Original Paper

Intervirology

Intervirology 2008;51:210–216
DOI: 10.1159/000151732

Received: January 7, 2008
Accepted after revision: July 14, 2008
Published online: August 28, 2008

New Insight into the Species Barrier of Mammalian Prions Obtained from Structure-Based Conservation Analysis

Yaofeng Wang Bingjie Hu Yuan Liu Huili Zhang Youtao Song

College of Life Science, Liaoning University, Shenyang, PR China

Key Words

Prions, mammalian · Structural conservation analysis · Phylogenetic trees · Species barrier

Abstract

Objectives: The structures of several mammalian prions have been determined by NMR; however, their transmission and evolutionary features are still unclear. The objective of this study is to explore species barrier information from their structures. ***Methods:*** In this study, we compared the functional domains (121–231) of 8 mammalian prions with structure-based conservation analysis using visual molecular dynamic software (version 1.8.5). ***Results:*** Residues 184, 203 and 205, which play important roles in modulating the transmission of prion diseases, were identified as a structure-conserved domain rather than sequence-conserved domain of mammalian prions. The phylogenetic tree was reconstructed based on structure conservation. The topology of the tree explains the species barrier during the transmission of mammalian prions which cannot be deduced by the sequence phylogenetic tree. ***Conclusions:*** These results suggest that structure-based analysis is more accurate and informative than sequence-based analysis in the study of prion transmission and evolution.

Y.W. and B.H. contributed equally to this work.

Introduction

Prion proteins (PrPs) are proteinaceous infectious particles that are probably the primary pathogens of a class of diseases known as transmissible spongiform encephalopathy (TSE) [1, 2]. Because of their unique features as pathogenic proteins, PrPs are the most intensively studied proteins [2–6]. However, little is known precisely about their normal cellular function and pathogenesis [7]. A comparison of mammalian PrPs might help to understand the enigmatic functional and pathogenic properties of these proteins. Previous studies based on sequence comparison between PrPs of different species of mammals suggested their evolutionary relationship [7]. However, the fact that the overall topology of the PrP gene tree agrees with the species tree suggests that no dramatic sequence changes have occurred to avoid cross-species TSE infectivity such as that occurring among humans, bovine and sheep [7]. It is believed that interspecies susceptibility is not simply determined by overall sequence similarity [8]. The PrP strain is more important with regard to the tertiary structure of nascent PrP^{Sc} which is determined by the conformation of PrP^{Sc} together with the PrP sequence of the recipient [9].

It has long been recognized that protein structure is more conserved than sequence [10]. Many potential evolutionary relationships that were not apparent on the basis of sequence analysis have been discovered since three-

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0300–5526/08/0513–0210$24.50/0

Accessible online at:
www.karger.com/int

Youtao Song
College of Life Science, Liaoning University
No. 66 Chongshanzhong Road, Huanggu District
Shenyang 110036 (PR China)
Tel. +86 24 6220 2280, Fax +86 24 8686 4476, E-Mail ysong@lnu.edu.cn

dimensional structures became available [10]. Domain 121–231 of PrPs, which contains most of the point-mutation sites that have been linked in the human PrP to the occurrence of familial PrP diseases, is made up of five secondary structure elements [11]. It is known that all mutations with pathological significance occur either within or adjacent to the regions of secondary structure, notably associated with the second and third α-helix and in most cases appear to destabilize the PrP structure [2, 12, 13]. In the work done by Zahn et al. [14] in 2000, the tertiary structure of this region of the human PrP was determined. However, due to the limited structure data of mammalian PrPs, only the structure of mouse and hamster PrPs were cited then as a comparison to illustrate the evolution and transmission of mammal PrPs. Recent studies done by Lysek et al. [15] and Gossert et al. [16] have made PrP structures of several other mammalian PrPs available, and subsequently made it possible to investigate their structural similarities.

Having noticed the crucial role that domain 121–231 plays in the occurrence of PrP diseases and the theory that more information is conserved in the tertiary structure rather than the primary sequence, we carried out a detailed investigation based on structure alignment of this region among 8 mammalian PrPs. Multiple structural alignments were accurately applied to investigate structure conservation of this domain. As a comparison, multiple sequence alignments were also conducted. Subsequently a robust structural measurement was used to reconstruct the evolutionary history of the PrPs. The consequent phylogenetic tree based on structure conservation analysis helped us gain new insight into the species barrier occurring between the 8 investigated mammalian PrPs.

Material and Methods

PrP Domains and Coordinates

The 8 NMR structures used in this study were from the National Center for Biotechnology Information (Bethesda, Md., USA): mouse PrP 121–231 (PDB ID code 1XYX) [16]; bovine PrP fragment 121–230 (PDB ID code 1DWZ) [17]; human PrP 121–230 (PDB ID code 1H0L) [14]; cat PrP 121–231 (PDB ID code 1XYJ); hamster PrP 121–231 (PDB ID code 1B10); pig PrP 121–231 (PDB ID code 1XYQ); elk PrP 121–231 (PDB ID code 1XYW) and sheep PrP 121–231 (PDB ID code 1XYU) [15].

Structure and Sequence Conservation Analysis

MultiSeq [18] in visual molecular dynamic (VMD) software [19] was used to analyze both the structures and the sequence data. For structural alignment, the multiple structural alignment program STAMP [20] was used with the parameters npass = 2, scanscore = 6, and scanslide = 5. The multiple alignments generated were then measured with Qres [21] to calculate the structural similarity of each residue. The RMSD value was also calculated by the script in VMD.

For sequence data, CLUSTAL W [22] is used, and the sequence identity of each residue was subsequently calculated for this multiple sequence alignment by Multiseq.

Homology Measure and Phylogenetic Analysis

Method Q_H [21] is used for homology measurements [20]. In order to investigate the structural evolutionary relationships between PrPs, we applied the unweighted pair group method with arithmetic averages (UPGMA) [23] for cluster analysis.

All figures were made in VMD version 1.8.5 (Theoretical and Computational Biophysics Group, UIUC, NIH Resource for Macromolecular Modeling and Bioinformatics, Bethesda, Md., USA).

Results and Discussion

Structure-Based Conservation Analysis Is More Informative than Sequence-Based Conservation Analysis of Mammalian PrPs

A previous study using sequence-based conservation analysis has shown a lot of obvious features of domain 121–231 of mammalian PrPs [7]. However, due to the high sequence similarity among species, simple sequence analysis cannot give us enough information. Rather, structure-based analysis can be more informative and structure alignment is much more likely to be accurate in terms of biological function and evolution [20].

In this study, these two approaches were both applied to investigate the relationships among the human, bovine, sheep, mouse, hamster, elk, pig and cat mammalian PrPs. In figure 1A, the 8 superimposed PrPs are color coded according to their conservation level which is measured by the Q method. The Q value per residue for every PrP is quantitatively illustrated in figure 1B, and ranges from 0 to 1 with Q = 0 referring to totally different structures and Q = 1 referring to identical proteins. As a comparison, figure 2A is color coded according to sequence identity per residue, which is based on multiple sequence alignments, and the relative sequence identity for every PrP is quantitatively illustrated in figure 2B.

Comparing figure 1A with 2A, the similarity between them is apparent. Their structural cores are well conserved with increasing variation towards the protein surface. The secondary structures are conserved much more than non-secondary structures, such as the Ala-Gly-rich region 143–163 which is the hydrophobic transmembrane segment, Cys179 and Cys214 which involve in the disul-

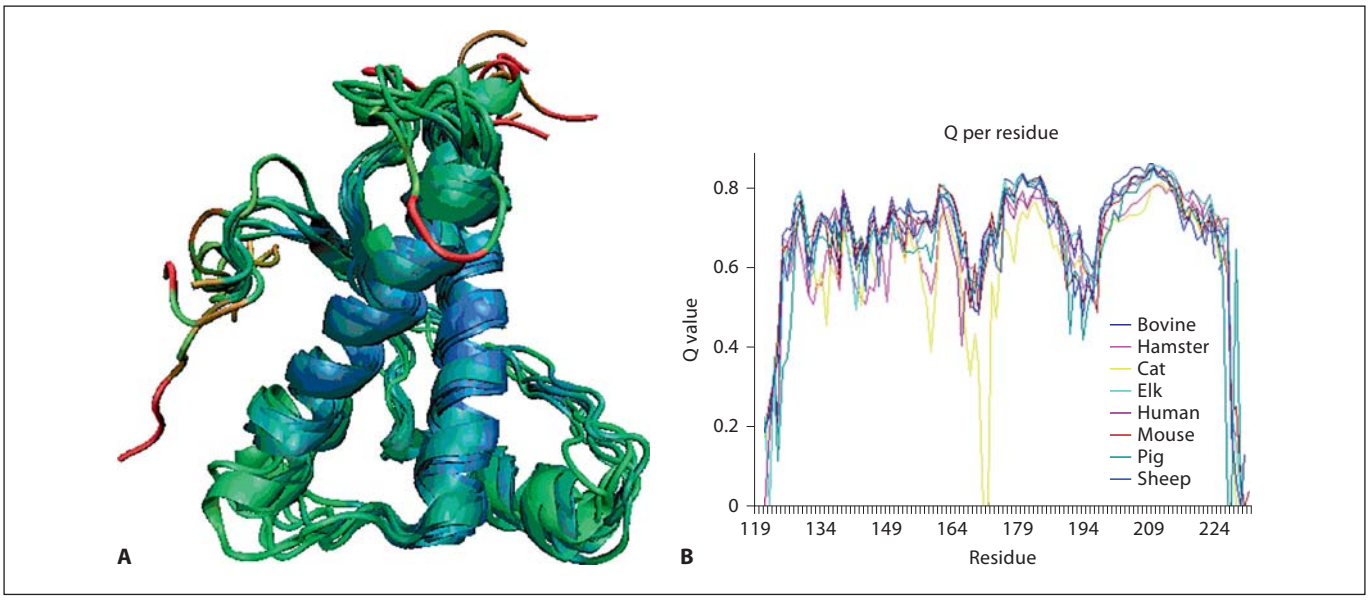


Fig. 1. Structural alignment of the mammalian prions. **A** Superimposed prions color coded by Q per residue. The brighter (bluer in the online version) the color the higher the conservation level; the dark black (red in the online version) color is the relatively lowest conservation level. A striking feature is that the color is more abundant than in the plot shown in figure 2A. **B** 2D plot representing Q per residue based on structure alignment. The residue numbers are given in the x-axis, and the Q values are given in the y-axis (Q = 0 refers to totally unconserved residues, while Q = 1.0 refers to structurally identical residues). Residues 184, 203 and 205 are located in the regions with a high Q value which illustrates their high structure conservation (their Q values are 0.80, 0.78 and 0.80, respectively).

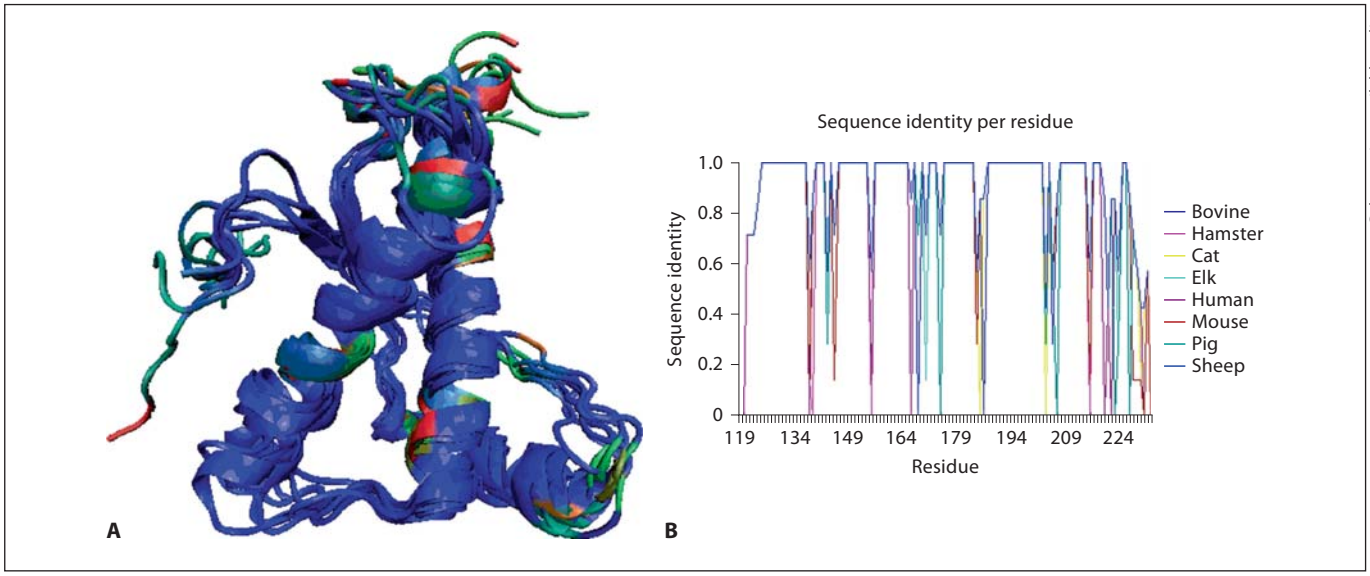
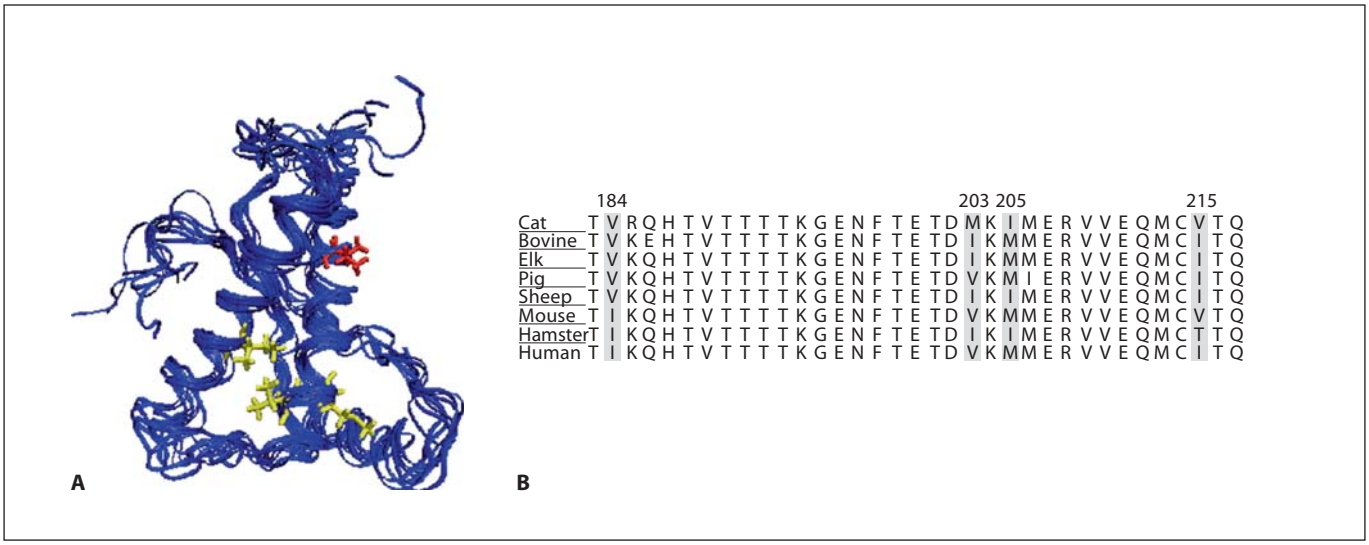


Fig. 2. Multiple sequence alignment of the mammalian prions. **A** Superimposed prions color coded by sequence identity per residue. The brighter (bluer in the online version) the color the higher the conservation level; the dark black (red in the online version) color is the relatively lowest conservation level. The color hierarchy in this map is more uniform than that in figure 1A. **B** 2D plot representing sequence identity per residue based on multiple sequence alignment. The residue numbers are given in the x-axis, and sequence identity values are given in the y-axis (0 refers to the most divergent sequence, while 1.0 refers to sequence identically the same). Residues 184, 203 and 205 are located in the region with a low sequence similarity (the identity values are 0.46, 0.32 and 0.46, respectively).



Color version available online

Fig. 3. Particular residues selected by structure-based conservation analysis show their 3D orientation. **A** Superimposed mammalian prions indicate Q 0.75. The black (yellow in the online version)-coded residues represent residues 184, 203 and 205 which form a putative epitope. The black (red in the online version)-coded residue represents residue 215. These residues are thought to be structurally conserved in a sequence unconserved region. **B** Corresponding sequence alignment shows the sequence identity of the three residues. Residues 184, 203, 205 and 215 are highlighted. It is evident that these residues are polymorphic among the 8 mammalian prions.

fide-bridged helix-loop-helix motif H2–H3, and Asn181 and Asn197 which are the Asn-X-Thr motifs required for N-glycosylation. Besides the secondary structure conservation, an interesting phenomenon is that regions of sequence similarity are 'encapsulated' by larger regions of structural similarity which is in accordance with the guiding principle that sequence conservation determines structure conservation. However, it is not always true that some residues with a high Q value can possess a low sequence identity value, such as the regions around 184, 203 and 205 that will be analyzed in detail below. These differences have prompted us to pay more attention to the characteristics of those structurally conserved residues that may be easily neglected by sequence comparison, and showed us that structure-based conservation analysis is more informative than the sequence-based method.

Due to the difference between basic computational methods, the sequence identity value for some residues can be roughly uniform. The Q method is sensitive enough to differentiate between them. For example, the sequence identity values of the region 125–137 are identical at 1.0, while their Q values gradually increase. This can also be readily and visually seen in figure 1A: the color coding by the Q method is more abundant than that by sequence identity. Thus, Q measurement can give much more meaningful information than the traditional sequence comparison method.

In general, structure-based conservation analysis is more informative than sequence-based conservation analysis of mammalian PrPs. According to this notion, we studied the characteristics implicated in structure-based conservation analysis in more detail.

Structurally Conserved but Sequence Nonconserved Residues Can Be Involved in Modulating Cross-Species Transmission

Motivated by the fact that structure-based analysis is more informative than sequence-based analysis, we studied those residues with high Q values in more detail. We chose Q 0.75, which is the highest structure conservation (more than 75% of residue pairs at the correct distance) [21]. Residues 161–162 (located in the β-sheet), 176–185 (located in α-helix 2), and 203–216 (located in α-helix 3) were selected (fig. 3), and the relative average Q values and sequence identity values of the 8 species are listed in table 1.

Investigating the data in table 1, a conspicuous phenomenon is the contradiction between structure and sequence conservation for residues 184, 203, 205 and 215. These four structurally conserved residues possess a low sequence similarity. Interestingly, in the study by Scott

Table 1. Average conservation value list

Res. No.	161	162	163	176	177	178	179	180	181	182	183	**184**	185	201
QV	0.77	0.78	0.76	0.79	0.77	0.76	0.79	0.81	0.79	0.80	0.80	**0.80**	0.78	0.77
SIV	1	1	1	1	1	1	1	1	1	1	1	**0.46**	0.85	1
Res. No.	202	**203**	204	**205**	206	207	208	209	210	211	212	213	214	**215**
QV	0.76	**0.78**	0.79	**0.80**	0.81	0.81	0.82	0.83	0.84	0.83	0.83	0.83	0.82	**0.80**
SIV	1	**0.37**	1	**0.46**	0.86	1	1	1	1	1	1	1	1	**0.45**

The residues whose average Q values (QV) on the 8 mammals are more than 0.75 were selected, and their relative sequence identity values (SIV) are also illustrated. The contradictions between structure conservation and sequence conservation of residues 184, 203, 205 and 215 are indicated by bold print. Res. No. = Residue number.

et al. [24], residues 184, 186, 203 and 205 were expected to form an epitope modulating PrP transmission, but the mutation of these residues created a species barrier. It is also expected that residue 184, which is not homologous in bovine and mouse PrP and lies at the C-terminal end of the chimeric region, might account for the differences in the susceptibility of transgenetic (Tg) MHu2M and Tg MBo2M mice to PrP. Alternatively residue 203, which is a Val in Mo and HuPrP and is an Ile in BoPrP, might be responsible for the difference between these species in the susceptibility to PrPs. In Tg MBo2M mice, residue 203 is a Val and thus might prevent the conversion from MBo2M PrP^{C} to PrP^{Sc} [24]. However, no evolutionary features of these residues have been reported yet from a structural point of view. According to our results, the sequentially disparate but spatially proximal epitope formed by 184, 203 and 205 was structurally conserved throughout the evolutionary process of the PrP. Though they are highly flexible in sequence, it is evident that the sequence polymorphism did not dramatically affect their tertiary structure, and this is reflected by their biological significance. Subsequently, their structural conservation may insure the infectious ability of PrPs across species, while their sequence polymorphism may contribute to the species specificity of PrPs. Thus together with the previous studies, our study indicates that these functionally important but sequentially variable residues might play some role in the formation of the species barrier.

As for residue 215, it was found to be involved in the discontinuous epitope consisting of residues 168, 172, 215, and 219 that bind to protein X [25]. The striking feature of residues 184, 203 and 205 suggests to us that residue 215 may also be important in terms of the transmission across the species barrier besides being involved in forming the protein X-binding sites. Of course, more evidence is needed to support our theory.

Phylogenetic Order of the Mammalian PrPs Based on Structure-Conservation Analysis Explains the Species Barrier Occurring between Mammalian PrPs

A previous study based on multiple sequence alignment has outlined a phylogenetic tree, the topology of which agrees with the species tree [7]. This agreement suggests that no dramatic sequence changes have occurred to avoid cross-species TSE infectivity [7]. A phylogenetic tree based on structure homology analysis has been reconstructed here to help us gain new insight into the species barriers between these mammalians.

The dendrogram in figure 4 was drawn based on structure analysis. By comparison with the gene tree drawn by van Rheede et al. [7], the most striking information we obtained was the relationship between bovine, sheep and human PrPs. In the gene tree, the bovine and sheep PrPs are relatively more closely related, but were separated from human PrPs at an early evolutionary age. In other words, the evolutionary distance of bovine and sheep PrPs is equal to that of human PrPs. However, studies on transmissions to mice indicate that the novel human PrP disease, variant CJD, is caused by the same PrP strain causing BSE [26–28]. Scrapie of sheep, which has been endemic in United Kingdom for more than 200 years, has never been reported to be able to transmit to humans. Besides this evidence, the transmission of BSE to mice has made them show remarkably uniform disease characteristics, which was different from the contemporary transmission of sheep and goat scrapie [29]. According to our structure-based phylogenetic tree, human and bovine PrPs formed a monophyletic group sharing a common ancestor, while sheep PrPs diverged from human

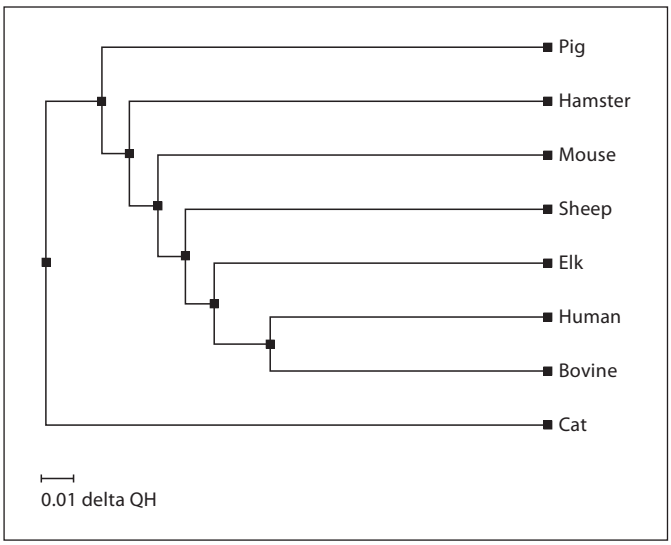


Fig. 4. Polygenetic order based on the QH method of the 8 mammalian prions. The delta QH bar measures the evolutionary distance between the species as detailed in Material and Methods. This dendrogram was drawn by applying UPGMA cluster analysis.

and bovine PrPs at an early evolutionary age. The fact that the BSE agent is able to cross the so-called 'species barrier' and infect humans can be directly reflected by our dendrogram, which indicates that the overall structure of the human PrP in the investigated functional domain is more similar to bovine than sheep PrPs.

The relationship between hamster, mouse and human PrPs can also be explained using our structure-based phylogenetic tree. A previous study has shown that human and mouse PrP proteins were fully compatible, HuPrP fibrils readily seeded mouse protein, and MoPrP fibrils acted as an efficient seed for fibrillization of human PrP [30]. In contrast, a strong species barrier was found between human and hamster PrPs. The lag phases for HuPrP fibrillization in the presence of ShaPrP seeds, or ShaPrP polymerization in the presence of HuPrP seeds were very similar to those for nonseeded reactions, which indicated that no cross-seeding occurs between human and hamster PrPs [30]. According to our tree, mouse PrPs are more homologous to human PrPs than hamster PrPs in their tertiary structure, thus the transmission more likely occurred between human and mouse PrPs rather than between human and hamster PrPs.

Besides the above conclusions, instructive information would be obtained from our tree, such as the relationship between elk and human PrPs. As learned from our tree, the evolutionary distance between elk and human PrPs is closer than that of sheep PrPs but farther than that of bovine PrPs, thus whether the chronic wasting diseases found in cervids is able to transmit to humans or not is of special interest. A recent study has found that there is a substantial species barrier for transmission of elk CWD to humans [31]. However, it has long been thought that feeding cattle with meat and bone meal has allowed the scrapie PrPs from sheep to survive rendering and pass into cattle, which consequently causes the danger in the human food chain [2]. Thus under the instruction of our phylogenetic tree, whether the infectious PrP particles can 'jump' across the species barrier from elk to human via bovine is of outstanding interest and needs to be investigated further.

The TSE species barrier is still a mystery for the time being. However, our phylogenetic order deduced from the tertiary structure has shed more light on the species barrier than the simple sequence phylogenetic tree. Thus under the notion that more information on strain type is encoded conservatively on the tertiary structure rather than on the primary structure [9], more information may be discovered by studying PrP tertiary structure and more insight into the function of the 'mischievous' PrPs will be gained.

Acknowledgement

This work is supported by National Natural Science Foundation of China (No. 30600113) and partially supported by Natural Science Foundation of Liaoning Province (No. 20072055).

References

1 Prusiner SB: Novel proteinaceous infectious particles cause scrapie. Science 1982;216:136–144.
2 Prusiner SB: Prions. Proc Natl Acad Sci USA 1998;95:13363–13383.
3 Brown DR: Prion and prejudice: normal protein and the synapse. Trends Neurosci 2001;24:85–90.
4 Collinge J: Prion diseases of humans and animals: their causes and molecular basis. Annu Rev Neurosci 2001;24:519–550.
5 Hope J: Prions and neurodegenerative diseases. Curr Opin Genet 2000;10:568–574.
6 Rudd PM, Wormald MR, Wing DR, Prusiner SB, Dwek RA: Prion glycoprotein: structure, dynamics, and roles for the sugars. Biochemistry 2001;40:3759–3766.
7 van Rheede T, Smolenaars OM, Madsen O, de Jong WM: Molecular evolution of the mammalian prion protein. Mol Biol Evol 2003;20:111–121.

8 Goldmann W, Hunter N, Somerville R, Hope J: Prion phylogeny revisited. Nature 1996; 382:32–33.
9 Scott M, Foster D, Mirenda C, Serban D, Coufal F, Wälchli M, Torchia M, Groth D, Carlson G, DeArmond SJ: Transgenic mice expressing hamster prion protein produce species-specific scrapie infectivity and amyloid plaques. Cell 1989;59:847–857.
10 Chothia C, Lesk A: The relation between the divergence of sequence and structure in proteins. EMBO J 1986;5:823–826.
11 Prusiner SB: Genetic and infectious prion diseases. Arch Neurol 1993;50:1129–1153.
12 Krakauer DC, Zanotto PM, Pagel M: Prion's progress: patterns and rates of molecular evolution in relation to spongiform disease. J Mol Evol 1998;47:133–145.
13 Liemann S, Glockshuber R: Influence of amino acid substitutions related to inherited human prion diseases on the thermodynamic stability of the cellular prion protein. Biochemistry 1999;38:3258–3267.
14 Zahn R, Liu A, Lührs T, Riek R, von Schroetter C, Lopez Garcia F, Billeter M, Calzolai L, Wider G, Wüthrich K: NMR solution structure of the human prion protein. Proc Natl Acad Sci USA 2000;97:145–150.
15 Lysek DA, Calzolai L, Pe'rez DR: Prion protein NMR structures of chickens, turtles, and frogs. Proc Natl Acad Sci USA 2005;102: 651–655.
16 Gossert AD, Bonjour S, Lysek DA, Fiorito F, Wüthrich K: Prion protein NMR structures of elk and of mouse/elk hybrids. Proc Natl Acad Sci USA 2005;102:646–650.
17 Hosszu LL, Baxter NJ, Jackson GS, Power A, Clarke AR, Waltho JP, Craven CJ, Collinge J: Structural mobility of the human prion protein probed by backbone hydrogen exchange. Nat Struct Biol 1999;6:740–743.
18 Roberts E, Eargle J, Wright D, Luthey-Schulten Z: MultiSeq: unifying sequence and structure data for evolutionary analysis. BMC Bioinformatics 2006;7:382–392.
19 Humphrey W, Dalke A, Schulten K: VMD – visual molecular dynamics. J Mol Graph 1996;14:33–38.
20 O'Donoghue P, Luthey-Schulten Z: On the evolution of structure in aminoacyl-tRNA synthetases. Microbiol Molec Biol Rev 2003; 67:550–573.
21 Eastwood MP, Hardin C, Luthey-Schulten Z, Wolynes PG: Evaluating protein structure-prediction schemes using energy landscape theory. IBM J Res 2001;45:475–497.
22 Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.
23 Sokal RR, Michener CD: A statistical method for evaluating systematic relationships. Univ Kansas Sci Bull 1958;28:1409–1438.
24 Scott MR, Safar J, Telling G, Nguyen O, Groth D, Torchia M, Koehler R, Tremblay P, Walther D, Cohen FE, DeArmond SJ, Prusiner SB: Identification of a prion protein epitope modulating transmission of bovine spongiform encephalopathy prions to transgenic mice. Proc Natl Acad Sci USA 1997;94: 14279–14284.
25 Hill AF, Joiner S, Linehan J, Sidle KCL, Gowland I, Collinge J: Species-barrier-independent prion replication in apparently resistant species. Proc Natl Acad Sci USA 2000;18: 10248–10253.
26 Bruce ME, Will RG, Ironside JW, McConnell I, Drummond D, Suttie A, McCardle L, Chree A, Hope J, Birkett C: Transmissions to mice indicate that 'new variant' CJD is caused by the BSE agent. Nature 1997;389: 498–501.
27 Collinge J, Sidle KCL, Meads J, Ironside J, Hill AF: Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD. Nature 1996;383:685–690.
28 Hill AF, Desbruslais M, Joiner S, Sidle KCL, Gowland I, Collinge J: The same prion strain causes vCJD and BSE. Nature 1997;389:448–450.
29 Bruce M, Chree A, McConnell I, Foster J, Pearson G, Fraser H: Transmission of bovine spongiform encephalopathy and scrapie to mice: strain variation and the species barrier. Philos Trans R Soc Lond B Biol Sci 1994; 343:405–411.
30 Vanik DL, Surewicz KA, Surewicz WK: Molecular basis of barriers for interspecies transmissibility of mammalian prions. Mol Cell 2004;14:147–148.
31 Kong QZ, Huang SH, Zou WQ, Vanegas D, Wang ML, Wu D, Yuan J, Zheng MJ, Bai H, Deng HY, Chen K, Jenny A L, O'Rourke K, Belay ED, Schonberger LB, Petersen RB, Sy Man-Sun, Chen SG, Gambetti P: Chronic wasting disease of elk: transmissibility to humans examined by transgenic mouse models. J Neurosci 2005;25:7944–7949.