BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **238,** 185–190 (1997)
ARTICLE NO. RC977243

A Hypothesis Describing a Potential Link between Molecular Structure and TSE Strains

James Warwicker
Institute of Food Research, Reading Laboratory, Whiteknights Road, Reading RG6 6BZ, United Kingdom

Received August 4, 1997

In considering a protein-only model for prion pathogenesis in TSEs, one key challenge is to explain the existence of strains. These have traditionally been characterised by neuropathology and incubation times and more recently through biochemical analysis of prion protein (PrP), which shows differences in protease-resistant fragment size and glycoform ratios. It is now suggested that PrP possesses two faces which on the basis of conservation and non-polar nature could each (physiologically) interact either with membrane or with neighbouring protein. This model leads to the construction of two clearly different membrane-attached PrP orientations, with consequences for protease resistance and glycoform incorporation that qualitatively match to experiment. © 1997 Academic Press

The widespread outbreak of bovine spongiform encephalopathy (BSE) in the UK (1), and the possibility that this may have transmitted to humans in the form of new variant Creutzfeldt-Jakob disease (nvCJD) (2), has intensified research effort into the pathogenesis of transmissible spongiform encephalopathies (TSEs). Many questions remain unresolved, including the underlying issue of the nature of the infectious agent, although a large amount of data point to the protease-resistant form of PrP (3). This is denoted PrP^{Sc} in the current article, in contrast to the protease-sensitive form PrP^{C}; (alternative nomenclature for these PrP forms includes PrP-res and PrP-sen). However, development of TSEs has been observed in the absence of detectable amounts of PrP^{Sc}, supporting the view of some researchers that another element is involved (4,5). One challenge for a protein-only model is to give a plausible explanation for the existence of strains. This is most convienently considered within the framework of PrP^{Sc} biochemistry. Both the size of the deglycosylated protease-resistant fragment of PrP^{Sc} and the glycoforms composition (none, one or both N-linked sites glycosylated) vary (2,6,7). A complete one-to-one mapping between such analysis and strain characterisation by neuropathology and incubation time may ultimately not be possible (8), but a clear correlation exists, particularly in regard to the protease-resistant fragment size (9). Whilst the basis of any link between neuropathology and PrP^{Sc} biochemistry is unknown, it is possible that each cell-type is more susceptible when intrinsic glycoform ratios match those of the infecting PrP^{Sc} (2).

Within the framework of a PrP-based model for transmission, it is useful to consider what elements of the protein/glycosylation/membrane system are required. The presence of lipids appears to preserve scrapie infectivity within fractions of the PrP 27-30 fragment (10). Efficient formation of PrP^{Sc} requires glycosylphosphatidylinositol (GPI) anchors in scrapie-infected mouse neuroblastoma cells (11), although absence of the GPI anchor does not prevent conversion of PrP^{C} to PrP^{Sc} in a cell-free system (12). The observation that both PrP^{C} and PrP^{Sc} are attached to detergent-insoluble complexes or 'rafts' (13) is consistent with an important role for protein-membrane interactions in TSE pathogenesis. Although glycoform ratios appear to have some link to TSE strains, glycosylation of PrP^{C} at the two N-linked sites is not required for the cell-free conversion reaction (12). In PrP itself, residues around 115-220 of PrP^{Sc} form a folding core (14), consistent with folding (15) and nmr (16) studies of the mouse PrP fragment 121-231. Whilst mutations outside of this region can be associated with TSEs, for example P102L or insertion of additional copies of the amino-terminal (N-t) repeats (17), elements within the folding core appear to be critical for the conversion reaction (14). Disruption of the N-t repeats does not remove susceptibility to scrapie in transgenic mice (18). Interestingly, deletion of 141-176 within the folding core does not prevent the development of protease-resistant protein in scrapie-infected mouse neuroblastoma cells (19), although it is more soluble than wild-type PrP^{Sc}.

This article asks which features within the putative PrP^{C} folding core could, in the context of membrane attachment, contribute to the biochemical characteristics associated with strains. Our modelling is based

0006-291X/97 $25.00

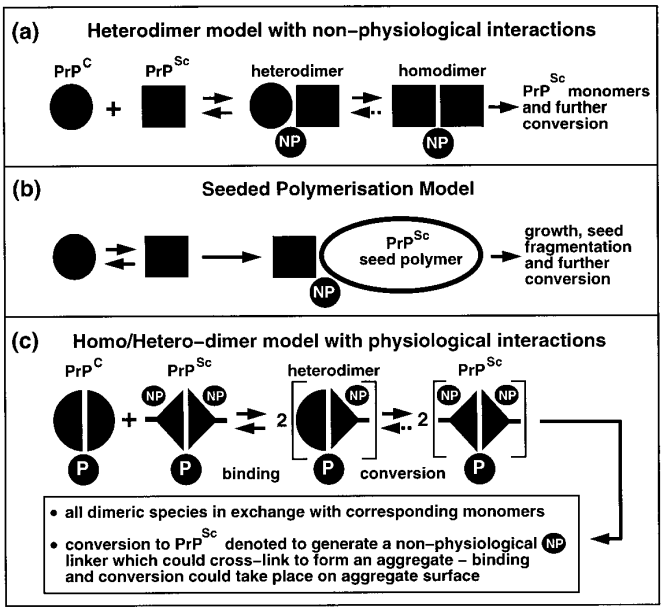

FIG. 1. Comparison of models for PrP^C to PrP^{Sc} conversion. (a) The heterodimer model (23) uses non-physiological (NP) interactions and conversion occurs after binding. (b) Binding follows conversion in the seeded polymerisation model (24), which also uses non-physiological interactions. (c) The homo/hetero-dimer model (22) includes elements of (a) and (b), but adds physiological (P) interactions for PrP^C-PrP^{Sc} binding. Conversion follows binding and may be aided by the presence of an aggregate.

on a protein-only theory of pathogenesis, within which some part of the PrP^C/PrP^{Sc} interaction is hypothesised to mimic a physiological PrP^C/PrP^C interaction (20-22). Figure 1 illustrates the relationship between this (physiological dimer) model and those related to a heterodimer (23) and to seeded polymerisation (24). Non-physiological links, derived from the PrP^C to PrP^{Sc} structural transition, are incorporated into the physiological dimer model to build an aggregate of dimers. Key parts of this model are that binding of PrP^C to a PrP^{Sc} aggregate precedes the conversion step, that it mimics a physiological interface, and that the aggregate could facilitate the conversion. The model therefore borrows from both heterodimer (23) and seeded polymerisation models (24). Crucially, the addition of a physiological interface gives a molecular mechanism for the infectivity of TSEs. Similarities exist between TSEs and other amyloid forming diseases, but unlike PrP^{Sc}, the amyloid protein itself is generally not a transmissible agent (3,25). The physiological interface of the model in fig. 1c, and the binding of PrP^C to PrP^{Sc} before conversion, could explain the tendency towards infectivity in TSEs relative to other amyloidoses. The mechanism of fig. 1(b), with conversion prior to binding across a non-physiological interface, may be more prominent in amyloidoses which are less associated with infectivity. It should be noted that the model elements of fig. 1(c) could also be derived from physiological oligomers of higher order than dimers.

This article will show that the structure of the PrP^C folding core is consistent with two distinct interaction faces, which could couple with membrane binding to give two orientations of the protein on the membrane, and account for the observed size characteristics of protease-resistant fragments of PrP^{Sc}.

METHODS

The refined C_α coordinates that correspond to the nmr structure for 121-231 of mouse PrP (16) have now been released, so that the earlier low resolution reconstruction (22) is not required. The highly conserved non-polar element (109-122), that is sequentially adjacent to the nmr structure, has been placed in the region of a non-polar surface in the 3D structure (22). Molecule display used the program QUANTA (Molecular Simulations Inc., 200 Fifth Avenue, Waltham, MA 02154 USA) running on a Silicon Graphics Indigo workstation.

RESULTS AND DISCUSSION

Figure 2 shows the refined C_α coordinates for residues 124-226 of mouse PrP (16), with 3 different characteristics plotted by colour-coding. Display of amino acid polarity in panel (a) is interesting with regard to the non-polar regions that could form part of membrane or protein interfaces. Large non-polar stretches are seen on the extended chain running into helix 1 on the left hand side and at the fragment N-t, recalling that the conserved segment adjacent to the nmr structure N-t is also non-polar. Shorter non-polar regions occur largely on the remaining two α-helices and may be associated with packing the folding core (16). Panel (b) shows conserved amino acids, colouring separately those around the disulphide and the two N-linked glycosylation sites. Conservation is plotted for mammalian and avian sequences (21,26). This display again highlights the folding core, the fragment N-t region and the region around helix 1. Fig. 2(c) shows the location of 6 amino acids that may be associated with species barriers or with disease modifying polymorphisms, within the PrP fragment (22). Five of these are around the helix 1 and fragment N-t regions indicated in panels (b) and (c).

Non-polar surface areas around helix 1 and around the fragment N-t are interesting with regard to the potential for membrane interactions that are additional to the GPI anchor. Positive charge (see fig. 2(a)) could also play a role. In addition, a relatively high degree of amino acid conservation in these regions, together with the siting of species barrier residues and disease-modifying polymorphisms, argues that they could also be involved in protein-protein interactions. Our modelling of the infectious process suggests that this could involve PrP^C/PrP^C as well as PrP^{Sc}/PrP^C. A study based on mammalian PrP sequence variation, as opposed to conservation across a wider species range,

has been used to argue that helix 1 and the region around 168 and the GPI anchor (but not the fragment N-t) could be PrP^{Sc}/PrP^{C} (but not necessarily PrP^{C}/PrP^{C}) interaction sites (26). This philosophy of looking for variation in a non-physiological interaction is different to our strategy of searching for conservation and physiological interactions that could be recruited by the disease process.

Figure 3 demonstrates that the two interaction faces suggested by fig. 2 are roughly orthogonal. Fig. 3(a) shows an orientation similar to that of fig. 2, with the two putative interaction faces and the conserved non-polar segment (109-122) marked schematically. Panels (b) and (c) show that modelling either face in contact with the membrane would be consistent with GPI anchor location (5 amino acids from the fragment C-t), and would leave the other face in a position to make interactions across a symmetry axis perpendicular to the membrane. Face 2 denotes the region around helix 1, and face 1 that around the fragment N-t and the non-polar segment 109-122 and also including the region around 168, (thus accounting for the sixth marked site in fig. 2(c)), since it lies close to the presumed membrane plane with face 1 docked (fig. 3(b)). This theoretical analysis suggests that PrP^{C} could accommodate two different orientations on a membrane, both consistent with GPI-anchoring, by swapping faces 1 and 2 between PrP/membrane and PrP/PrP interactions.

In the context of a model for disease pathogenesis based on PrP dimerisation, the new hypothesis introduces a membrane environment and two physiological PrP-PrP links to replace one, and again it is easiest to interpret them in terms of dimerisation, although other oligomerisations or protein-protein interactions would also be possible. These two physiological links are added to the non-physiological dimer cross-linking interactions that the model presumes to accompany α-helix to β-sheet transition and to be associated with pathogenic mutations in the PrP folding core (16). Figure 4 shows the proposed interaction faces of fig. 3 incorporated into schematic dimers and dimer aggregates on a membrane surface. The proposed pathogenic mechanism of fig. 1 now applies to each of the dimeric species (face 1 or face 2 dimers) drawn in figure 4. The PrP^{C} to PrP^{Sc} transition has been reported to involve PrP-PrP interactions between amino acids 96 and 167, and interaction with another protein, which may bind near to the PrP carboxy-terminus (C-t) (27). Figures 1 and 4 illustrate a model for the transition which uses PrP-PrP interfaces away from the C-t, and in which other proteins could be involved in PrP α-helix to β-sheet conformational change that is separate from these interfaces.

Currently, there is no direct experimental evidence for functional PrP^{C} dimerisation, with or without membrane interactions. From the various indirect lines of evidence shown in figure 2 and the hypothesis that TSE infectivity could be explained by recruitment of physiological interactions, the model developed in this article predicts that at least two interaction surfaces exist, either of which could interact with membrane or protein. Since the prediction relies on amino acid polarity and conservation, it can be tested with mutagenesis experiments linked to (perhaps cell-free) assays of PrP conversion. In what way two orientations of PrP^{C} on the membrane would contribute to function is not clear, but conformational switching is common in biological systems, often in response to ligand binding or signalling. For example, the two predicted orientations could be related to function resident in the N-t repeat regions. The two sets of N-t repeats within a putative face 1 dimer (fig. 3(c)) would be opposed and unlikely to interact, whereas there would be more scope for interaction within a putative face 2 dimer (fig. 3(b)). It is unlikely that the N-t repeats fold to a single average structure or that they would follow the strict symmetry of any PrP^{C} folding core dimerisation/oligomerisation, giving a general indication of the important role that modelling could play in bringing the various structural features of PrP^{C} together into a molecular basis for function.

Since the N-t repeats are not essential for the PrP^{C} to PrP^{Sc} conversion (18), it is likely that the folding core holds the key, and it is sensible to ask whether the predicted orientations of figure 3 correlate with the properties of protease-resistance and glycoform ratios that are associated with strains. The two predicted types of homogeneous PrP^{Sc} aggregate are shown in figure 4. It is clear (see also fig. 3) that PrP would extend farther from the membrane in the face 2 dimer than it would in the face 1 dimer, so that proteolysis of a face 2 dimer aggregate would, (when considering closest protease approach to the membrane), be likely to yield larger residual fragments than proteolysis of a face 1 dimer aggregate. This is consistent with characterisations of deglycosylated protease-resistant PrP^{Sc} fragments of variable size (2,6,7), which appear to be dominated by just two overall molecular weight (MW) categories in some cases (9).

The two N-linked glycosylation sites (181 and 197) face more toward the predicted aggregate exterior within the face 1 dimer than within the face 2 dimer (fig. 3), and therefore glycosylated PrP may be more easily accommodated. The model therefore predicts that more glycosylation should be correlated with lower MW (deglycosylated) protease-resistant forms, since both features would derive from the face 1 dimer aggregate. This is largely the case (2,9), although glycoform ratios vary across protease-resistant fragments that collapse to a similar low MW when deglycosylated. Such variation is not surprising within the model since the face 1 dimer aggregate could accommodate a range of glycoform ratios, which would presumably be modulated by specific cell-type and glycoform production.

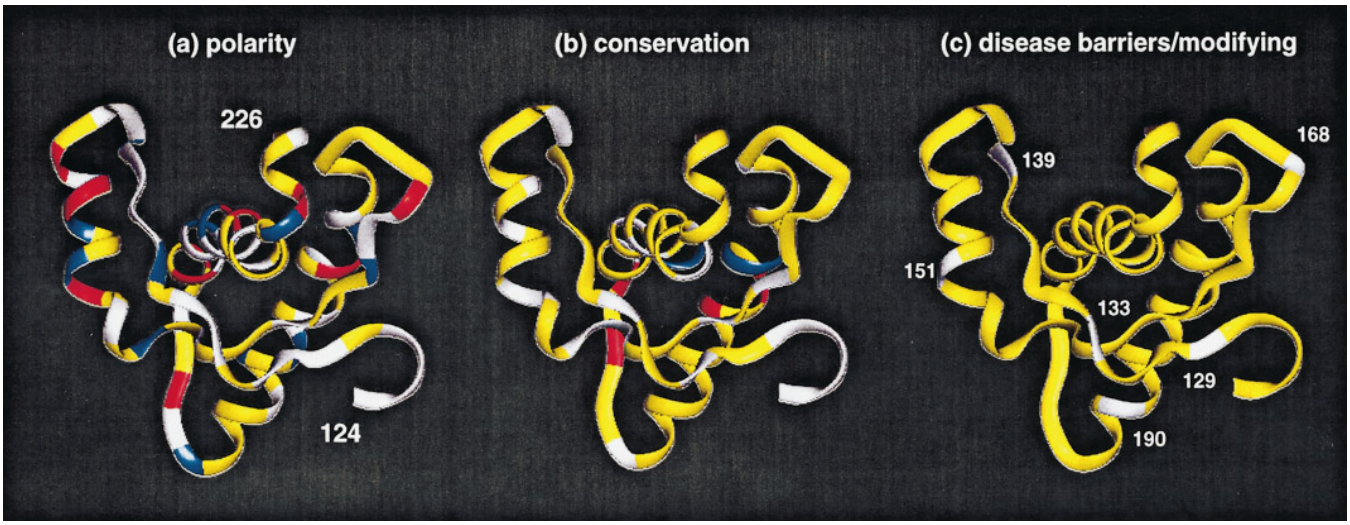

FIG. 2. Determination of potential interaction surfaces in the PrP fragment structure (16). One view is repeated in all panels, with helix 1 of the nmr structure running approximately vertically on the left-hand side. (a) The refined C_α trace is shown colour-coded by amino acid polarity: white, non-polar; yellow, polar non-charged; red, negatively charged; blue, positively charged. The N-t (124) and C-t (226) of the coordinate set are marked. (b) Conserved amino acids across mammalian and avian sequences are shown: in red for those associated with the N-linked glycosylation sites; in blue for those associated with the disulphide; and in white for other conserved residues. (c) Six amino acids associated with species barriers or disease-modifying polymorphisms are marked in white.

The model predicts that the face 2 dimer aggregate, with the higher MW deglycosylated protease-resistant fragment, would be more restrictive and tend towards less glycosylation, consistent with experiment (2,9).

The homogeneous PrP^{Sc} aggregates drawn in figure 4 would each give an approximately uniform size of deglycosylated protease-resistant fragment, which is consistent with some PrP^{Sc} preparations (9). The model would not necessarily exclude a situation in which the various thermodynamic and kinetic factors that control PrP^{Sc} formation combine to direct proportions of both face 1 and face 2 dimers into the same aggregate, if the non-physiological cross-link could be made between the two dimer forms. Since the model would still require binding before conversion, a relatively uniform proportion of face 1 and face 2 dimers could be propagated, consistent with preservation of TSE strains. In this regard it will be interesting to study the range of sheep scrapie variants with detailed MW characterisations (28). If reported TSE strain alteration (5) relates to changes in faces 1 and 2 dimer composition, then the model should incorporate some flexibility with regard to modification within an aggregate.

Where TSE develops as a result of infection, an important model factor would be matching the existing aggregate dimer interfaces (surmounting the species barrier). Both BSE and nvCJD give lower MW deglycosylated protease-resistant PrP^{Sc} forms and a higher proportion of doubly glycosylated PrP^{Sc} (2), consistent with face 1 dimers in the model. Should BSE turn out to be the cause of nvCJD, then the apparently increased susceptibility for human MM homozygotes at 129 (2) would in this model derive from matching the bovine sequence across the face 1 dimer interface.

Separation of phenotypes for the D178N mutation dependent on M or V at 129 (29) is also consistent with the MW/glycoforms ratio correlation of the model. The 129M polymorphism on the D178N mutant allele links to fatal familial insomnia (FFI), whereas combination with 129V links to one familial type of CJD (29). Smaller MW (protease-resistant fragment) and more glyclosylation is associated with FFI relative to D178N CJD (30), consistent with FFI forming face 1 dimers and D178N CJD forming face 2 dimers in the model. In contrast to the case of matching dimer interfaces with an infectious inoculate, the model provides the outline but not yet the detail associated with directing PrP^{Sc} aggregate formation for inherited or sporadic forms of TSE. For example, in sporadic CJD, MM 129 homozygotes tend towards larger MW fragments (face 2 dimers in our model), and VV towards smaller MW fragments (face 1 dimers in our model) (9). The complexity of TSEs pathogenesis is shown by this comparison of the various human susceptibilites in regard to the polymorphism at 129. It should be noted that the model can in principle generate matching complexity by placing 129 in an interface with protein or with membrane, either of which could assume greater importance depending on the particular mode of pathogenesis. If face 1 and face 2 dimers are both involved in PrP natural function, then any differences in dimer energetics associated with the naturally occurring polymorphisms cannot be significant, whereas in the dis-

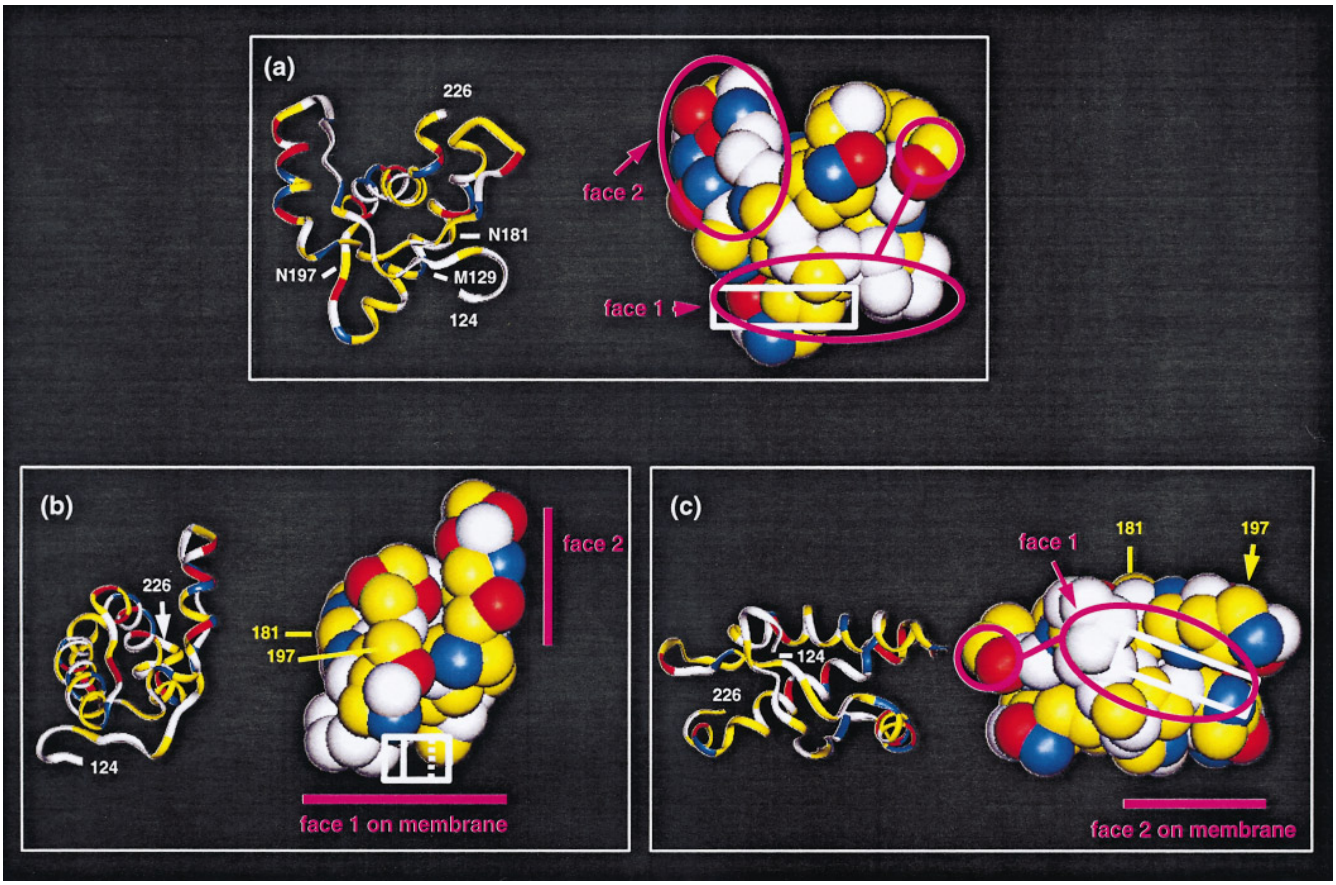

FIG. 3. Putative interaction surfaces and PrP orientation on a membrane. Each panel shows one view as a C_α trace and also one space filled around each residue. Colour coding is by amino acid polarity (see fig. 2(a)) throughout. The fragments N-t and C-t are marked on each C_α trace. N-linked glycosylation sites (181 and 197) are marked on the C_α trace of (a), and on the space-filled views of (b) and (c). The conserved non-polar sequence element 109–122 is marked schematically as a white box, docked to a non-polar surface (22). (a) A view similar to that of figure 2, with the proposed interaction faces (1 and 2, see text) marked in purple. Residue 129 is marked on the C_α trace. (b) A view with face 1 adjacent to the membrane and looking along face 2, which is roughly perpendicular to face 1. (c) A view with face 2 adjacent to the membrane and looking into face 1.

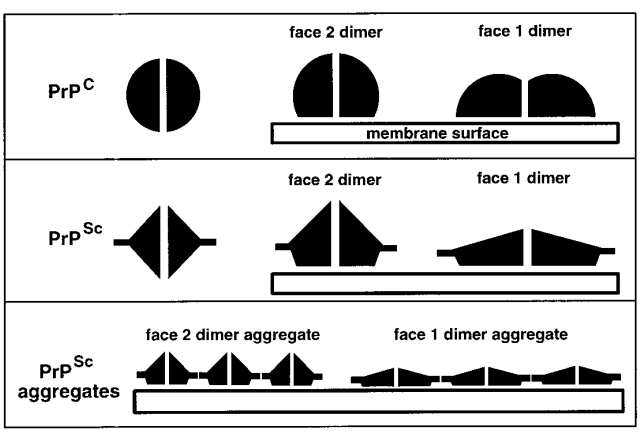

FIG. 4. Schematic of proposed PrP face 1/2 dimer forms and aggregates on a membrane. Representations of PrP^C and PrP^{Sc} dimers in fig. 1 are split into face 1 and face 2 dimers that are also proposed to bind to the membrane surface. PrP^{Sc} aggregates are drawn from cross-linked face 1 dimers or cross-linked face 2 dimers.

ease process a 'freezing out' of specific dimer forms would be dependent on such differences.

Although the detailed mapping from observed MW forms to multiple glycoform ratios and to multiple strains remains obscure, accumulated experimental evidence shows the importance of two underlying MW forms. For example, transmission of one form of sporadic CJD to transgenic animals consistently reproduces the size but not the glycoform pattern of the original PrP^{Sc} inoculate (7,9). The current model, based on sequence and structural studies, gives one possible molecular basis for the observed MW forms, and therefore provides a framework in which to conduct further experiments.

REFERENCES

1. Wells, G. A. H., and Wilesmith, J. W. (1995) *Brain Pathology* **5,** 91–103.
2. Collinge, J., Sidle, K. C. L., Meads, J., Ironside, J., and Hill, A. F. (1996) *Nature* **383,** 685–690.

3. Caughey, B., and Chesebro, B. (1997) *Trends in Cell Biology* **7,** 56–62.
4. Lasmézas, C. I., Deslys, J.-P., Robain, O., Jaegly, A., Beringue, V., Peyrin, J.-M., Fournier, J.-G., Hauw, J.-J., Rossier, J., and Dormont, D. (1997) *Science* **275,** 402–405.
5. Manuelidis, L., Fritch, W., and Xi, Y.-G. (1997) *Science* **277,** 94–98.
6. Bessen, R. A., Kocisko, D. A., Raymond, G. J., Nandan, S., Lansbury, P. T., and Caughey, B. (1995) *Nature* **375,** 698–700.
7. Telling, G. C., Parchi, P., DeArmond, S. J., Cortelli, P., Montagna, P., Gabizon, R., Mastrianni, J., Lugaresi, E., Gambetti, P., and Prusiner, S. B. (1996) *Science* **274,** 2079–2082.
8. Somerville, R. A., Chong, A., Mulqueen, O. U., Birkett, C. R., Wood, S. C. E. R., and Hope, J. (1997) *Nature* **386,** 564.
9. Parchi, P., Capellari, S., Chen, S. G., Petersen, R. B., Gambetti, P., Kopp, N., Brown, P., Kitamoto, T., Tateishi, J., Giese, A., and Kretzschmar (1997) *Nature* **386,** 232–233.
10. Riesner, D., Kellings, K., Post, K., Wille, H., Serban, H., Groth, D., Baldwin, M. A., and Prusiner, S. B. (1996) *J. Virol.* **70,** 1714–1722.
11. Kaneko, K., Vey, M., Scott, M., Pilkuhn, S., Cohen, F. E., and Prusiner, S. B. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 2333–2338.
12. Kocisko, D. A., Priola, S. A., Raymond, G. J., Chesebro, B., Lansbury, P. T., and Caughey, B. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3923–3927.
13. Naslavsky, N., Stein, R., Yanai, A., Friedlander, G., and Taraboulos, A. (1997) *J. Biol. Chem.* **272,** 6324–6331.
14. Kocisko, D. A., Lansbury, P. T., and Caughey, B. (1996) *Biochemistry* **35,** 13434–13442.
15. Hornemann, S., and Glockshuber, R. (1996) *J. Mol. Biol.* **261,** 614–619.
16. Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R., and Wüthrich, K. (1996) *Nature* **382,** 180–182.
17. Baldwin, M. A., Cohen, F. E., and Prusiner, S. B. (1995) *J. Biol. Chem.* **270,** 19197–19200.
18. Fischer, M., Rülicke, T., Raeber, A., Sailer, A., Moser, M., Oesch, B., Brandner, S., Aguzzi, A., and Weissmann, C. (1996) *EMBO J.* **15,** 1255–1264.
19. Muramoto, T., Scott, M., Cohen, F. E., and Prusiner, S. B. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 15457–15462.
20. Hardy, J. (1991) *Trends in Neurosciences* **14,** 423–424.
21. Warwicker, J., and Gane, P. J. (1996) *Biochem. Biophys. Res. Comm.* **226,** 777–782.
22. Warwicker, J. (1997) *Biochem. Biophys. Res. Comm.* **232,** 508–512.
23. Prusiner, S. B. (1991) *Science* **252,** 1515–1522.
24. Caughey, B., Kocisko, D. A., Raymond, G. J., and Lansbury, P. T. (1995) *Chemistry & Biology* **2,** 807–817.
25. Horwich, A. L., and Weissman, J. S. (1997) *Cell* **89,** 499–510.
26. Billeter, M., Riek, R., Wider, G., Hornemann, S., Glockshuber, R., and Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 7281–7285.
27. Telling, G. C., Scott, M., Mastrianni, J., Gabizon, R., Torchia, M., Cohen, F. E., DeArmond, S. J., and Prusiner, S. B. (1995) *Cell* **83,** 79–90.
28. Bossers, A., Belt, P. B. G. M., Raymond, G. J., Caughey, B., de Vries, R., and Smits, M. A. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 4931–4936.
29. Petersen, R. B., Parchi, P., Richardson, S. L., Urig, C. B., and Gambetti, P. (1996) *J. Biol. Chem.* **271,** 12661–12668.
30. Monari, L., Chen, S. G., Brown, P., Parchi, P., Petersen, R. B., Mikol, J., Gray, F., Cortelli, P., Montagna, P., Ghetti, B., Goldfarb, L. G., Gajdusek, D. C., Lugaresi, E., Gambetti, P., and Autilio-Gambetti, L. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2839–2842.