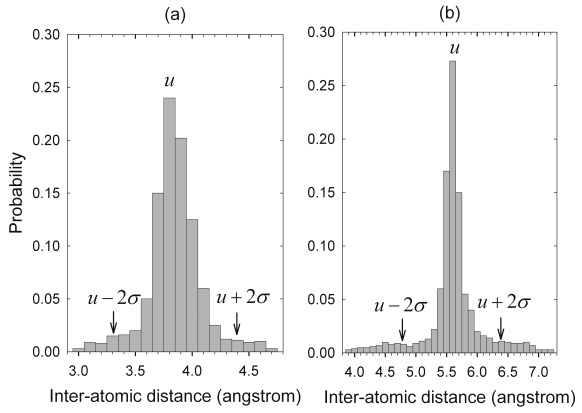


Fig 1. Sample distance distribution functions. (a) Distances between the C atom of ARG at position *i* and the O atom ILE at position *i* + 1. (b) Distances between the C^{β} atom of ALA at position *i* and the N atom of LEU at position *i* + 2.

3. Results

3.1. Refining NMR structures of the disease-related hPrP^{c} (E200K)

The NMR structure of the E200K variant of the human prion protein was refined by adding the database-derived distance constraints. Two ensembles of 50 accepted structures were collected: one was obtained with experimental constraints (denoted as <E200K>^{NMR}) and the other with experimental data plus database-derived distance constraints (denoted as <E200K>$^{NMR+D}$). The statistics of resulting ensembles in terms of the agreement with the experimental constraints, optimal covalent geometry, and the local and global potential energy are given in Table 2-4.

The quality analysis by PROCHECK on the average and energy-minimized structures[2] and the lowest energy structures of the ensembles showed a significantly higher percentage (89.6%) of residues in most favorable regions of the Ramachandran plot of the structures in <E200K>$^{NMR+D}$ compared to the 85.4% of these residues in <E200K>^{NMR} (Table 5). Note that the latter percentage (85.4%) is consistent with what was reported by Zhang *et al.* for their experimental structures (85.7%) . This increase in the percentage of residues in the most favorable regions indicated an improvement of the structural conformations coming from the use of the database-derived distance constraints.

[2]It is obtained by minimizing the total energy of the structure using the CNS energy minimization routine with the averaged structure of the ensemble as the starting structure.

Table 1. Experimental restraints of E200K

Experimental Restraints of Loop 1 (residue 167-171)

Residue	NOE†	Torsion	H-bond	J-coupling‡
167	3	0	0	0
168	3	0	0	0
169	0	0	0	0
170	1	0	0	0
171	21	1	2	0

Experimental Restraints of Loop 2 (residue 195-199)

Residue	NOE	Torsion	H-bond	J-coupling
195	17	0	0	0
196	45	0	0	1
197	34	0	0	0
198	76	0	2	1
199	32	1	0	1

	NOE	Torsion	H-bond	J-coupling
Total	3157	177	96	44
Per Res.	29.8	1.7	0.9	0.4

Note: Total number of Residue is 106.

†Total distance restraints, ‡J HNHA-coupling constants

3.2. Comparison between the refined E200K structures and the wild-type hPrP^{c} NMR and X-Ray structures

The residue-residue comparison of the average and energy-minimized structures of <E200K>$^{NMR+D}$ and <E200K>^{NMR} with the experimentally-determined NMR and X-ray structures of wild-type PrP^{c} was conducted. The residual RMSD values between the refined E200K structures and wild-type structures 1QM0, 1HJM, 1UW3, and 1I4M were plotted in Fig. 2 (see Fig. 2 legends for details).

The residue RMSD values of Loop 1 and 2 in all these plots are all significantly higher compared to other regions of the structure (> 4 Å), indicating that the loop regions are more flexible compared to the helical and sheet regions. The <E200K>$^{NMR+D}$ structure is closer to the hPrP^{c} X-ray structure (1I4M) at Loop 1 with smaller residue RMSD values compared with <E200K>^{NMR} (in Fig. 2(d)). This indicates that the database-derived distance constraints indeed 'improve' the loop conformations. As to Loop 2, the residue RMSD values in the <E200K>$^{NMR+D}$ structure are consistently smaller than those in <E200K>^{NMR} structure (in Fig. 2(a-d)), showing that the database-derived distance constraints make the loop more consistent with the experimental NMR and X-ray structures.

In addition to the under-determined loop regions, differences between <E200K>$^{NMR+D}$ and <E200K>^{NMR} were observed in well-defined helical regions (Helix 2 and Helix 3) (Figure 3). The dimeric hPrP^{c} X-ray structure was used as a reference structure for this comparison. The residual RMSD values for <E200K>$^{NMR+D}$ are slightly smaller than those for <E200K>^{NMR} in the N terminal of Helix 2 (residues 172-190) and Helix 3 (residues 201-228), showing that the helical regions of <E200K>$^{NMR+D}$ are closer to the corresponding X-ray structure compared with those of <E200K>^{NMR} (Fig. 3(a), 3(c)). As to the region between the β-sheet and Helix 2 (residues 191-199), the <E200K>$^{NMR+D}$ structure is closer to the X-ray structure compared with <E200K>^{NMR} (Fig. 3(b)). Overall, the <E200K>$^{NMR+D}$ structure has been 'improved' in both the under-determined loop regions and the well-defined helix regions.

Table 2. Average RMSD from experimental restraints

	†<E200K>^{NMR}	‡<E200K>^{NMR}	†<E200K>$^{NMR+D}$	‡<E200K>$^{NMR+D}$
Distances (Å)	0.0046 ± 0.0018	0.0030	0.0047 ± 0.0016	0.0040
Angles (degrees)	0.1664 ± 0.0368	0.1540	0.1589 ± 0.0340	0.1380
J-couplings (Hz)	0.3787 ± 0.0951	0.2470	0.2105 ± 0.0186	0.2550

†Average RMSD ± standard deviations for the ensemble of structures, ‡RMSD for the lowest energy structure in the ensemble

Table 3. Average RMSD from idealized geometries

	†<E200K>^{NMR}	‡<E200K>^{NMR}	†<E200K>$^{NMR+D}$	‡<E200K>$^{NMR+D}$
Bond lengths (Å)	0.0014 ± 0.0002	0.0012	0.0014 ± 0.0002	0.0012
Bond Angles (°)	0.3128 ± 0.0212	0.2990	0.3108 ± 0.0153	0.3040
Improper-Angles (°)	0.2148 ± 0.0236	0.2000	0.2105 ± 0.0186	0.2120

†Average RMSD ± standard deviations for the ensemble of structures, ‡RMSD for the lowest energy structure in the ensemble

Table 4. Potential energy of different types

	†<E200K>^{NMR}	‡<E200K>^{NMR}	†<E200K>$^{NMR+D}$	‡<E200K>$^{NMR+D}$
Total Energy	104.16 ± 24.80	82.08	102.31 ± 23.09	86.30
Bonds	3.30 ± 1.11	2.45	3.54 ± 1.58	2.70
Bond-Angles	46.78 ± 6.67	42.53	46.11 ± 4.92	44.11
Improper-Angles	6.78 ± 1.54	5.80	6.49 ± 1.21	6.53
Van der Waals	34.44 ± 9.44	26.23	31.81 ± 6.97	26.29
NOE	5.85 ± 4.80	2.11	6.97 ± 6.57	3.60
Dihedral-Angles	0.31 ± 0.14	0.26	0.28 ± 0.14	0.21

†Average energy ± standard deviations for the ensemble of structures, ‡energy for the lowest energy structure in the ensemble (kcal/mol)

Table 5. Percentage of residues in different Ramachandran plot regions

	†<E200K>^{NMR}	‡<E200K>^{NMR}	†<E200K>$^{NMR+D}$	‡<E200K>$^{NMR+D}$
Most favorable	85.40%	84.40%	89.60%	88.50%
Additional allowed	14.60%	14.60%	10.40%	11.50%
Generously allowed	0.00%	0.00%	0.00%	0.00%
Disallowed	0.00%	1.00%	0.00%	0.00%

†Average percentage of residues in different (φ,ψ)-regions for the average and energy-minimized structure in the ensemble, ‡the percentage of residues in different (φ,ψ)-regions for the lowest energy structure in the ensemble

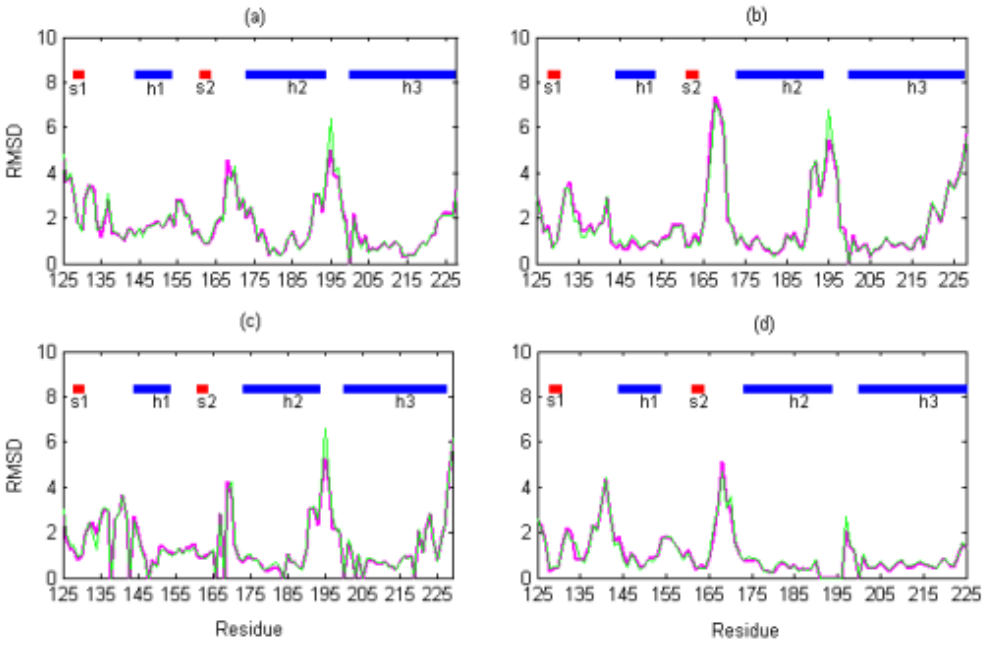


Fig 2. Residue-residue comparison between the refined E200K and wild-type NMR and X-ray structures. The graphs show the residue RMSD values (with the backbone atoms, N, C^{α}, C', and O) for the average and energy-minimized structures of <E200K>$^{NMR+D}$ (magenta line) and <E200K>^{NMR} (green line) compared with the structures of (a) the wild-type hPrP^{c} NMR structure (1QM0) at mildly acidic condition (pH 4.5); (b) the wild-type hPrP^{c} NMR structure (1HJM) at neutral condition (pH 7.0); (c) the wild-type shPrP^{c} X-ray structure (1UW3) in a monomeric form; (d) the wild-type hPrP^{c} X-ray structure (1I4M) in a dimeric form. The secondary structures are indicated along the top of each part of the figure with *h* representing alpha helix and *s* beta sheet.

3.3. Conformational analysis of the refined E200K loop regions

To further elucidate the improvements of the under-determined loop regions of E200K, the sequential φ and ψ angles for each residue between <E200K>^{NMR} and <E200K>$^{NMR+D}$ were compared (Fig. 4) and displayed in Ramachandran plots (Fig. 5). As shown in Fig. 4, the angles φ and ψ of <E200K>$^{NMR+D}$ (magenta square) and <E200K>^{NMR} (green square) were extremely close to one another, except in Loop 1 and 2. In the Loop 1 of <E200K>^{NMR}, the residues lie far outside the most favourable regions of the Ramachandran plot (Fig. 5(a)). However, most of these residues have moved into the most favourable regions of Ramachandran plot (Fig. 5(b)) in the <E200K>$^{NMR+D}$ structure. We obtained a similar result for the Loop 2 (data not shown).

3.4. Backbone and side-chain packing analysis of the refined E200K loop regions

To illustrate the difference of backbones and side-chains in the loop regions between the <E200K>$^{NMR+D}$ and <E200K>^{NMR} structures, we superimposed the backbones of the whole structure (Fig. 6(a)) and of the loop structures (Fig. 6(a)-6(e). Apparently, the Loop 1 of <E200K>$^{NMR+D}$ differ significantly from the same region of <E200K>^{NMR} (Fig. 6(b)), implying that the database-derived distance constraints affect the backbone conformations in regions where experimental restraints are insufficient (Table 1). The conformations of Loop 2 in both structures appear quite similar except at residue GLY195 (Fig. 6(c)). It has been shown that the conformations of glycine and neighbouring residues can be improved by using database-derived distance constraints [28-29]. Examination of side-chain packing in the two loop regions showed that overall conformations of side-chains are quite similar between <E200K>$^{NMR+D}$ and <E200K>^{NMR} (Fig. 6(d-e)). No change in either hydrogen bonds or salt bridges was observed. It suggests that the database-derived distance constraints

do not affect the side-chain packing, which is not so surprising since only the constraints between backbone heavy atoms were utilized.

4. Discussion

Prion diseases can be sporadic (spontaneous), inherited, or transmitted by infectious agents. Some prion diseases in humans, such as the familiar CJD, FFI, and GSS, are inherited and linked to mutations in the PrP^c-coding gene, PRNP. More than 20 mutations in this gene have been associated with prion diseases . The mutation E200K, where a glutamic acid is substituted with a lysine at residue 200, is a major cause of familiar CJD . The tertiary structure of the variant E200K of human PrP^c is almost identical to the wild-type prion . This mutation changes not only the surface potential and charge distribution of the protein, which might affect its interaction with Protein X or other cellular components and the conversion from PrP^c to PrP^{Sc} , but also the stability of prion protein . On the other hand, this mutation by itself cannot lead to the conversion, which could require additional modification of the protein and additional cofactors . The possible involvement of Protein X in the pathogenic process of the familiar CJD makes the structure of the binding sites between the mutated prion (E200K variant) and Protein X, which include the loop regions encompassing residues 167-171, urgently needed. However, the lack of experimental constraints in this critical region makes the structures poorly defined [14-16]. Here, we have employed distance constraints that are derived from a database of high-resolution protein structures and significantly improved the conformation of the loop region. With the database-derived distance constraints, both loop regions (residues 167-171 and 195-199) in $<E200K>^{NMR+D}$ showed more generally acceptable conformations. The comparisons between the generated structures and the experimentally determined structures of the prion protein confirmed a convergence between the structures, which implies that an increase in the accuracy of the poorly determined loop structures is possible with database-derived distance constraints. The improved structures may provide a better structural understanding of prion proteins and hence facilitate insights into its conversion in the spongiform encephalopathies.

5. Acknowledgement

The work is partially supported by the research funds provided by the Department of Mathematics, the Graduate Program on Bioinformatics and Computational Biology, and the Lawrence Baker Center for Bioinformatics and Biological Statistics at Iowa State University and the NIH/NIGMS grant R01GM081680. The work is also partially supported by the Institute of Mathematics and Its Applications with funds provided by the NSF.

6. References

[1] P. Brown and D.C. Gajdusek, The human spongiform encephalopathies: kuru, Creutzfeldt-Jakob disease, and the Gerstmann-Straussler-Scheinker syndrome, *Curr. Top. Microbiol. Immunol* **171** (1991), pp. 1-20.

[2] A.L. Horwich and J.S. Weissman, Deadly conformations – protein misfolding in prion disease, *Cell* **89** (1997), pp. 499-510.

[3] S.B. Prusiner, Prions, *Proc. Natl. Acad. Sci. USA* **95** (1998), pp. 13363-13383.

[4] S.B. Prusiner, Novel proteinaceous infectious particles cause scrapie, *Science* **216** (1982), pp. 136-144.

[5] N. Stahl, M.A. Baldwin, D.B. Teplow, L. Hood, B.W. Gibson, A.L. Burlingame and S.B. Prusiner, Structural studies of the scrapie prion protein using mass spectrometry and amino acid sequencing. *Biochemistry* **32** (1993), pp. 1991-2002.

[6] R.K. Meyer, M.P. McKinley, K.A. Bowman, M.B. Braunfeld, R.A, Barry and S.B. Prusiner, Separation and properties of cellular and scrapie prion proteins, *Proc. Natl. Acad. Sci. USA* **83** (1986), pp. 2310-2314.

[7] B. Oesch, D. Westaway, M. Walchli, M.P. Mckinley, S.B. Kent, R. Aebersold, R.A. Barry, P. Tempst, D.B. Teplow, L.E. Hood, *et al.*, A cellular gene encodes scrapie PrP 27-30 protein, *Cell* **40** (1985), pp. 735-746.

[8] K.M. Pan, M. Baldwin, J. Nguyen, M. Gasset, A. Serban, D. Groth, I. Mehlhorn, Z. Huang, R.J. Fletterick, F.E. Cohen, *et al.*: Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins, *Proc. Natl. Acad. Sci. USA* **90** (1993), pp. 10962-10966.

[9] B.W. Caughey, A. Dong, K.S. Bhat, D. Ernst, S.F. Hayes and W.S. Caughey, Secondary structure analysis of the scrapie-associated protein PrP 27-30 in water by infrared spectroscopy, *Biochemistry* **30** (1991), pp. 7672-7680.

[10] J. Safar, P.P. Roller, D.C. Gajdusek and C.J. Gibbs Jr., Thermal stability and conformational transitions of scrapie amyloid (prion) protein correlate with infectivity, *Protein Sci.* **2** (1993), pp. 2206-2216.

[11] R. Riek, S. Hornemann, G. Wider, M. Billeter, R. Glockshuber and K. Wüthrich, NMR structure of the mouse prion protein domain PrP (121-321), *Nature* **382** (1996), pp. 180-182.

[12] P.M. Rudd, M.R. Wormald, D.R. Wing, S.B. Prusiner and R.A. Dwek, Prion glycoprotein: structure, dynamics, and roles for the sugars, *Biochemistry* **40** (2001), pp. 3759-3766.

[13] A.C. Apetri, K. Surewicz and W.K. Surewicz, The effect of disease-associated mutations on the folding pathway of human prion protein, *J. Biol. Chem.* **279(17)** (2004), pp. 18008-18014.

[14] Y. Zhang, W. Swietnicki, M.G. Zagorski, W.K. Surewicz and F.D. Sönnichsen, Solution structure of the E200K variant of human prion protein, *J. Biol. Chem.* **275** (2000), pp. 33650-33654.

[15] R. Zahn, A. Liu, T. Lührs, R. Riek, C. von Schroetter, F.L. Garcia, M. Billeter, L. Calzolai, G. Wider and K. Wüthrich, NMR solution structure of the human prion protein, *Proc. Natl. Acad. Sci. USA* **97** (2000), pp. 145-150.

[16] L. Calzolai and R. Zahn, Influence of pH on NMR structure and stability of the human prion protein globular domain, *J. Biol. Chem.* **278** (2003), pp. 35592-35596.

[17] K. Knaus, M. Morillas, W. Swietnicki, M. Malone, W.K. Surewicz and V.C. Yee, Crystal structure of the human prion protein reveals a mechanism for oligomerization, *Nat. Struct. Biol.* **8 (**2001), pp. 770-774.

[18] L.F. Haire, S.M. Whyte, N. Vasisht, A.C. Gill, C. Verma, E.J. Dodson, G.G. Dodson and P.M. Bayley, The crystal structure of the globular domain of sheep prion protein, *J. Mol. Biol.* **336** (2004), pp. 1175-1183.

[19] G.C. Telling, M. Scott, K.K. Hsiao, D. Foster, S.L. Yang, M. Torchia, K.C. Sidle, J. Collinge, S.J. DeArmond and S.B. Prusiner, Transmission of Creutzfeldt-Jakob disease from humans to transgenic mice expressing chimeric human-mouse prion protein, *Proc. Natl. Acad. Sci. USA* **91** (1994), pp. 9936-9940.

[20] G.C. Telling, M. Scott, J. Mastrianni, R. Gabizon, M. Torchia, F.E. Cohen, S.J. DeArmond and S.B. Prusiner, Prion propagation in mice expressing human and chimeric PrP transgenes implicates that interaction of cellular PrP with another protein, *Cell* **83** (1995), pp. 79-90.

[21] M. Billeter, R. Riek R, G. Wider, S. Hornemann, R. Glockshuber and K. Wüthrick, Prion protein NMR structure and species barrier for prion diseases, *Proc. Natl. Acad. Sci. USA* **94** (1997), pp. 7281-7285.

[22] K. Kaneko, L. Zulianello, M. Scott, C.M. Cooper, A.C. Wallace, T.L. James, F.E. Cohen and S.B. Prusiner, Evidence for protein X binding to a discontinuous epitope on the cellular prion protein during scrapie prion propagation, *Proc. Natl. Acad. Sci. USA* **94** (1997), pp. 10069-10074.

[23] A.D. Gossert, S. Bonjour, D.A. Lysek, F. Fiorito and K. Wuthrich, Prion protein NMR structures of elk and of mouse/elk hybrids, *Proc Natl Acad Sci USA* **102** (2005), pp. 646-650.

[24] R.A. Laskowski, M.W. MacArthur, D.S. Moss and J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Cryst.* **26** (1993), pp. 283-291.

[25] A.L. Morris, M.W. MacArthur, E.G. Hutchinson and J.M. Thornton, Stereochemical quality of protein structure coordinates, *Proteins* **12** (1992), pp. 345-364.

[26] G.N. Ramachandran and V. Sasisekharan, Conformation of polypeptides and proteins, *Adv. Protein Chem.* **23** (1968), pp. 283-437.

[27] G.M. Clore, Improving the quality of NMR and crystallographic protein structures by means of a conformational database potential derived from structure databases, *Protein Sci.* **5** (1996), pp. 1067-1080.

[28] F. Cui, R.L. Jernigan and Z. Wu, Refinement of NMR-determined protein structures with database-derived distance constraints, *J. of Bioinformatics and Computational Biology* **3** (2005), pp. 1315-1331.

[29] F. Cui, R.L. Jernigan and Z. Wu, Knowledge-Based versus Experimentally-Acquired Distance and Angle Constraints for NMR Structure Refinement, *J. of Bioinformatics and Computational Biology* 2007, submitted.

[30] S. Miyazawa and R.L. Jernigan, Estimation of effective inter-residue contact energies from protein crystal structures: quasi-chemical approximation, *Macromolecules* **18** (1985), pp. 534-552.

[31] S. Miyazawa and R.L. Jernigan, Residue-residue potentials with a favourable contact pair term and an unfavourable high packing density term for simulation and threading, *J. Mol. Biol.* **256** (1996), pp. 623-644.

[32] M.J. Sippl, Calculation of conformational ensembles from potentials of mean force, *J. Mol. Biol.* **213** (1990), pp. 859-883.

[33] M.J. Sippl and S. Weitckus, Detection of native-like models for amino acid sequence of unknown three-

dimensional structure in a database of known protein conformations, *Proteins* **13** (1992), pp. 258-271.

[34] A. Rojnuckarin and S. Subramaniam, Knowledge-based potentials for protein structure, *Proteins* **36** (1999), pp. 54-67.

[35] M.E. Wall, S. Subramaniam and G.N. Phillips Jr., Protein structure determination using a database of inter-atomic distance probabilities, *Protein Sci.* **8** (1999), pp. 2720-2727.

[36] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, L.N. Shindyalov and P.E. Bourne, The Protein Data Bank, *Nuc. Acids Res.* **28** (2000), pp. 235-242.

[37] J.F. Doreleijers, S. Mading, D. Maziuk, K. Sojourner, L. Yin, J. Zhu, J.L. Makley and E.L. Ulrich, BioMagResBank database with sets of experimental NMR constraints corresponding to the structures of over 1400 biomolecules deposited in the Protein Data Bank, *J. Biomol. NMR* **26** (2003), pp. 139-146.

[38] A.T. Brünger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, S. Jiang, J. Kuszewski, N. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson and G.L. Warren, Crystallography and NMR System: A new software suite for macromolecular structure determination, *Acta Cryst.* **D54 (**1998), pp. 905-921.

[39] L.G. Goldfarb, P. Brown, L. Cervenakova and D.C. Gajdusek, Genetic analysis of Creutzfeldt-Jakob disease and related disorders, *Phil. Trans. Roy. Soc. Lond. B Biol. Sci.* **343** (1994), pp. 379-384.

[40] K. Hsiao, Z. Meiner, E. Kahana, C. Cass, I. Kahana, D. Avrahami, G. Scarlato, O. Abramsky, S.B. Prusiner and R. Gabizon, Mutation of the prion protein in Libyan Jews with Creutzfeldt-Jakob disease, *N. Engl. J. Med.* **324** (1991), pp. 1091-1097.

[41] Y. Levy and O.M. Becker, Conformational polymorphism of wild-type and mutant prion proteins: Energy landscape analysis, *Proteins* **47**(4) (2002), pp. 458-468.

[42] S. Capellari, P. Parchi, C.M. Russo, J. Sanford, M.S. Sy, P. Gambetti and R.B. Petersen, Effect of the E200K mutation on prion protein metabolism, *Amer. J. Pathol.* **157** (2000), pp. 613-622.

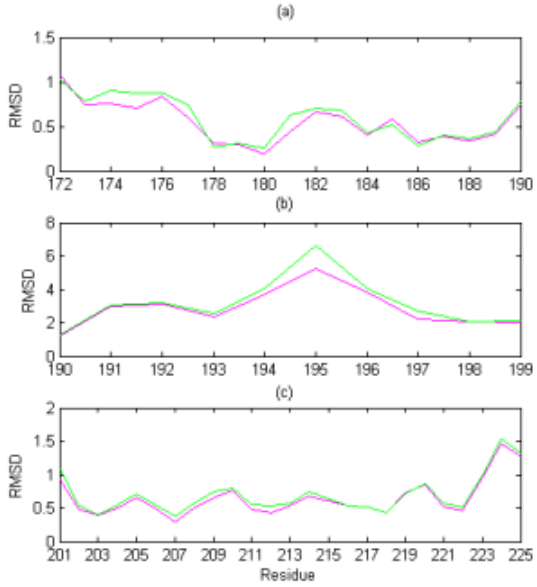


Fig 3. Detailed residue RMSD plots for helix and loop regions. The graphs show the detailed residue RMSD values (of the backbone atoms, N, C^{α}, C', and O) for the average and energy-minimized structures of $\langle E200K \rangle^{NMR+D}$ (magenta line) and $\langle E200K \rangle^{NMR}$ (green line) at (a) N-terminal of Helix 2 (residues 172-190) when compared with the $hPrP^{c}$ X-ray structure; (b) C-terminal of Helix 2 and the loop between Helix 2 and Helix 3 (residues 191-199) compared with the $hPrP^{c}$ X-ray structure; and (c) Helix 3 (residues 201228) when compared with the $hPrP^{c}$ X-ray structure.

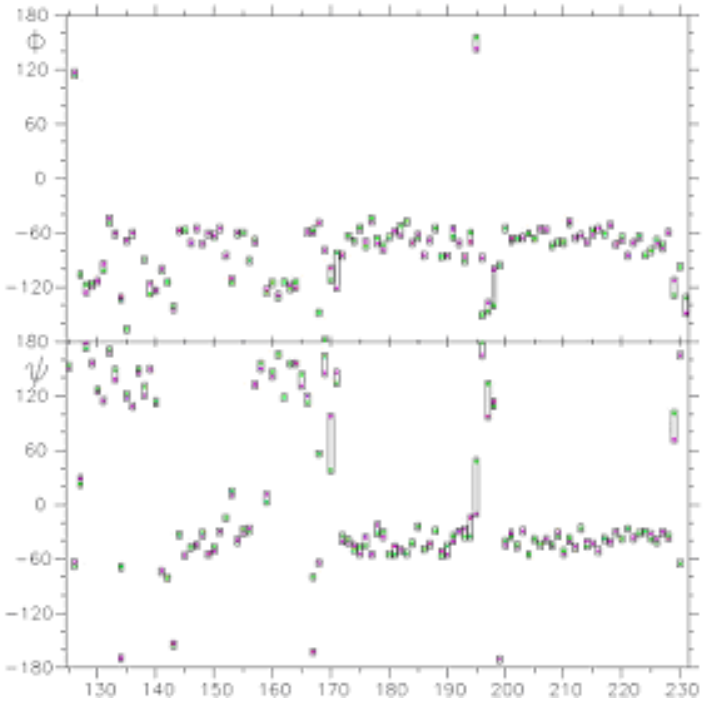


Fig 4. The φ and ψ angles of the refined E200K structures. The angles of <E200K>$^{NMR+D}$ are represented by magenta squares and for <E200K>^{NMR} by green squares.

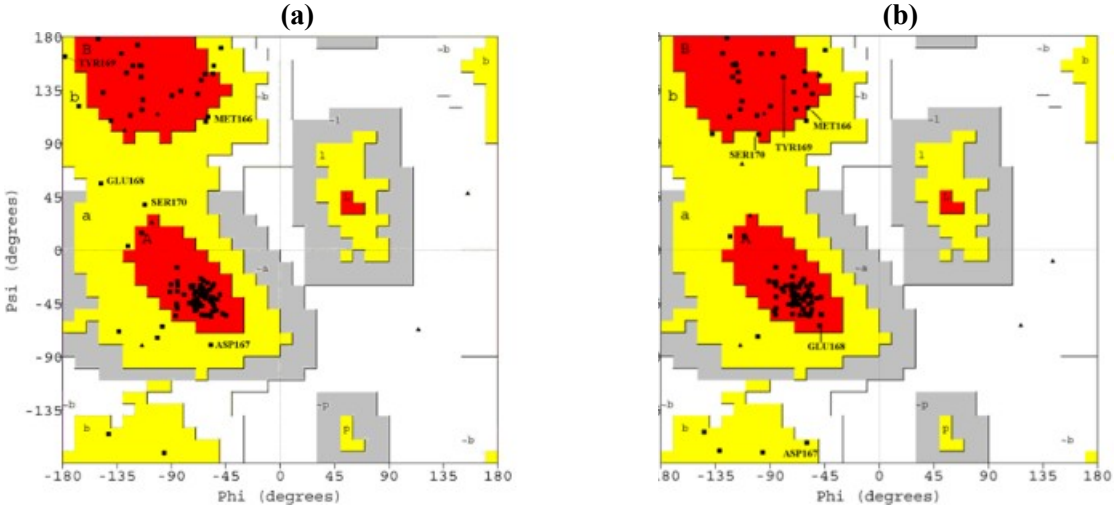


Fig 5. Ramachandran plots of <E200K>^{NMR} and <E200K>$^{NMR+D}$. (a) Ramachandran plot showing the values of (φ, ψ) angles of the average and energy-minimized structure of <E200K>^{NMR}; (b) Ramachandran plot of <E200K>$^{NMR+D}$. Only the residues of Loop 1 are shown. In (a), most of the residues in Loop 1 are found outside the most favourable (red) regions, while in (b), most of these residues lie within the most favorable regions.

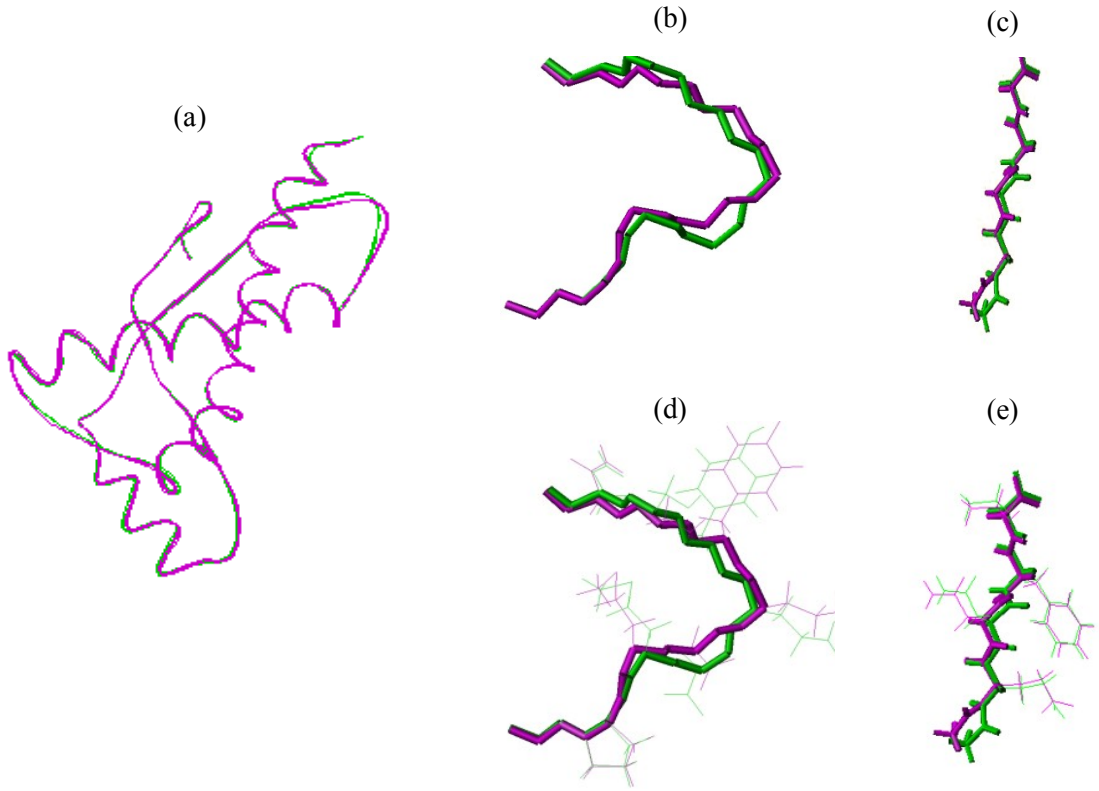


Fig 6. Superimposition of tertiary structures. Tertiary structures of the average and energy-minimized structures of <E200K>$^{NMR+D}$ (small magenta cylinders) and <E200K>^{NMR} (small green cylinders) are superimposed in (a) the backbone of the whole protein; (b) the backbone of Loop 1; (c) the backbone of Loop 2; (d) the backbone and side-chain of Loop 1; (e) the backbone and side-chain of Loop 2.