Evaluating Prion Models Based on Comprehensive Mutation Data of Mouse PrP

Tsuyoshi Shirai,1,2,* Mihoko Saito,1,2,5 Atsushi Kobayashi,3 Masahiro Asano,3 Masaki Hizume,3 Shino Ikeda,3 Kenta Teruya,3,6 Masanori Morita,4,4,7 and Tetsuyuki Kitamoto3,*

1Department of Computer Bioscience, Nagahama Institute of Bio-Science and Technology, Nagahama, Shiga 526-0829, Japan
2Bioinformatics Research Division, Japan Science and Technology Agency, 5-3, Yonbancho, Chiyoda-ku, Tokyo 102-8666 Japan
3Department of Neurological Science, Tohoku University Graduate School of Medicine Research, 2-1 Seiryo-machi, Aoba-ku, Sendai 980-8575, Japan
4Research and Development Division, Benesis Corporation, Kitahama, Chuo-Ku, Osaka 541-850, Japan
5Present address: Infectious Pathogen Research Section, Central Research Laboratory, Japan Blood Products Organization, Kobe KIMEC Center Building 8F, 1-5-2 Minatojima-Minamimachi, Chuo-ku, Kobe, 650-0047, Japan
6Present address: Laboratory of Functional Analysis in Silico, Human Genome Center, The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan
7Present address: Department of Chemistry, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kita-ku, Kyoto 603-8334, Japan
*Correspondence: t.shirai@nagahama-i-bio.ac.jp (T.S.), kitamoto@med.tohoku.ac.jp (T.K.)

http://dx.doi.org/10.1016/j.str.2013.12.019

SUMMARY

The structural details of the essential entity of prion disease, fibril prion protein (PrPSc), are still elusive despite the large body of evidence supporting the prion hypothesis. Five major working models of PrPSc structure, which are not compatible with each other, have been proposed. However, no systematic evaluation has been performed on those models. We devised a method that combined systematic point mutation with threading on knowledge-based amino acid potentials. A comprehensive mutation experiment was performed on mouse prion protein, and the PrPSc conversion efficiency of each mutant was examined. The models were evaluated based on the mutation data by using the threading method. Although the data turned out to be rather more consistent with the models that assumed a conversion of the N-terminal region of core PrP into a β helix than with others, substantial modifications were also required to further improve the current model based on recent experimental results.

INTRODUCTION

The prion hypothesis assumes conversion of prion protein (PrP) from the innocuous cellular form (PrPC) into a cytotoxic fibril form (PrPSc) during infection by prion diseases such as Creutzfeldt-Jakob disease, scrapie, and bovine spongiform encephalopathy (Prusiner, 1991, 1998). The protein in the latter form is hypothesized to be the sole pathogenic entity that transmits the disease without mediation by other substances such as nucleic acids. Much biochemical, biomolecular, and pathological evidence that supports this hypothesis has been accumulated to date.

The single difference between PrPC and PrPSc is thought to be their molecular conformations. Therefore, the three-dimensional (3D) structure of PrPSc should represent the most essential information to understand, treat, and prevent prion diseases. Despite the importance of the structure, however, little consensus on PrPSc structure has been established (Diaz-Espinoza and Soto, 2012). It is currently known that, first, the β strand content increases during conversion from PrPC to PrPSc (Caughey et al., 1991; Gasset et al., 1993). Second, a periodic structure at 4.8 Å intervals exists in the PrPSc structure, which is interpreted to be a long cross-β sheet perpendicular to the fibril axis of PrPSc (Nguyen et al., 1995; Wille et al., 2009). Third, PrPC to PrPSc conversion increases the resistance of residues V121–S231 of core PrP (unless otherwise noted, amino acids and residue numbers are those of mouse PrP, and the region S102–P240 of PrP is called core PrP in the following sections) to proteinase K digestion, which is the major hallmark of prion infection, by reflecting the structural changes from PrPC to PrPSc (Prusiner, 1998). To date, a series of electron microscopic, X-ray diffraction, and biochemical studies of PrPs have resulted in several 3D structural working models of PrPSc that are not compatible with each other (Figure 1).

The structure of the core region of mouse PrPC (residues V121–S231) was determined with nuclear magnetic resonance spectroscopy in 1996 (Figure 1A; Riek et al., 1996). Govaerts and colleagues proposed the first molecular model of PrPSc based on the electron microscopic (EM) image of a two-dimensional (2D) crystal in 2004 (Govaerts et al., 2004; Wille et al., 2002). They detected compatibility between the N-terminal region of core PrP (residues G89–H176) and the β helix structure by using a threading method. Based on microscopic images of the wild-type PrP, a deletion mutant (PrPΔ141–176), and carbohydrate-depleted proteins, they proposed a model (called BH1 in this study; Figure 1B), in which the N-terminal region of core PrP was converted into a triangular β helix containing four turns (four-run model), and was associated with the intact C-terminal α2–α3 helix bundle. The converted PrP subunits were assembled into a trimer by using the sidewall of the β helix as an interface,
and the PrP trimers were stacked into fibrils by forming an inter-subunit β helix.

Langedijk and colleagues proposed the second model (BH2; Figure 1C) in 2006 (Langedijk et al., 2006). They also assumed conversion of the N-terminal region of core PrP (residues P104–D143) into a β helix, according to sequence comparisons and molecular dynamics (MD) simulations. The β helix consisted of two turns (two-rung model) with eight residues in one edge of the triangle in contrast to the five residues in the BH1 model. The BH2 model was also assumed to form a fibril in the same manner as the BH1 model.

DeMarco and Daggett observed extension of a native β sheet consisting of β1 and β2 strands and recruitment of a nascent β strand to the β sheet in an MD simulation, and proposed a model in which subunits were assembled into a spiral fibril by forming intersubunit β sheets (SP model; Figure 1D; DeMarco and Daggett, 2004). These three models assumed that conformational conversion was limited to the N-terminal region of core PrP, and that the C-terminal α2-α3 helix bundle was largely intact.

The 3D reconstruction of an EM image of a mouse PrP fibril obtained by Tattum and colleagues demonstrated a ladder-shaped fibril structure, which seemed very different from the preceding 2D image (Tattum et al., 2006). Cobb and colleagues observed the electron paramagnetic resonance (EPR) of human PrP fibrils converted in vitro, and proposed an in-register-stacking model (IS model; Figure 1F), like that also proposed for several amyloid proteins, based on the interresidue distances (Cobb et al., 2007; Lu et al., 2007). In this model, the C-terminal region (residues Q159–K219 in mouse PrP) is converted to a hairpin structure, and the fibril is formed as a long two-layered β sheet, which is moderately similar to the 3D EM image.

Kunes and colleagues proposed another model (BH3; Figure 1E) based on the ladder-shaped 3D EM image, in which the N-terminal region (residues G89–D143) and the C-terminal region (residues Y154–G227) of core PrP were independently converted into β helices. The β helices were stacked into rails, and the linker peptide between the β helices formed the steps of the ladder (Kunes et al., 2008).

Although these five major working models of PrPSc are apparently not consistent with each other, they have not been systematically compared. The fibril nature of PrPSc prevented elucidation of its structure with ordinary structure determination methods such as X-ray crystallography. In such cases, a systematic site-directed mutagenesis study such as alanine scanning or β value evaluation can be used to probe structures by correlating a substituted residue or site position to the effect of substitution (Fersht and Daggett, 2002; Sidhu and Kossiakoff, 2007). This method has been applied to investigate protein folding intermediates or temporary complex structures for which a direct structure determination is difficult.
In this study, we devised a method that combines comprehensive point mutation data with threading based on knowledge-based amino acid potentials, which is used to evaluate the compatibility of amino acids to a molecular environment in a defined protein structure (Lüthy et al., 1992). A systematic point mutation experiment was executed on mouse PrP, and the PrPSc conversion efficiencies in vitro (in-cell conversion) were examined by using a scrapie-infected mouse cell line.

The results of the experiment on the working models of PrPSc were further analyzed by using the threading method to judge whether the models were consistent with the comprehensive mutational data. With this method, each site of the molecular models was tested against 20 amino acids for compatibility with the molecular environment of the site. The results demonstrated that the mutation data were rather more consistent with the models that assumed an N-terminal β helix of core PrP than with the others. However, molecular modeling also indicated that a substantial improvement of the current models is required to make them consistent with the recently accumulated experimental data.

RESULTS

Comprehensive Point Mutations of Mouse PrP

The PrPSc conversion efficiencies for comprehensive point mutants of mouse PrP within the region S102–P240 were evaluated. The mutant genes were expressed in scrapie-infected mouse neuroblastoma (ScN2a) cells (Figure 2; Figures S1 and S2 available online). The PrP expressed and converted into PrPSc in the ScN2a cell is infective in mice (Butler et al., 1988; Race et al., 1989). Also, a positive correlation between the conversion efficiency of PrP mutants in ScN2a cells and that of the mutants in transgenic animals has been demonstrated (Kaneko et al., 1997; Perrier et al., 2002). Therefore, the experimental system of this study would be useful to evaluate PrPSc working models on the basis of conversion efficiencies of the PrP mutants.

The conversion efficiencies were calculated as relative amounts of a mutant PK-resistant PrP molecule, which was corrected by the expression levels of the PrP molecule without PK digestion to that of the transfected wild-type (Figure 2). The linear correlation between expression levels of PrP and the amounts of converted PrP, or constancy of conversion rate, was confirmed for the wild-type PrP (Figures 2D and 2E). The transfected PrPs (PrPSc) are readily converted by intrinsic PrPSc nuclei, which sustained in the ScN2a cell. Therefore, the conversion efficiency should represent the tendency of mutant PrPs to be integrated into PrPSc, rather than the tendency to spontaneously convert into PrPSc.

As a result, the residue sites in which point mutations tended to increase conversion efficiency were mainly localized in residues K103–A119 and S131–M137 in the primary structure (Figure 3 and Table S2). The ten most effective sites (with the highest average of fold-increase of PrPSc conversion over all amino acids) were A115, A117, K105, K109, M137, K103, V121, S131, S134, and V111. The ten most efficient mutants were V121I (23.3-fold increase), K105R (17.1), A115S (16.5), H176R (15.6), A115W (15.5), A117V (14.8), Y127I (13.4), A115N (12.7), P104T (12.1), and A115Y (10.6). Thus, the
Figure 3. Conversion Efficiency of PrP Mutants

The fold differences of PrP<sup>Sc</sup> conversion efficiency of mouse PrP mutants from the wild-type are shown in a heat map (red, increased; green, no change; blue, decreased). The mutants that were excluded from analyses (conversion efficiency < 0.2) are indicated with ‘/’. The horizontal axis is residues from G89 to P240.

(legend continued on next page)
N-terminal region of core PrP appeared to play an important role in PrP\textsuperscript{Sc} conversion.

In contrast, mutations in the C-terminal helical regions of PrP\textsuperscript{C} rarely increased efficiency or reduced protein production by much, except for few residues close to the helix C termini, such as T189 or S231. It has been demonstrated that the mutants at S231 are GPI-modified in this experimental system, therefore might be used for analyses (Hizume et al., 2010). It was also shown that S230 and S232 were critical for GPI-anchoring, and some of the mutants (W, Y, R, F, H, M, E, K, Q, I, L, and C for S230 and W, Y, F, M, I, L, and P for S232) were not GPI-modified, and were not localized to membrane. These mutants have been excluded from the analyses due to a low expression (Figure 3).

**Constructing PrP\textsuperscript{Sc} Working Model Coordinates**

Evaluations of the mutation data against various PrP\textsuperscript{Sc} working models were then attempted. The atomic coordinates of the models were required for this purpose. Unfortunately, however, none of the model coordinates turned out to be publicly available. Therefore, model coordinates were prepared from scratch, as detailed in the Supplemental Experimental Procedures, referring to the original reports. Information explained in texts or presented in graphics was generally not sufficient to reproduce atomic coordinates. Consequently, those used in this study resemble models of the proposed ones to a certain extent, and the authors of this paper are responsible for any discrepancy from the original models. Other proposed models that are considered only a small part of PrP or for which enough information on coordinates was not provided were not considered in this study.

The BH1 (Govaerts et al., 2004), BH2 (Langedijk et al., 2006), SP (DeMarco and Daggett, 2004), BH3 (Kunes et al., 2008), and IS (Cobb et al., 2007) models were constructed and refined by combining manual model constructions and energy calculations. The subunits contained between 74 and 143 acidic residues, and they had \textit{N}-glycan moieties (Man \textit{z}1-6[GlcnAc \textit{b}1-4][Man \textit{z}1-3]Man \textit{b}1-4-GlcNAc \textit{b}1-4-[Fuc \textit{z}1-6]GlcNAc-Asn; Stimson et al., 1999) at N180 and N196, if possible. Every model had at least 98% of the residues in the generously arrowed region in the Ramachandran plot and less than 0.22 Å and 2.5° SD from the ideal geometries for bond distance and angle, respectively (Figure S1 and Table S1). The fibril models contained from 8 to 16 subunits where at least one of the subunits was fully surrounded by other subunits so that all interfaces could be examined (Figure S2).

**Evaluating Models by Threading Method**

A simple but probable assumption was made for the model evaluation; that is, a point mutation that increased conversion efficiency would destabilize the PrP\textsuperscript{C} structure and/or stabilize the PrP\textsuperscript{Sc} structure, and the highly influential sites would be in different molecular environments in PrP\textsuperscript{C} and PrP\textsuperscript{Sc}.

To avoid biases from aberrant PrP samples, the data on mutants were discarded when a defect in protein expression, glycosylation, or conversion into PK-resistant form was detected (Figures 2 and S2). The confirmation of glycosylation in mutant PrP should guarantee proper cellular sorting of the proteins, and that of conversion into PK-resistant form was required to exclude folding-deficient mutants of PrP\textsuperscript{C}.

PrPs have been known to interact with various intracellular cofactors (Caughey et al., 2009). For that reason, some of the PrP mutants might have affected the conversion efficiency owing to the interactions with cofactors rather than destabilization or stabilization of PrP\textsuperscript{C} or PrP\textsuperscript{Sc}. However, the mechanisms and interactions of cofactors that directly affect PrP\textsuperscript{C} to PrP\textsuperscript{Sc} conversion have not been sufficiently clarified at this point. Additionally, PrP\textsuperscript{C} is known to convert into PrP\textsuperscript{Sc} through several intermediates, although the structural detail is not clarified (Stöhrl et al., 2008). Because the fitness of mutant sequences to initial (PrP\textsuperscript{C}) and final (PrP\textsuperscript{Sc}) structures have been evaluated, the presented analysis is analogous to an equilibrium evaluation, and the interactions in intermediate structures have not been explicitly taken into account.

As already mentioned, transfected PrPs are readily converted by intrinsic PrP\textsuperscript{Sc} templates in the conversion system of this study. It implies that the energies of intermediate structures are very much lowered. Without this condition, the conversion efficiency might be largely ruled by conversion kinetics depending on the formation and stability of intermediate structures. The consistency of this analysis obviously depends on an important condition, that is, PrP\textsuperscript{C} to PrP\textsuperscript{Sc} conversion pathway is not impaired by a mutation. If a mutant makes PrP\textsuperscript{Sc} state unreachable, then PrP\textsuperscript{Sc} is not observed. For this reason, the data of the mutants with low conversion efficiency (less than 0.2) were totally discarded, although these mutants might contain considerable number of interesting cases, in which PrP\textsuperscript{Sc} structure was largely destabilized without affecting conversion pathway. Additionally, PrP\textsuperscript{C} to PrP\textsuperscript{Sc} conversion has not been observed in N2a cells, the parent strain of ScN2a (Hiraga et al., 2009). Thus, the most influential cofactor in this system is thought to be PrP\textsuperscript{Sc} itself. In this study, therefore, the conversion efficiencies of mutants were interpreted entirely as a function of initial (PrP\textsuperscript{C}) and final (PrP\textsuperscript{Sc}) structures, although interactions in intermediate structures and with cofactors remain to be a major limitation of current evaluation.

The evaluation was made in both coarse and fine scopes. In the former evaluation, the sites that made a larger difference in conversion efficiency and in different molecular environments between PrP\textsuperscript{C} and PrP\textsuperscript{Sc} in each PrP\textsuperscript{Sc} working model were thought to be consistent with the corresponding model. As a result, the excesses of the fraction of consistent sites over those...
Table 1. PrPSc Model Evaluation

<table>
<thead>
<tr>
<th>Model Evaluationsa</th>
<th>Fine (Threading) Evaluation</th>
<th>Protease Restriction Site Consistency</th>
<th>Model Clearanceb NT NG SS CT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Consistent (%)</td>
<td>Inconsistent (%)</td>
<td>Excess (%)</td>
</tr>
<tr>
<td>PrPSc</td>
<td>18.4</td>
<td>44.7</td>
<td>−26.3</td>
</tr>
<tr>
<td>BH1</td>
<td>37.2</td>
<td>21.3</td>
<td>16.0</td>
</tr>
<tr>
<td>BH2</td>
<td>55.3</td>
<td>26.6</td>
<td>28.7</td>
</tr>
<tr>
<td>SP</td>
<td>40.4</td>
<td>14.9</td>
<td>25.5</td>
</tr>
<tr>
<td>BH3</td>
<td>56.4</td>
<td>24.5</td>
<td>31.9</td>
</tr>
<tr>
<td>IS</td>
<td>64.3</td>
<td>26.6</td>
<td>27.7</td>
</tr>
<tr>
<td>BH4</td>
<td>53.2</td>
<td>25.5</td>
<td>27.7</td>
</tr>
</tbody>
</table>

aPercentage of sites for which mutation data are consistent or inconsistent with each model, excess of consistent fraction over inconsistent fraction, and average of effective correlation coefficients over sites in each model (AECC).
bJudgment on whether a model can accommodate an N-terminal peptide (NT), N-glycans (NG), disulfide bond (SS), and C-terminal peptide, and GPI (CT).
cPercentage of protease restriction sites of PrPSc modeled as accessible residues in each model.
dPercentage of protease restriction sites of PrPSc modeled as accessible residues or not in each model.

of inconsistent sites were 16%, 29%, 26%, 32%, and 28% of the total for BH1, BH2, SP, BH3, and IS models, respectively (Table 1 and Figure 4). The value for BH1 was significantly lower than that of the other models.

In the fine evaluation, threading profiles were calculated for each model and used to correlate the difference in amino acid fitness (the difference in the pseudo-free energy of amino acids before and after point mutation) to the changes in PrPSc conversion efficiency (Ota et al., 2001). A site was regarded as consistent if the conversion efficiency increment was negatively correlated to the difference in pseudo-free energy on the PrPSc working model (because it stabilized PrPSc) and/or positively correlated to that on PrPSc model (because it destabilized PrPSc). The excesses of the consistent site fraction over that of the inconsistent site were 43%, 38%, 38%, 39%, and 14% of the total for BH1, BH2, SP, BH3, and IS models (Table 1 and Figure 4). In this case, the value for IS was significantly lower than that for the other models.

The averages of correlation coefficients over all sites in each model were −0.09, −0.13, −0.12, −0.10, and 0.0 for BH1, BH2, SP, BH3, and IS models, respectively. Altogether, the mutation data appeared to be relatively coherent with the BH2 or BH3 models, followed by the SP model (Figure 4).

Interestingly, correlation analysis showed that destabilization of the PrPSc structure did not always increase conversion efficiency. The consistent site fraction was much smaller than that of inconsistent site fractions for PrPSc (Table 1). A close inspection of the result suggested that this was mainly due to the contribution of mutations that destabilized the C-terminal α2-α3 helix bundle (Figure 5). The result implied that the helix bundle structure is required for the PrPSc conversion process.

Another explanation of this result might be provided by the recently found interaction between the C-terminal α2-α3 helix bundle and the octarepeat at N-terminal region of full-length PrP (SPEVACEK et al., 2013). It was suggested that the mutations at C-terminal α2-α3 helix bundle promoted PrPSc conversion by disrupting the interaction with octarepeat region, rather than by destabilizing the helix bundle. At this point of time, however, evaluation of the comprehensive mutation data under the presence of the structural model of interacting octarepeat should remain as an issue in the future, because atomic details of the interaction are not provided.

Constructing Refined PrPSc Working Model

Although the BH2 or BH3 models were judged to be compatible with the mutation data, the triple-helix fibril structure of the BH2 model and ladder structure of the BH3 model did not resemble the 3D and 2D EM images, respectively, of the mouse PrP fibril (GOVAERTS et al., 2004; TATUM et al., 2006). Therefore, modification of the BH2 model was attempted to construct a model more consistent with the recent experimental data (BH4 model; Figure 1G). The BH2 model was selected as a starting point because preservation of the C-terminal α2-α3 helix bundle was suggested by the analyses described above.

The modeling process of the BH4 model is detailed in the Supplemental Experimental Procedures (Figure S2G). Briefly, the N-terminal β helices of core PrP were arranged so that they comprised the rails of a ladder, and the α2-α3 helix bundles were placed on steps of the ladder. The interfaces between subunits were selected according to the experimental results of PrPSc to PrPSc/PrPSc interactions obtained with the hydrogen exchange, antibody inhibition, and peptide-array methods (RIGTER et al., 2007, 2009; SOFORSOI et al., 2007).

Four different interfaces were assumed for a subunit in the BH4 model. The interfaces 1 and 2 were at the top and bottom of the N-terminal β helix of core PrP, and were also assumed in the BH1, BH2, and BH3 models (Figure 6A). Interface 3 provided lateral interaction between β helices from neighboring subunits, which was also hypothesized in the BH1 and BH2 models. By using interfaces 1, 2, and 3, like the BH2 model, the fibril of the BH4 model could be constructed as observed in the 2D EM image (Figure S6A). The additional interface 4 was located...
between α2-α3 helix bundles, which was suggested from the peptide-array experiment for sheep PrP. The ladder-shaped fibril could be formed with interfaces 1, 2, and 4.

The excess of consistent over inconsistent site fractions for the BH4 model was 28% (coarse evaluation), and 42% (fine evaluation). The average of the efficient correlation coefficient was −0.19. All of these values were comparable to the best values from other models (Figure 4). Also in this model, the experimentally detected interfaces between PrP	extsuperscript{C} and PrP	extsuperscript{Sc} (Rigter et al., 2007, 2009; Solforsø et al., 2007), which might be important to initiate the template-assisted conversion (Figure 6B), and highly influential sites observed in the mutation experiment (Figure 6C) were thoroughly used for intersubunit interactions.

Additionally, the models were examined by referring to the results of protease restriction of mouse GPI-less PrP	extsuperscript{Sc} (Sajnani et al., 2008; Vázquez-Fernández et al., 2012). The percentages of protease restriction sites, which were solvent-accessible in each model, were relatively large for the BH3 and BH4 models (Table 1). When the out-of-model sites were also counted as correct protease restriction sites, because those sites might be in a disordered region, the IS model was the most consistent with the experimental results because most of the protease restriction sites were outside of this model. Under this criterion, the BH4 model was the second most consistent among the other models (Table 1).

**DISCUSSION**

**Coherence of Mutation Data to β Helix Models**

The five major PrP	extsuperscript{Sc} working models were evaluated using the comprehensive mutation data of mouse PrP and compared to each other in atomic detail. An intriguing methodology that correlated comprehensive mutation data to threading profiles was devised for this purpose. The experimental data revealed high consistency of particular models—those that involved conversion of the N-terminal region of core PrP into a β helix.

Preparation of in vivo PrP	extsuperscript{Sc} samples that are suitable for the ordinary structure determination methods is quite difficult. Therefore, some of the PrP	extsuperscript{Sc} working models were based on the analyses of in vitro generated PrP fibrils. The difference between converted PrP structures grown in vitro and in vivo has long been discussed, mainly in terms of their infectivity, because PrP converted in vitro (in cell-free conversion systems) tended to demonstrate very low infectivity (Baskakov and Breydo, 2007; Collinge and Clarke, 2007; Prusiner, 1998). Nevertheless, several lines of evidence have suggested common structural features shared by PrP	extsuperscript{Sc} and in vitro-converted PrP. For example, the 3D EM image, on which the BH3 model was based, was obtained from an in vitro generated PrP fibril. However, the fibril shared a similar morphology with the PrP	extsuperscript{Sc} fibril extracted from the brains of prion-infected animals (Sim and Caughey, 2009), and the antibody against the in vitro fibril component was shown to be reactive with PrP	extsuperscript{Sc} from human and mouse prion-infected brains (Makarava et al., 2010).

EPR spectra endorsing the IS model were also obtained from PrP fibrils prepared in vitro (Cobb et al., 2007; Lu et al., 2007). It was shown that the PrP	extsuperscript{Sc} extracted from the brain of a prion-infected animal demonstrated an H/D exchange profile consistent with the IS model (Smirnovas et al., 2011). Additionally, recent studies also demonstrated that some of the PrP	extsuperscript{Sc}-specific antibodies were reactive with PrPs converted in vitro (Biaisini et al., 2008), and highly infective PrP	extsuperscript{Sc} can be generated in vitro (Díaz-Espinoza and Soto, 2010; Shikiya and Bartz, 2011). In this study, therefore, the five major proposed models and two EM images were used for model evaluation by assuming that they shared structural features with infective PrP	extsuperscript{Sc}.

At this point, however, the conformational details of the PrP	extsuperscript{Sc} working models are mostly theoretical, except that the IS model is based on rigid geometric data, distance constraints between residues, obtained by EPR spectroscopy of human PrP (Cobb et al., 2007; Lu et al., 2007). The IS model assumed a drastic structural change that involved complete refolding of α helices and their conversion into in-register β sheets. The absence of α helices in PrP	extsuperscript{Sc} was recently proposed based on the reassignment of an infrared absorbance at 1,650 cm⁻¹, which was interpreted to assign 10%–17% of the α helix in PrP	extsuperscript{Sc} in previous studies, to a certain nonprotein component (Baron et al., 2011). As mentioned above, the H/D exchange mass spectroscopy experiments detected strong protection of the C-terminal regions in mouse PrP	extsuperscript{Sc} (residues W144–A223) extracted from the brain of a prion-infected animal, which could be explained only by a tightly packed β sheet structure (Lu et al., 2007; Smirnovas et al., 2011). The result of protease restriction of mouse GPI-less PrP	extsuperscript{Sc}, by indicating strong protection in the C-terminal region of core PrP, was consistent with this model (Sajnani et al., 2008; Vázquez-Fernández et al., 2012). The in-register-stacking model provides common structural bases for PrP	extsuperscript{Sc} and several amyloid proteins (Lührs et al., 2005; Sawaya et al., 2007). Therefore, it is quite interesting to investigate why the mutation data indicated a preference for the β-helical models rather than the in-register-stacking β sheet model.
The mutation data appeared to be incompatible with the IS model on two points. First, the data suggested that stabilization of the C-terminal α helices promoted conversion because 71% of inconsistent sites were found on the α helices of PrPC (Figure 5). Second and more significantly, the sites that largely affect conversion efficiency were localized to the N-terminal region of core PrP (residues ~V160) and involved in β sheet formation in the β-helix-based models (Figure 3), while the corresponding region was assumed to be disordered in the IS model (Smirnovas et al., 2011).

Several lines of evidence have accumulated for involvement of the N-terminal regions of core PrP in the formation of PrPSc. Mouse-hamster chimeric PrP revealed that residues N99–P104 were required for PrPSc conversion (Ihara et al., 2012). The motif-grafted antibodies against peptides W88–V111 and R135–P157 were shown to prevent mouse PrPSc conversion, and those against K23–T33, W98–H110, and P136–N158 prevented mouse PrPSc to PrPSc interaction (Moroncini et al., 2004; Solforosi et al., 2007). Similarly, peptide array experiments detected that sheep PPr residues W102–K104, P140–Y148, Y152–R154, Y165–R167, N177–V179, and S225–Y228 (W98–K100, P136–W144, Y148–R150, Y161–R163, N173–V175, and S221–Y224 in mouse PrP) were the self-interaction of PrPC to K100, P136–W144, Y148–R150, Y161–R163, N173–V175, and S225–Y228 (W98–K109), respectively, in (A). The α helices α1–α3, β strands β1–β2, and N and C termini (N and C) are indicated in (B). N-glycans (gray wire), and disulfide bonds (yellow ball-and-stick) are also shown in (B).

The PrPSc strains isolated from infected mouse brains revealed significant differences in infrared absorbance and protease digestion patterns (Caughey et al., 1998). In addition, a large variation in fibril morphology of PrPSc strains was detected in EM analyses (Sim and Caughey, 2009). To date, two different types of EM images have been observed for PrP fibrils. The 2D image showed 3- (or 6-) fold symmetric fibrils, where N- and C-terminal regions resided inside and outside the fibril, respectively (Wille et al., 2002). The 3D reconstructed image, however, revealed a 2-fold symmetric ladder-like structure of a mouse PrP fibril (Tattum et al., 2006), which resembled some of the twisted filament EM images observed for the mouse PrPSc (Sim and Caughey, 2009).

These images are not apparently consistent with each other. The current PrPSc working models were originally devised to be consistent with one of these images; the BH1, BH2, and SP models were consistent with the former, and the BH3 with the latter. The IS model is rather close to the 3D image, although a periodic structure with ~60 Å intervals in the 3D EM image would not be explained by simple in-register stacking of β strands (Figure 1F).

Thus, the lines of evidence suggest the requirement of multiple PrPSc models to explain the observed polymorphism.

Reconciling β Helix Models to Experimental Data

Because PrPSc conversion occurred in scrapie-infected cells in the mutation experiment of this study, the results might reflect an in vivo structure of PrPSc. Thus, reconciliation of the current β-helix-based model was also attempted in this study to find an alternative model of the PrPSc structure. It was intended to remove the questionable features of the current models, which were observed in the modeling processes or would be raised in light of recent experimental data, as far as possible.

For example, although mouse PrPSc has been shown to convert to PrPSc with core PrP (L108–L243) only, a conversion also took place for the intrinsic cellular form of PrP, which consists of highly disordered N-terminal residues of full-length PrP (K23–K109), N-glycans, and a glycosylphosphatidylinositol (GPI) moiety attached to the C terminus (Chesebro et al., 2005;
In addition, the C178–C213 disulfide bond is known to be important for PrP conversion (Herrmann and Caughey, 1998). Therefore, a plausible model should have enough space for these molecular peripherals. In this regard, the BH1 and SP models were questionable because they did not provide sufficient space for the N-terminal residues of full-length PrP to escape. In addition, the disulfide bond could not be formed in the BH3 model (Table 1).

On the other hand, the detected core structure of the N-glycan of PrP is considerably large (Stimson et al., 1999), and its packing in a molecular model is not a minor problem. For the BH1 model, spaces left for N-glycans were not sufficient due to the close packing of β2-α3 helix bundles. In the IS model, although the N-glycans point toward the outside of the complex, the in-register conformation inevitably placed both asparagines in a line, making it quite difficult to avoid an atomic crash between N-glycans if PrP is highly modified with N-glycans (Figure S2E).

The BH2 model was selected as the starting model for refinement because it seemed to be free from these problems (Table 1). Taking these problems in account, the N-terminal β helix of core PrP of the model was thought to comprise the rails of the 3D EM image because this part is essential for fibril extension (Figure 5A). The C-terminal β2-α3 helix bundle was put into the steps. The helix bundles from the subunits on different rails were arranged so that the experimentally detected self-interface residues were used for the interactions, and so that the helix bundles could be swapped between subunits (Figure 5C). This is because dissociation between the N-terminal β1-α1-β2 and C-terminal β2-α3 domains of core PrP was detected in the initializing step for sheep PrPSc formation, implying that a domain-swapped PrP dimer is a building block of PrPSc (Eghiaian et al., 2007; Lührs et al., 2005).

The persistence of PrPα helices in PrPSc might be critical because PrPβ with PrPSc interaction was shown to be required for a formation of intermediate structure in the initial step of conversion (Rigter et al., 2007, 2009; Solforosi et al., 2007). It implied that at least part of the interface in PrPSc consists of the native PrPβ structure. In the BH4 model, the C-terminal α helices, which contained the region implied for PrPβ with PrPSc interaction, take part in the intersubunit interactions. Therefore, this model can explain the introduction of PrPβ to PrPSc through this region.

Fitting a model to both 2D and 3D EM images was another intricate problem, for which none of the current models was appropriate. The analogy of crystal growth has been often mentioned for PrPSc formation: inoculation of a small fragment of a PrPSc fibril as a seeding nucleus initiates conversion of PrPβ by causing crystal-like growth (Sasaki et al., 2008). Therefore, the apparently incompatible features between the EM images were interpreted to be due to a difference in crystal systems. The BH4 model could explain the ladder-shaped 3D EM image if interface 4 was used. In the ladder structure, the sidewall of the β helix was open for interfacing, so that a trimer that resembled the 2D EM image could be formed readily in the same manner as the BH1 and BH2 models, when interface 3 was used. Thus, switching between interfaces was assumed to explain different EM images, although this analogy of protein crystallization might not explain entire of the prion polymorphism.
As the result, the BH4 model ranked highest among the current models in evaluations against the mutation data by using most of the suggested self-interface residues and the highly influential sites detected in the mutation experiment for contacts with neighboring subunits (Figures 6C and 6D). It could also explain both 2D and 3D EM images, except that the model thick-ness appeared to exceed that of the 3D EM image if the C-termi-nal helices were fully intact (Figure S5).

The BH4 model might partially explain the structural polymorphism of PrPSc with preferential usage of interfaces. The confor-mational polymorphism among prion strains has been well known (Caughey et al., 1998; Safar et al., 1998). However, the structural differences at atomic detail and the direct cause of the polymorphism are not understood. A similar structural polymorphism depending on pH difference has been observed in HET-s amyloid fibril (Mizuno et al., 2011). Because the stability of each interface might differ depending on environmental con-ditions, such as pH or amino acid sequence variation among organisms, the subunits of BH4 model would result in different fibril structures under different conditions, because a protein can be crystallized into various crystal systems depending on crystallization conditions. Therefore, the BH4 model might serve as the 3D model of PrPSc that can partially explain the conformational polymorphism. Because the PrPSc conversion was executed in ScN2a cell, the evaluated compatibility of the models should be that for RML scrapie prion, and it still remained a possibility that a similar comprehensive mutation study would provide a different result for other strains. Consequently, application of the introduced method to other prion strains would be quite interesting for further study.

Conclusions
The comprehensive mutation study of mouse PrP highlighted the involvement of the N-terminal region of core PrP in structure con-version, which was compatible with the β helix models. The revised BH4 model could explain the results of mutation exper-iments of this study, PrPc to PrPSc interactions, and PrPSc fibril polymorphism better than the previous models. The method introduced in this article is rather elaborate because it requires the existence of comprehensive mutation data and model struc-tures. Nevertheless, this method would be generally applicable for these cases in which current structural biology methods are hardly applicable.

EXPERIMENTAL PROCEDURES
Preparation of PrP Point Mutants and PrPSc Conversion Efficiency Assay
The open reading frame of the mouse PrP gene was cloned into a pBluescript plasmid (Stratagene), and 19 single amino acid substitutions at every amino acid residue from codon 102 to 240 were introduced into the mouse PrP gene by site-directed PCR mutagenesis (Stratagene) as described (Iida et al., 2008). To detect specifically the transfected mutant PrP against endog-enous mouse PrP, the epitope for the monoclonal antibody 3F4 (L108M and pSPOX (Scott et al., 1992). A scrapie-infected mouse neuroblastoma cell line (ScN2a) was kindly provided by Dr. Stanley B. Prusiner (Butler et al., 1988). ScN2a cells were transiently transfected with the plasmid constructs using the FuGENE6 transfection reagent (Roche Diagnostics) and harvested 48 h after transfection. For detection of the total cell-associated PrP, the cell lysates were digested with 300 U of PNGase F (New England Biolabs) at 37°C for 2 hr. For detection of the protease K (PK)-resistant core of PrPSc, the cell lysates were digested with 20 μg/ml PK at 37°C for 30 min and ultracentrifuged at 100,000 × g at 20°C for 1 hr, and then the pellets were resuspended in sample buffer. The samples were subjected to 13% SDS-PAGE and western blotting as described (Asano et al., 2006; Iida et al., 2008). The signal intensities of western blotting were quantified, and the amounts of PK-resistant PrP molecules were corrected by expression levels of the PrP molecule without PK digestion and normalized them to the efficiency of translation and expression efficiencies. The conversion efficiencies were calculated as the relative amount of the mutant PK-resistant PrP molecules to that of the transfected wild-type (Figure 2). The glycosylation on every mutant was also examined by PNGase F (New England Biolabs) digestion at 37°C for 2 hr.

The experiments for the mutants at 19 residue sites, namely, A115, A117, N180, T182, N196, T198, R228, R229, S230, S231, S232, T233, V234, L235, P236, S237, P238, P239, and P240, were repeated three times, and those at 12 sites, namely, A112, G113, A114, V120, G125, G126, Y127, L129, S131, W144, H176, and D177 were executed twice from transfection procedure, to examine reproducibility. For other sites, the data were collected from a single transfection experiment. Reproducibility of conversion efficiency evaluation was judged to be fine from the data of the triplicate or duplicate experiments. The example of reproducibility is shown for the A115 site in Figure 2C.

A constancy of conversion efficiency against total expression level of PrP was examined by expressing varying amounts of wild-type PrP with the 3F4- or T41-epitope-tags in the ScN2a cells (Figures 2D and 2E).

PrPSc Working Model Construction
Five of the PrPSc working models were constructed mostly from scratch by referring to the descriptions in the relevant literature (Figures 1 and S5; Cobb et al., 2007; DeMarco and Daggett, 2004; Govaerts et al., 2004; Langedijk et al., 2006; Tatum et al., 2006), MOE (Ryoka Systems) and COOT/REFMAC5 applications were used for the manual construction and refinement of the models (CCP4, 1994; Emsley et al., 2010; Murshudov et al., 1997). The core of the N-glycan moieties (Man x1-6[Glcnac x1-4]Man x1-3)Man x1-4Glcnac x1-4[Fuc x1-6]Glcnac-Asn) was modeled at N180 and N196, and a disulfide bond was made between C178 and C213 whenever possible. The models were evaluated by using Verify3D and PROCHECK (Figure S1 and Table S1). The construction procedures for each model are detailed in the Supplemental Experimental Procedures. Molecular graphics presenta-tions were prepared using CHIMERA (Pettersen et al., 2004).

Evaluation of Point Mutation Effects on 3D Models
The mutation data (Figure 3 and Table S2) were used to evaluate the PrPSc working models. To avoid possible biases from aberrant samples, the data of the mutant PrPs were discarded when they showed very low conversion efficiency or a defect in glycosylation. The mutant PrPs with very low conversion efficiency were excluded because a defect in protein folding was sus-pected. The 1,331 (47.9%) mutant and wild-type sites out of a total 2,780 (20 amino acids for 139 sites) were used for the model evaluation (Figure 3). The confirmation of proper glycosylation on mutant PrPs was used to ensure the processing of mutant proteins through Golgi apparatus. The evaluation was made in both coarse and fine scopes. In the coarse eval-uation, the amino acid residues in the models were assigned to surface, inter-o, or interface categories. If the average conversion efficiency of a site ranked within the top 50% and the categories of the site differed between PrPc and PrPSc working models or ranked within the bottom 30% and stayed in same category, the site was thought to be consistent with the corresponding model. If a site ranked in the top 50% and the site categories did not differ or ranked in the bottom 30% and the categories were different, it was considered inconsistent. The other sites were assigned as ambivalent (Table 1).

In the fine evaluation, threading profiles were calculated for each model and used to correlate the difference in fitness (difference in virtual free energy) of amino acids to the changes in PrPSc conversion efficiency on point mutation (Ota et al., 2001). The tendency toward stabilization/destabilization of a point mutation was evaluated as ΔG = [score of substituted amino acid in model] – [score of wild-type amino acid in model] for PrPc and PrPSc working models. Then, the effective correlation coefficient ρp = −log p between ΔG
and the logarithm of conversion efficiency $\log_{10} C$ were evaluated, where $p_i$ and $r_i$ are the p value and the correlation coefficient of site $i$, respectively. Sites with $p_i < 0.1$ (promoting conversion by stabilizing PrPSc) or $p_i > 0.1$ (preventing conversion by destabilizing PrPPr) were regarded as consistent or inconsistent with the corresponding PrPPrSc working model, respectively. For PrPSc experimental structures, a site was consistent or inconsistent when $p_i > 0.1$ (promoting conversion by destabilizing PrPSc) or $p_i < -0.1$ (preventing conversion by destabilizing PrPPr), respectively (Table 1). The average of the correlation coefficients over the sites, $\bar{r}$; $p_i/n$, was also evaluated for each PrPPrSc working model, where $n$ was the number of counted sites in each model.

Additionally, the models were evaluated by referring to the results of protease restriction of mouse GPI-less PrPSc (Sajnani et al., 2008; Vázquez-Fernández et al., 2012). The percentage of exposed protease restriction sites was obtained for each model. Because disordered residues might be potential restriction sites, the percentage of restriction sites, which were exposed on a model or were not modeled, was also evaluated for each model (Table 1).

The methods are detailed in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.str.2013.12.019.

AUTHOR CONTRIBUTIONS

T.S. and M.S. designed and executed the computational analyses; A.K., M.A., M.H., S.I., K.T., and M.M. performed the mutation and assay experiments; K.T, M.M., and T.K. designed the experiments; and T.S. wrote the manuscript. All authors commented on the manuscript.

ACKNOWLEDGMENTS

This work was partly supported by research grants from the Institute for Bioinformatics Research and Development (IBiRD); the Japan Science and Technology Agency; and Platform for Drug Design, Informatics, and Structural Lifescience (PDIS), the Ministry of Education, Culture, Sports, Science & Technology in Japan.

Received: August 21, 2013
Revised: December 21, 2013
Accepted: December 28, 2013
Published: February 20, 2014

REFERENCES


Please cite this article in press as: Shirai et al., Evaluating Prion Models Based on Comprehensive Mutation Data of Mouse PrP, Structure (2014), http://dx.doi.org/10.1016/j.str.2013.12.019. Structure 22, 1–12, April 8, 2014 ©2014 Elsevier Ltd All rights reserved 11


