Evolutionarily Conserved Proline Residues Impede the Misfolding of the Mouse Prion Protein by Destabilizing an Aggregation-competent Partially Unfolded Form

Suman Pal [1] **and Jayant B. Udgaonkar** [1,2]∗

1 - *Indian Institute of Science Education and Research,* Pune, India
2 - *National Centre for Biological Sciences,* Tata Institute of Fundamental Research, Bengaluru, India

Correspondence to Jayant B. Udgaonkar:∗Indian Institute of Science Education and Research, Pune, India.
jayant@iiserpune.ac.in (J.B. Udgaonkar)
https://doi.org/10.1016/j.jmb.2022.167854
Edited by Sheena E. Radford

Abstract

The misfolding of the prion protein has been linked to several neurodegenerative diseases. Despite extensive studies, the mechanism of the misfolding process remains poorly understood. The present study structurally delineates the role of the conserved proline residues present in the structured C-terminal domain of the mouse prion protein (moPrP) in the misfolding process. It is shown that mutation of these Pro residues to Ala leads to destabilization of the native (N) state, and also to rapid misfolding. Using hydrogen–deuterium exchange (HDX) studies coupled with mass spectrometry (MS), it has been shown that the N state of moPrP is in rapid equilibrium with a partially unfolded form (PUF2*) at pH 4. It has been shown that the Pro to Ala mutations make PUF2* energetically more accessible from the N state by stabilizing it relative to the unfolded (U) state. The apparent rate constant of misfolding is found to be linearly proportional to the extent to which PUF2* is populated in equilibrium with the N state, strongly indicating that misfolding commences from PUF2*. It has also been shown that the Pro residues restrict the boundary of the structural core of the misfolded oligomers. Overall, this study highlights how the conserved proline residues control misfolding of the prion protein by modulating the stability of the partially unfolded form from which misfolding commences.

© 2022 Published by Elsevier Ltd.

Introduction

The mammalian prion protein (PrP) is a highly conserved glycoprotein expressed mainly on the neuronal cell surface.[1] The mouse prion protein (moPrP) is 208 residues long, GPI (glycophosphatidylinositol)-anchored, and α helix rich.[2–3] It consists of two distinct domains, an unstructured N-terminal domain (NTD) and a structured C-terminal domain (CTD). The CTD consists of three α helices (α1, α2, α3) and two anti-parallel β-strands (β1 and β2).[3] A disulfide bond between Cys178 and Cys213 (mouse numbering is used throughout) links α2 and α3.[3–4] The physiological function of the prion protein is yet to be determined.

Studies done so far suggest that the prion protein may not be an essential protein and might have a redundant function.[5–8] Misfolding of the prion protein has been linked to several fatal neurodegenerative diseases, collectively known as Transmissible Spongiform Encephalopathies (TSE).[9]

The misfolded prion protein can exist in at least two β-sheet-rich forms: an amyloid fibrillar form and a soluble oligomeric form.[10] Notably, misfolding of the prion protein has been shown to take place in the endocytic pathway, when the monomeric protein enters into the lysosome and encounters a low pH.[11] *In vitro* studies have shown that the prion protein readily undergoes misfolding and oligomerization at acidic pH, upon His186 becoming

protonated with a characteristic transition midpoint at pH 4.7.[12] Hence, while oligomerization is observable at one pH unit above the transition midpoint,[12] it is not observable at physiological pH at which only a miniscule fraction of prion protein molecules would have the His186 side-chain in its protonated form. It is important to note that the propensity of different mammalian prion proteins to form β-sheet-rich oligomers at acidic pH correlates well with susceptibility to prion disease.[13] Oligomers are also known to be more infectious than fibrils.[12] Hence, understanding the molecular mechanism of prion misfolding leading to oligomer formation is important for the development of therapeutics against prion diseases, where little success has been achieved so far despite extensive efforts.

One approach that can be taken to understand the mechanism of misfolding of the prion protein is to study the role of evolutionarily conserved amino acid residues in controlling the misfolding process. It is known that the sequences of proteins have evolved to safeguard against misfolding and aggregation.[14–16] Evolutionarily conserved residues at specific sequence positions, which act as "gatekeeper" residues[15] that impede misfolding, include the charged amino residues Arg, Lys, Asp and Glu,[14–15] as well as Pro.[14–16]

It is well known that a Pro residue may play an important role in protein folding reactions,[17–18] but Pro also appears to be the most effective gatekeeper residue against misfolding.[14,16] Since its backbone nitrogen is linked covalently to its side chain, generating a cyclic pyrrolidine ring,[19] its presence disrupts secondary structural units, both α-helices and β-strands, by introducing a strong local twist to the peptide bond.[16] Consequently, Pro residues are often found at the ends of α-helices and β-strands, and they disfavor intermolecular β-β interactions when present at the end of the β-strands.[16] A Pro residue also introduces rigidity to the peptide backbone,[16,19] by restricting the main chain dihedral angle, φ, of the peptide bond; thus, it can play an important role in the misfolding process, as observed in the case of fibronectin type-III[20] and α-synuclein.[21] Notably, the cyclic pyrrolidine ring of Pro often causes the *trans* conformer of the peptide bond to be not as strongly favored as it is by other amino acid residues[22]; consequently, a Pro residue can adopt a *cis* conformation more readily than any other amino acid residue.[22] *Cis/trans* peptidyl proline isomerization can modulate protein misfolding and aggregation in multiple ways. For example, an increase in the *trans* population of Pro32 in the case of β2-microglobulin induces the formation of higher-order structures.[23] By altering the binding of phosphatase or kinase enzymes, proline isomerization can also indirectly contribute to the propensity to form aggregates.[24]

Elucidation of the role played by conserved Pro residues in modulating the misfolding of the prion protein, is expected to provide a better molecular understanding of the misfolding process, in which α-helical structure is converted into β-sheet structure.

The sequence of the mammalian prion protein has several conserved Pro residues: residues 50, 59, 67, 75, 83, 101 and 104 in the NTD, as well as 136, 157, and 164 in the CTD. Of the seven Pro residues in the NTD, five are present in the highly conserved five-octapeptide repeat region that spans residues 50–90. This repeat region is thought to be important in prion function[25] and in modulating aggregation.[26–27] Proline residues 101 and 104 are present in the highly conserved central lysine cluster spanning residues 100–109. Mutation of Pro101 and Pro104 is associated with the most common prion disease, Gerstmann-Straussler-Scheinker disease.[28] All three Pro residues in the CTD are highly conserved across different mammalian species; moreover, sequence segment XPNXVY that contains Pro157 has a higher than average sequence conservation.[29] Pro136 is present in the loop between β1 and α1, Pro157 is present in the loop between α1 and β2 (Figure 1(a) and (b)), and Pro164 is located on the edge of β2. β2 is flanked by Pro157 and Pro164 (Figure 1(a)). It seems that Pro164 may terminate the expansion of the β2 strand due to conformational constraints imposed on the peptide backbone. Unlike in the case of the conserved Pro residues in the NTD, which have been implicated in misfolding,[26–28,30–31] little is understood of the role of highly conserved Pro residues in the CTD of the prion protein.

In the current study, it is shown that the conserved proline residues in the structured CTD of moPrP play an important role in maintaining the integrity of the protein structure. Replacement of the proline residues by alanine destabilizes the native protein, and accelerates the misfolding process by up to 900-fold. Hydrogen deuterium exchange (HDX) studies coupled with mass spectrometry (MS) show that the native (N) state of wt moPrP is in dynamic equilibrium with a partially unfolded form (PUF2*) in which α1, the loop between α1 and β2, β2 and the N-terminal end of α3 are disordered. PUF2* is stabilized and therefore becomes more accessible from the N state upon replacement of the Pro residues with Ala. It is also shown that the Pro residues play an important role in determining the boundary of the structural core of the misfolded oligomers, and that the Pro to Ala mutations result in an expanded and more ordered core. The results of the current study suggest that the Pro residues in the CTD have been evolutionarily selected for reducing the probability of misfolding of the prion protein.
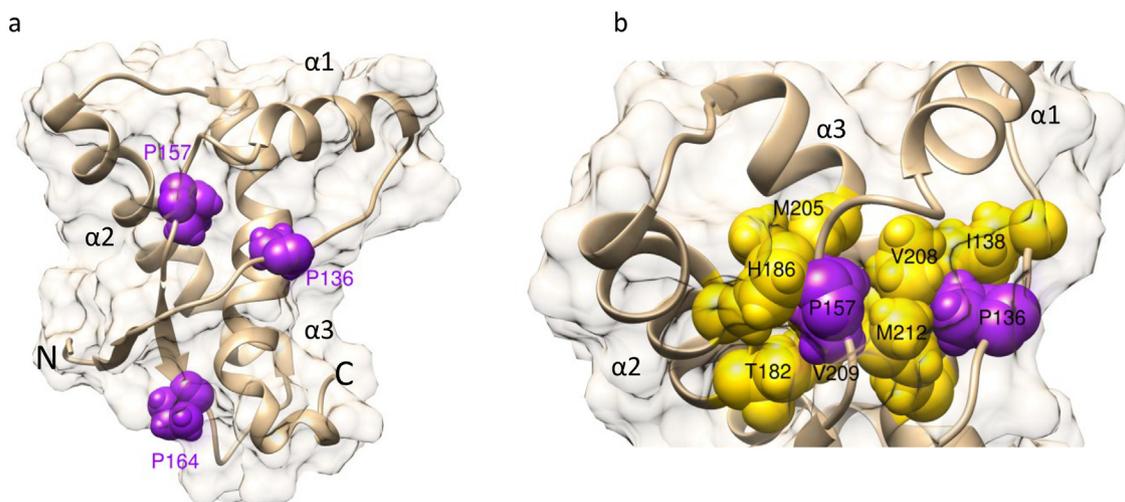
Figure 1. Structure of the C-terminal domain of the mouse prion protein showing the locations of the three proline residues (Pro136, Pro157 and Pro164). The N-terminal domain comprises residues 23–120, and is unstructured in the full-length protein. (a) The three proline residues that were mutated in this study are colored purple. All three are in the *trans* conformation. N and C represent the termini of the CTD. (b) The side-chains of Pro136 and Pro157 are buried to the extent of 92% and 97%, respectively, and appear to interact with Ile138, His186, Thr182, Met205, Val208, Val209, and Met212 (yellow). The side-chain of Pro164 is buried to the extent of 60%. The figure was drawn using Chimera and Protein Data Bank entry 1AG2.

Results

The Pro to Ala mutations do not affect secondary structure but destabilize the native state

All four proline mutant variants (P136A, P157A, P164A and 3PA) showed far-UV CD spectra that are very similar to those of wt moPrP (Figure 2 (a)). This suggests that the Pro to Ala mutations did not affect the secondary structure of the protein in any significant manner. However, urea-induced equilibrium studies indicated that the Pro to Ala mutations led to significant destabilization of the N state of the protein (Figure 2(b)). P157A moPrP showed the highest degree of destabilization among the single proline mutant variants ($\Delta\Delta G_u = 0.6$ kcal mol^{-1}), suggesting that it plays an important role in maintaining the stability of the N state of the protein.

The Pro to Ala mutations result in rapid misfolding of moPrP

The prion protein remains in its native α helical monomeric form at pH 4, but becomes prone to misfolding at low pH because of the protonation of critical residues, His186, Asp177, Asp201.[12,32–33] Misfolding is triggered by the addition of a physiological concentration (150 mM) of salt.[34] The addition of salt is known to disrupt a stabilizing interaction between K193 and E195 in the mono-

meric protein.[33–34] At pH 4, monomeric wt moPrP has an overall charge of +27, and the addition of salt would also screen the electrostatic repulsion between monomers, facilitating monomer association.[34] In the absence of salt, misfolding is very slow (10^{-4} h^{-1}).[34] Hence, the effects of the Pro to Ala mutations on misfolding were studied at pH 4, 37 °C in 150 mM NaCl. The kinetics of misfolding was monitored by measuring the change in the CD signal at 222 nm. The observed misfolding kinetics of all moPrP variants studied here, appeared monophasic and fit well to a single exponential equation (Figures 3(a) and (b)). It should be noted that the misfolding of the prion protein at pH 4 is accompanied by oligomer formation. In the case of wt moPrP, for which the observed misfolding rate constant is relatively slow, the observed rate constant of misfolding matched the observed rate constant of oligomer formation, as probed by size exclusion chromatography (SEC) (data not shown). All four Pro to Ala mutant variants showed much faster misfolding rate constants compared to wt moPrP (Figure 3 and Table S1), suggesting that Pro residues play important role in impeding the misfolding process. P157A and 3PA moPrP showed the most drastic effects.

Far-UV CD spectra showed that the oligomers formed by the different Pro to Ala mutant variants have a β-sheet content similar to that in the oligomers formed by wt moPrP (Figure S1a). The hydrodynamic radii (R_H) of the oligomers formed
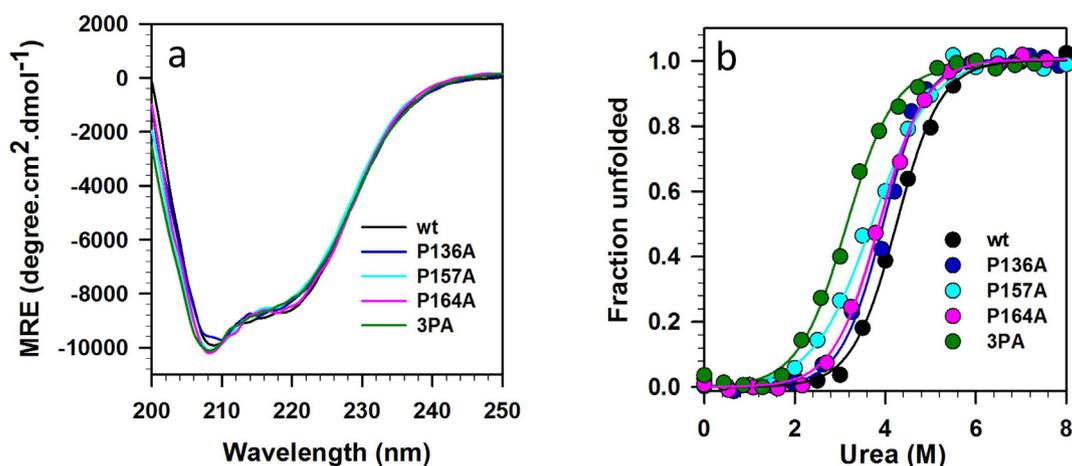
Figure 2. *Effect of Pro to Ala mutations on secondary structure and stability* (a) Far-UV CD spectra of the native monomeric forms of the moPrP variants acquired at pH 4, at 25 °C. (b) Urea-induced equilibrium unfolding transitions of the moPrP variants at pH 4, at 25 °C, as monitored by measurement of the far-UV CD signal at 222 nm. The signal change was normalized to obtain the fraction unfolded. The colors of the lines represent the different moPrP variants as indicated. The solid lines through the data points are fits to an equation describing a two state unfolding transition,[77] and the values obtained for the thermodynamic parameters are listed in Table S1. The urea-induced equilibrium unfolding transitions were measured in three independent experiments, and represerative data from one experiment is shown for each moPrP variant. The standard deviation in measurement determined from the three independent experiments for each moPrP variant is given in Table S1.
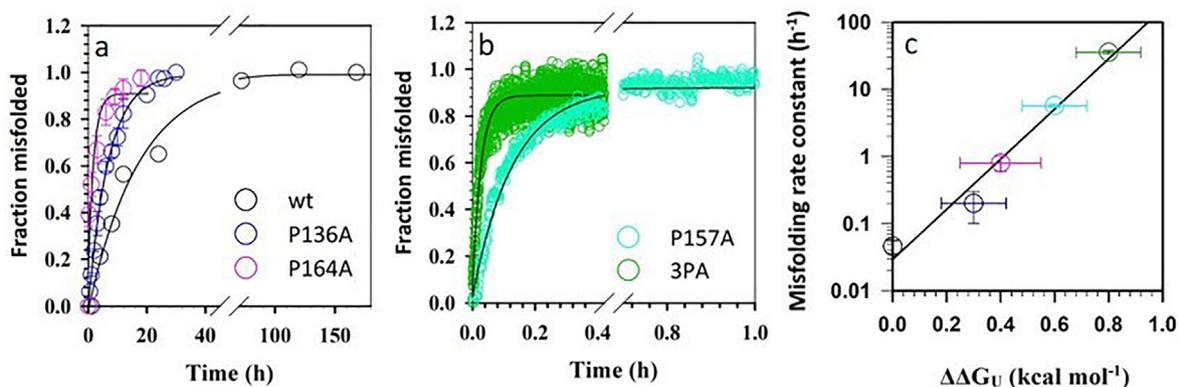
Figure 3. *Dependence of the observed rate constant of misfolding on native state stability* (a) The fraction misfolded form at different times of misfolding of 10 μM wt, P136A and P164A moPrP in 150 mM NaCl at pH 4, at 37 °C is shown. (b) The fraction misfolded form at different times of misfolding of 10 μM P157A and 3PA moPrP in 150 mM NaCl at pH 4,at 37 °C is shown. The fraction misfolded form was calculated from the fractional change in the CD signal at 222 nm. The continuous lines through the data points in both panels a and b are fits to a single exponential equation. (c) Plot of the observed rate constant of misfolding of the different moPrP variants at 10 μM concentration in 150 mM NaCl at pH 4, a t37 °C *versus* the decrease in free energy of unfolding of the variants ($\Delta\Delta G_u$) with respect to that of wt moPrP. The colors for the different moPrP variants are as indicated.The error bars represent the standard deviations from atleast two independent experiments.

by the different moPrP variants were also very similar, in the range of 10 to 15 nm, as measured by DLS (Figure S1b).

Dependence of the rate constants of misfolding on native state stability

To quantify the relationship between the misfolding propensity and thermodynamic destabilization, the observed misfolding rate constants are plotted against the decrease in free energy of unfolding ($\Delta\Delta G_u$) of the mutant variants with respect to wt moPrP (Figure 3(c)). The logarithm of the observed misfolding rate constants showed a linear dependence on $\Delta\Delta G_u$ (Figure 3(c)).

Effect of the Pro to Ala mutations on protein dynamics

To understand why the Pro to Ala mutations accelerate the misfolding of moPrP and

destabilize the native state of the protein, it was important to determine the effect of mutations not only on the structure but also on the dynamics of the protein. HDX-MS studies were carried out to obtain sequence specific information about changes in the structure and dynamics that occur upon mutation.

In the HDX-MS experiment, the unstructured, solvent-exposed amide sites on the main chain will become labeled by deuterium, whereas the amide sites in the structured buried regions of the protein will become deuterated to a lower extent and more slowly. Since the sequence segments which become deuterated will have a higher mass, they are easily identifiable by MS of the corresponding peptide fragments obtained by carrying out peptic fragmentation after HDX. For this purpose, a peptide map was generated using controlled proteolysis by pepsin as described in Materials and Methods (Figure S2). Since the HDX experiments were carried out at pH 4, where
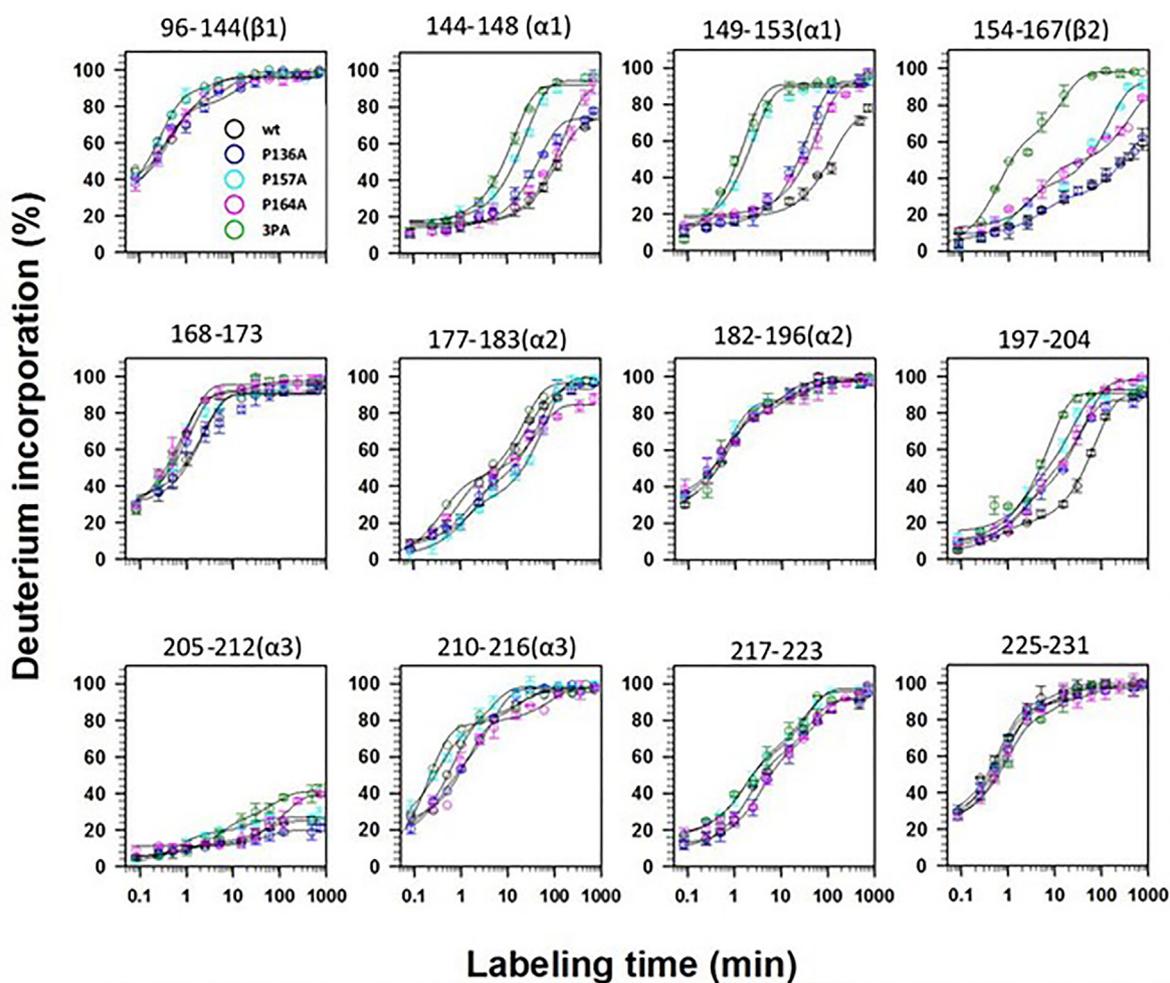
Figure 4. Effect of Pro to Ala mutations on the native state dynamics. The time courses of HDX into different sequence segments of the native monomeric moPrP variants at pH 4, at 25 °C. The differently colored circles represent the different moPrP variants, as indicated. The error bars represent the standard deviations determined from two independent experiments. The solid lines through the data represent fits to either monoexponential or biexponential equations.

the intrinsic rate constant of HDX is slow, structural changes in all the secondary structural elements could be monitored.

Figure 4 shows the kinetics of deuterium incorporation into different sequence segments of the CTD of moPrP, for the different mutant variants. In some sequence segments, the kinetics of deuterium incorporation were very similar for the different variants while some sequence segments showed small variations in the kinetics. However, the sequence segments 144–148, 149–153 (α1), 154–167 (the loop between α1 and β2 and the β2) and 197–204 (the N terminal end of α3) showed pronounced increases in the observed rate constants of deuterium incorporation for the mutant variants in comparison to wt protein (Figure 4). In some cases, the observed exchange rate constant was seen to have increased up to nearly a 1000-fold (Table S2).

For each of the four sequence segments that showed accelerated kinetics of deuterium incorporation in the mutant variants, compared to in wt moPrP, the observed rate constant (k_{obs}) was compared to the rate constant (k_{int}) that would be observed in the corresponding peptide fragment (Table S2) in a random coil state. The quantity $P_f = k_{int}/k_{obs}$ is the protection factor that slowed down HDX into the sequence segment, and $\Delta G_{op} = RT \ln P_f$, is the free energy of opening (unfolding) of the local structure of the sequence segment to HDX. ΔG_{op} is therefore the free energy of stability of the local structure. It is the difference in the free energy of the form with the local structure unfolded and the free energy of the native state with the local structure folded. Tables S3 and S4 show the effects of the Pro to Ala mutations on the P_f and ΔG_{op} values, respectively, for the sequence segments 144–148, 149–153, 154–167 and 197–204. They show that the mutations decrease the values of ΔG_{op}, thereby increasing the sampling of the partially unfolded form (PUF) in which the four segments are unfolded. It should be noted that ΔG_{op}, is the difference in free energies of the N state and the partially unfolded form (PUF) into which HDX occurs.

Dependence of the rate constants of misfolding on local segment stability

The Pro to Ala mutations maximally perturb the rate constants of HDX into four sequence segments: 144–148 and 149–153 (α1), 154–167 (the loop between α1 and β2, and β2) and 197–204 (the N terminal end of α3) (Figure 4). Hence the local segmental stabilities of these four sequence segments are decreased (Table S4). Figure S3 shows that when the local stabilities of these four most perturbed sequence segments are decreased incrementally by the Pro to Ala mutations, the logarithm of the observed rate

constant of misfolding increases linearly with the decrease in the ΔG_{op} values. For each of the moPrP variants, the ΔG_{op} values for the four sequence segments were averaged to yield ΔG_{op}^{av}. Figure 5 shows that when the logarithm of the observed rate constants of misfolding is plotted against the ΔG_{op}^{av} values, the linear dependence was still observed. This indicated that each mutation similarly perturbs the stability of the four sequence segments in the PUF into which HDX occurs.

Local and not the global stability dictates the misfolding rate constant

Several disease-linked mutant variants are known to misfold rapidly under similar conditions, and for them too, a linear dependence of the logarithm of the observed misfolding rate constants on the decrease in the free energy of unfolding ($\Delta \Delta Gu$) had been observed (Figure 6 (a)). The misfolding rate constants of the proline mutant variants have, however, a much stronger dependence on the decrease $\Delta \Delta Gu$, than do the disease-linked mutant variants (Figure 6(a)).

Interestingly, the logarithm of the observed rate constant of misfolding of the disease-linked mutant variants also correlates linearly and well
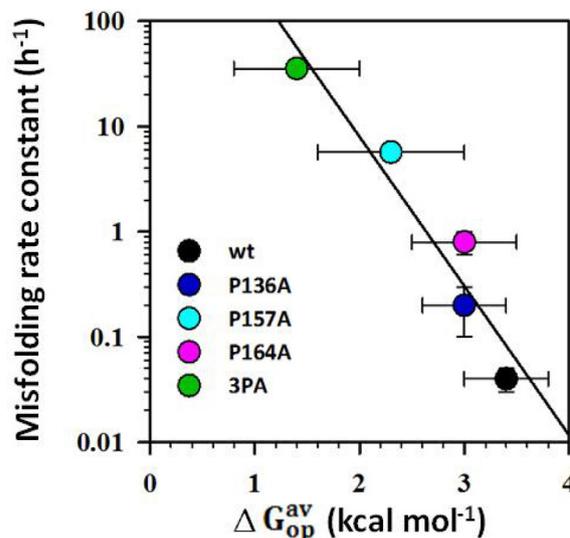
Figure 5. Dependence of the observed rate constant of misfolding of the different moPrP variants on ΔG_{op}^{av}. Misfolding of the different moPrP variants at 10 μM concentration was carried out in 150 mM NaCl at pH 4, at 37 °C. ΔG_{op}^{av} is the average of the ΔG_{op} values obtained for the four sequence segments 144–148, 149–153, 154–167 and 197–204 for each of the moPrP variants (Table S4). The colors of the symbols for the different moPrP variants are as indicated. The error bars are the standard deviations determined from two independent experiments.
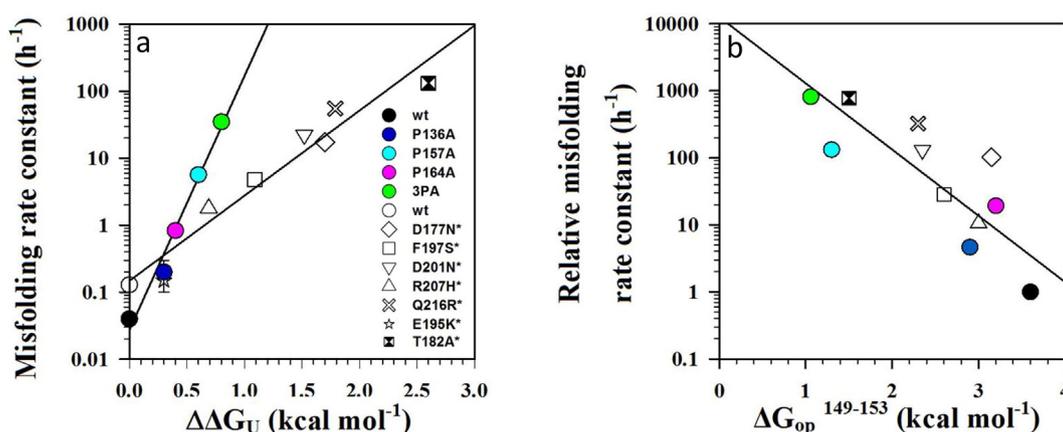
Figure 6. Proline mutations versus disease-linked mutations. (a) Plot of the observed rate constant of misfolding of different moPrP variants in 150 mM NaCl at pH 4, at 37 °C *versus* the decrease in global stability of the mutant variants ($\Delta\Delta G_U$) with respect to wt moPrP. The misfolding of the Pro to Ala mutant variants was studied at 10 μM concentration (filled symbols), and that of the disease-linked mutant variants had been studied at 100 μM concentration (empty symbols). The continuous lines through the data points are linear fits to the data points. (b) Plot of the relative observed rate constant of misfolding of the different moPrP variants with respect to that of wt moPrP in150 mM NaCl at pH 4, at 37 °C *versus* ΔG_{op} values for the 149–153 segment calculated from HDX labeling rates. The error bars represent the standard deviations from two independent experiments. Asterisks denote disease-linked mutant variants, for which the data was taken from references.[34,49,80]

with the local stability of sequence segment 149–153 (Figure 6(b)). The change in the local stability of sequence segment 149–153 is the more appropriate probe of the misfolding propensity, as the observed rate constant of misfolding has the same dependence on ΔG_{op} for both the Pro to Ala mutant variants and the disease-linked mutant variants.

The core of the oligomers formed by 3PA moPrP appears to be more extended than that formed by wt moPrP

Figure 7 compares the protection factors, P_f, (calculated as described in the Materials and Methods) determined from the extents of deuterium incorporation into the different sequence segments of oligomers formed by the moPrP variants, when the oligomers were subjected to 3000 s HDX labeling pulses at pH 4, 25 °C. The core of oligomers was defined as the longest continuous sequence stretch that had $P_f > 50$. By this criterion, the core comprised sequence stretch 168–212 in wt moPrP oligomers, and sequence stretch 144–231 in 3PA moPrP oligomers. As can be seen in Figure 7, the core became extended in the oligomers formed by 3PA moPrP because the Pro to Ala mutations increased the local stabilities of sequence segments 159–166 and 210–216.

Effect of the Pro to Ala mutations on the energy barrier of misfolding

Figure 8(a) shows the CD-monitored kinetics of misfolding of wt moPrP at different temperatures. At each temperature the kinetic curves fit well to a single-exponential equation. The observed misfolding rate constants increased with an increase in temperature (Figure 8(a)). Similar behavior was also observed for the 3PA mutant variant (Figure 8(b)). The logarithm of the observed rate constant of misfolding was plotted against 1/T; the Arrhenius plot (Figure 8(c)) indicated that the energy barrier misfolding slowing down misfolding was 19 kcal mol^{-1} for wt moPrP. Similarly large energy barriers have been observed to slow down the misfolding of other proteins, including α-synuclein,[35] amyloid-β,[36] human insulin[37] and barstar.[38] The energy barrier observed for the misfolding of wt moPrP was also very similar to that observed for the proline isomer-ization reaction.[19] All three Pro residues in the struc-tured CTD are known to be in the *trans* conformation in the moPrP monomer,[3] and it should be noted that the NMR spectra at pH 4 of the monomer[34,44] were found to be indistinguishable from those reported in the earlier study.[3] It is possi-ble therefore that the misfolding of moPrP is accom-panied by *trans* to *cis* isomerization of one or more of these Pro residues. When the three Pro residues were however mutated to Ala, the energy barrier of misfolding reduced only to 14 kcal mol^{-1} in 3PA moPrP (Figure 8(c)). This observation suggests that it is unlikely that any of the Pro residues switch from a *trans* to *cis* conformation during misfolding.
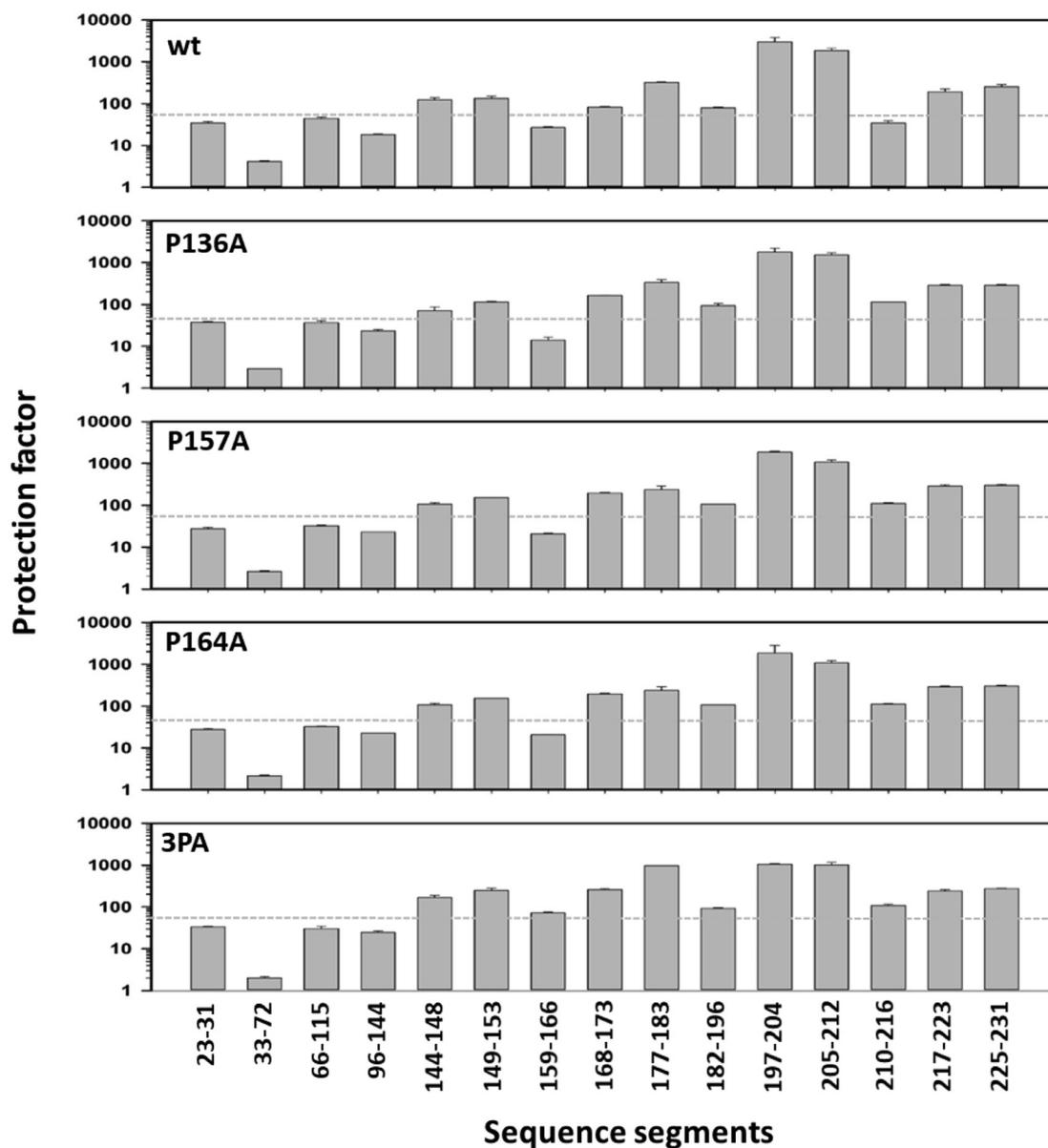
Figure 7. HDX-MS characterization of the oligomers of the different moPrP variants formed by 10 μM protein in 150 mM NaCl at pH 4, at 37 °C. The histogram bars show the protection factors (P_f) at 25 °C of the different sequence segments calculated as described in the Materials and Methods. The dashed line indicates $P_f = 50$, which was the threshold used to define the core. The error bars represent the standard deviations from two independent experiments.

Discussion

The effect of mutating a Pro residue on the stability of a protein is difficult to predict. The effect on the N state may be stabilizing or destabilizing, or neither, depending on the location of the Pro residue in the protein structure, on the nature of the neighboring residues, and on the nature of the residue replacing it.[39] A Pro residue cannot participate in hydrogen bonding, whereas any residue that replaces it would be able to. If the substituting residue does engage in hydrogen bonding, the mutation would be expected to have a stabilizing effect; if it does not, the unsatisfied hydrogen bonding potential would be expected to have a strongly destabilizing effect on the N state.[40] The effect on the U state is expected to be stabilizing, as the conformational entropy of U would be reduced because the pyrrolidine ring of a Pro residue restricts its main chain dihedral angle, ϕ, to $-63 \pm 15°$[41] and also restricts the conformation of the preceding residue.[41] Hence, replacing a Pro
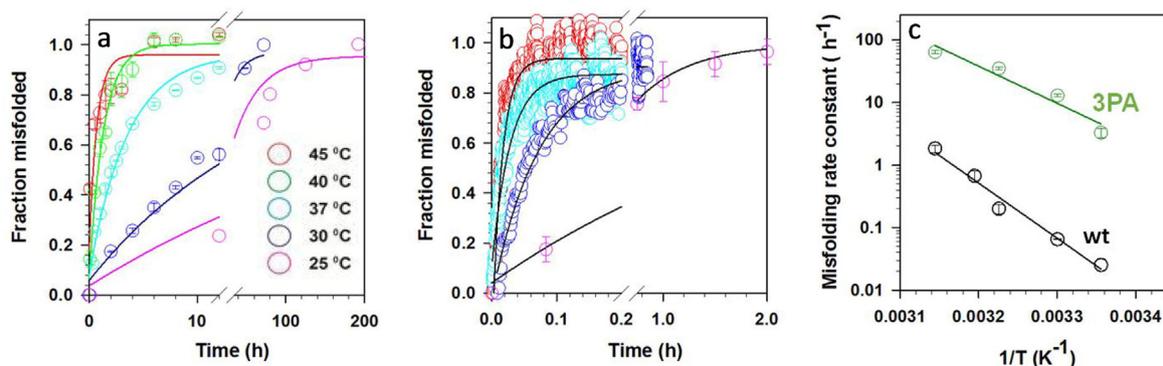
Figure 8. Temperature dependence of the kinetics of misfolding of 100 μM wt (a) and 10 μM 3PA (b) moPrP in 150 mM NaCl, at pH 4. Representative plots are shown for each temperature. The continuous lines through the data points are fits to a single-exponential equation. (c) Observed rate constants of misfolding are plotted against 1/T. Linear fits through the data points yield activation energies of 19 kcal/mol and 14 kcal/mol for wt and 3PA moPrP, respectively. The error bars in the panels represent the standard deviations from two independent experiments.

residue with any other residue would entropically destabilize the protein, that is reduce ΔG_U, the free energy difference between the U and N states.

The Pro to Ala mutations result in the destabilization of moPrP

The observation that the Pro to Ala mutations reduce ΔG_U can be rationalized to be the result of entropic stabilization of the U state, as discussed above. It is instructive, however, to examine whether ΔG_U could have also been reduced because the N state is enthalpically destabilized. The side-chains of Pro136 and Pro157 are buried in the hydrophobic core, and form van der Waal contacts with multiple residues present in α2 and α3 (Figure 1(b)). Substitution of these Pro residues with the smaller Ala residues could perturb the packing in the core and thereby destabilize the protein.[42] It would be unlikely that the substituting Ala residues remain buried, as hydrogen bonding partners for their main chain amides would not be available. Since burial of a main chain amide without a hydrogen binding partner would significantly destabilize the protein,[40] it is likely that the substituting Ala residues in the N state have their main chain amides not buried, but solvent-exposed, thereby resulting in a perturbation of core packing and consequent destabilization.

The buried side-chains of Pro157 and Pro164 also appear to confer rigidity to the sequence stretch 144–167, anchoring it in the hydrophobic core. This anchoring may be important for the positioning of Glu145 and Arg155 in the structure such that they can participate in the salt-bridges that comprise the highly conserved electrostatic network that stabilizes the interaction of the α1-β2 sub-domain with α2-α3 sub-domain (Figure 9). Substitution by Ala might perturb the positioning of Glu145 and Arg155, and thereby the critical

electrostatic network, resulting in the destabilization of the N state. It should be noted that the α1-β2 sub-domain has very few hydrophobic contacts with the α2-α3 sub-domain, as a result of the unusually high hydrophilicity of α1,[43] which has been reported to be one of the most soluble helices found in any protein.[43]

The Pro to Ala mutations stabilize a HX-competent partially unfolded form

Previous HDX-MS and HDX NMR measurements had identified two sparsely populated, partially unfolded forms (PUFs) into which HDX occurs in regions that have become unstructured and fully solvent-exposed.[44] β1, the loop between β1 and α1, as well as the N-terminal end of α3, were found to be such regions in PUF1. α1, the loop between α1 and β2, as well as β2 were found to be additional such regions in PUF2. It appeared that the α1-β2 sub-domain is separated from the α2-α3 sub-domain in PUF2. It was shown previously that the free energy of opening, ΔG_{op}, of structure in N to form PUF2, was 3.1 ± 0.5 kcal mol^{-1} in the case of the wt moPrP.[44] The previous studies could not, however, determine, whether PUF1 and PUF2 were unique conformations, or were ensembles of sub-populations that formed in parallel, and which differed in their conformations.

The current study shows that the Pro to Ala mutations predominantly affect four sequence segments encompassing the α1-β2 sub-domain and the N-terminal end of α3. These regions show increased dynamics and significant changes in ΔG_{op} (Figures 4 and 5; Table S4). The same regions were shown to be unstructured and solvent accessible in PUF2.[44] The Pro to Ala mutations do not, however, have any significant effect on other regions that were also shown to be unstructured in PUF2. In particular, the kinetics of deu-
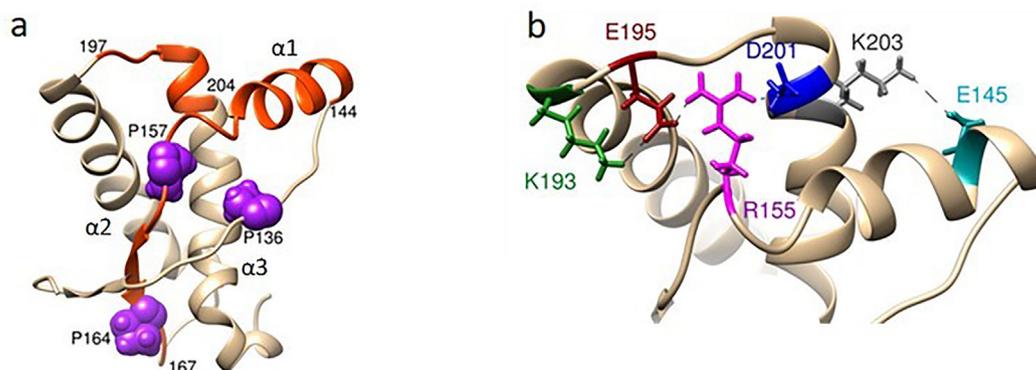
Figure 9. Effect of the Pro to Ala mutations on protein dynamics (a) Sequence segments 144–148, 149–153, 154–167 and 197–204 showing increased structural dynamics in native monomeric moPrP, upon Pro to Ala mutations are shown in red. The structural dynamics were determined from the kinetics of HDX into the different sequence segments at pH 4, at 25 °C (Figure 4). The proline residues which were mutated in this study are colored purple. (b) An expanded view of the well-conserved non-local salt bridges formed between Arg155 (purple) and Glu195 (brown), Arg155 and Asp201 (blue), Lys193 (green) and Glu195, Lys203 (grey) and Glu145 (cyan). The distances between the side-chains are shown in Å. The figure was drawn using Chimera and the Protein Data Bank entry 1AG2.

terium incorporation into these other regions (Figure 4 were affected minimally; consequently, values of ΔG_{op} for these regions were also affected only marginally (Table S5). This result suggests that PUF2 is an ensemble consisting of sub-populations of distinct conformations, each of which has become unstructured in different regions, but which are nevertheless very similar in energy. The sub-population of PUF2 which has become unstructured at the $\alpha1$-$\beta2$ sub-domain and the N-terminal end of $\alpha3$, and which is affected by the Pro to Ala mutations, is referred to as PUF2*. The value of ΔG_{op}^{av} for PUF2* in the case of wt protein, determined by averaging over the ΔG_{op} values of the four sequence segments encompassing the $\alpha1$-$\beta2$ sub-domain and the N-terminal end of $\alpha3$, is found to be 3.4 ± 0.4 kcal mol^{-1}, qualifying PUF2* to be a sub-population of the PUF2 ensemble.

The Pro to Ala mutations result in a significant reduction in the value ΔG_{op}^{av} for PUF2* (Figure 5 and Table S4). This reduction cannot be accounted for by the relatively small destabilization of the N state (see above). It can be accounted for only if the mutations significantly increase the stability of PUF2* with respect to the U state. At the present time, it is possible only to speculate how such stabilization might occur in PUF2*. It would appear that the substituting Ala residues have their main chain amides hydrogen bonded in PUF2*, either intra-molecularly with another part of the protein, or inter-molecularly with water.[45–46] This non-native hydrogen bonding would not be present in PUF2* formed by the wt protein, because the Pro residues do not have amides that can hydrogen bond. If the hydrogen bonding, and the stabilization afforded by it is absent in the U state, then PUF2* would be stabilized with respect to the U state.

PUF2* appears to be the precursor state from which misfolding commences

In an earlier study, an intermediate, which was populated at equilibrium during the urea-induced unfolding of a mutant variant of the isolated CTD of moPrP, was identified as the precursor conformation from which misfolding initiates.[47] This identification was on the basis of the observed rate constant of misfolding being linearly proportional to the population (concentration) of the intermediate. On the basis of similarities in stability and in gross structure, the equilibrium intermediate was inferred to be the same as PUF2.[44] However, in the case of full length moPrP, such an intermediate has not been identified, and it was not possible to show that PUF2 is the direct precursor to misfolding. In this study, it has been shown that the logarithm of the apparent misfolding rate constant is linearly proportional to ΔG_{op}^{av} for PUF2* (Figure 5), the free energy difference between the N state and PUF2*. In other words, the observed misfolding rate constant is directly proportional to the equilibrium population (concentration) of PUF2*. This observation strongly indicates that PUF2*, a sub-population of the PUF2 ensemble, is the direct precursor conformation from which misfolding commences.

The defining structural difference between PUF2* and N appears to be that in the former, the $\alpha1$-$\beta2$ sub-domain is separated from the $\alpha2$-$\alpha3$ sub-domain.[44] It is known that separation of the $\alpha1$-$\beta2$ sub-domain from the $\alpha2$-$\alpha3$ sub-domain is a critical event in the initiation of misfolding.[48–49] Several disease-linked mutations have been shown to facilitate sub-domain separation, increasing the population of PUF2*, and they thereby speed-up misfolding.[49] It is interesting to note that two anti-prion drugs that were found to be effective against

prion conversion[50–51] bind to the N state between α1 and the α2-α3 loop, thereby presumably stabilizing the N state with respect to PUF2*. It is therefore not surprising that the effect of the Pro to Ala mutations is largest in the two non-contiguous sequence stretches, 144-–167 spanning the α1-β2 subdomain, and 197–204 at the N-terminal end of α3 (Figure 9(a)). These two structural regions are both unfolded and solvent-accessible in PUF2*. In the N state, these two regions are close to each other, and are stabilized by an electrostatic interaction network comprised of salt bridges formed by residues in the two sequence stretches (Figure 9(b)).

Indeed, a distributed network of electrostatic interactions play a critical role in stabilizing the N state of the prion protein and guarding against misfolding. It is only when His186 becomes protonated at pH 4, leading to the perturbation of the electrostatic network, that the α1-β2 subdomain becomes destabilized.[32,52] Consequently, PUF2* becomes populated, albeit to a very sparse extent for wt moPrP, possibly driven by structural fluctuations between α1 and α2.[53] The electrostatic network is destabilized by the Pro to Ala mutations, which appear to perturb the structure in such a way so as to affect the positioning of charged residues that participate in critical salt-bridges that comprise the stabilizing electrostatic network (see above). The destabilization of the N state results in the reduction in ΔG_{op}^{av}, increases the population of PUF2* relative to N. A much larger reduction in ΔG_{op}^{av} occurs, however, as a consequence of the stabilization of PUF2* relative to the U state (see above).

Misfolding does not commence from the fully unfolded U state

The observation that the logarithm of the observed misfolding rate constant is also linearly proportional to $\Delta\Delta G_U$, for the Pro to Ala mutant variants (Figure 6(a)), could be interpreted to mean that the U state can also act as the precursor conformation from which misfolding initiates. However, if such an interpretation were true, then an identical dependence would be observed for all mutant variants that have been studied previously. The observation that while the logarithm of the apparent misfolding rate constant determined for disease-linked mutant variants is also linearly proportional to $\Delta\Delta G_U$, its dependence is nevertheless very different (Figure 6(a)), suggests that misfolding does not commence from the U state. It should also be noted that some of the disease-linked mutations destabilize the protein significantly more than do the Pro to Ala mutations, but yet cause less acceleration of the misfolding reaction. On the other hand, the observation that the logarithm of the observed misfolding rate constant for the Pro to Ala mutant variants and the disease-linked mutant variants (Figure 6(b)) studied earlier, has the same linear

dependence on ΔG_{op}, (determined as the free energy of opening of structure to HDX in the C-terminal part of α1),[49] is expected when PUF2* is the precursor conformation from which misfolding commences. It is clear that the probability of misfolding depends on ΔG_{op}, which dictates the extent to which PUF2* is populated, at equilibrium, and not on $\Delta\Delta G_U$, which dictates the extent to which U is populated in equilibrium with N.

The Pro to Ala mutations lower the energy barrier for oligomerization

While one mechanism by which the Pro to Ala mutations lead to faster misfolding is through their stabilization of the aggregation-competent PUF (PUF2*) from which misfolding commences, another mechanism could be that they lower the free energy barrier that slows down misfolding. This free energy barrier would have an entropic component arising primarily from molecules having to assemble together during the oligomerization reaction that occurs concomitantly with misfolding. There would also be an enthalpic component arising from having to break the interactions that stabilize the structure, so that conformational conversion can occur. In particular, α2 and α3, which convert into β-sheet, are stabilized by electrostatic interactions between them.[52] The protonation of His186 perturbs these interactions but not sufficiently for the two helices to start converting to β-strand, even though the gatekeeper helix α1 has become disordered in PUF2*. For misfolding to commence, the electrostatic interactions have first to be further destabilized. This happens upon the addition of 150 mM NaCl, which screens electrostatic charges and destabilizes N.[53] The immediate effect is the disruption of the electrostatic network that includes the salt bridge between Lys193 and Glu195 and electrostatic interactions involving Glu210-Gln211 and Lys219 in α3,[34] as well as many other polar residues.[54] Consequently, the stabilizing tertiary electrostatic contacts between α2 and α3 break, and the helices are destabilized. Once helical structure is lost, conformational conversion to β-sheet is initiated.[55] Conformational conversion is favored, as the sequence segments spanning α2 and α3 have low helical and high β-sheet propensities in the first place.[56] It is possible that the Pro to Ala mutations lower the activation energy for misfolding (Figure 8(c)) and accelerate misfolding by perturbing the structure of PUF2* in such a manner that the stabilization afforded by electrostatic interactions is decreased.

It should be noted that the activation energy for misfolding remains high even after all three Pro residues have been substituted with Ala (Figure 8(c)). Since the activation energy is similar to that seen for peptidyl proline isomerization reactions accompanying protein folding and unfolding reactions,[57] it is possible that the misfolding of

moPrP is accompanied by peptidyl proline isomerization. The three Pro residues in the CTD are in the *trans* conformation in native moPrP.[3] Since it is not known whether they are in the *cis* or *trans* conformation in the misfolded oligomers, it cannot be discerned whether misfolding is accompanied by *trans* to *cis* prolyl isomerization. The observation that the activation energy remains high even after all three Pro residues have been substituted with Ala (Figure 8(c)), makes it unlikely that *trans* to *cis* peptidyl proline isomerization plays a role in the misfolding of moPrP, as it does for other proteinss.[35–38] Certainly, *trans* to *cis* peptidyl proline isomerization is too slow ($<0.01 \ s^{-1}$) for it to occur during the fast unfolding reaction leading to the formation of the PUF2* from the N state.

PUF2* appears not to be on the direct unfolding pathway from the N to U state

Earlier studies had suggested that intermediates that are obligatory for folding[58–59] or unfolding[60] may initiate misfolding of the prion protein. Thus, an important question is whether PUF2* is on the direct pathway of unfolding from N to U, or whether it is a dead end, off-pathway intermediate accessible from N. Earlier studies had been unable to determine whether PUF1 and PUF2 were on– or off-pathway.[44] In the case of cytochrome c[61] and apomyoglobin,[62] the PUFs identified in HX NMR studies were designated as being on-pathway, as mutations that stabilized the N state also stabilized the PUFs.[63] In the case of apoflavodoxin, the identified PUFs appeared not to be present on the direct pathway of unfolding.[64] In the case of moPrP, the observation that the Pro to Ala mutations destabilize the N state (Figure 2) but stabilize PUF2* (Table S4) relative to U, suggests that PUF2* is an off-pathway dead-end intermediate sampled from the N state. The conjecture that PUF2* is stabilized by non-native hydrogen-bonding enabled by the Pro to Ala mutations is consistent with it being off-pathway. In this context, it should be noted that a molecular dynamics simulation study of the prion protein identified multiple partially unfolded states stabilized by non-native hydrogen bonds.[65] Single molecule force microscopy studies have also shown that the unfolded state of the prion protein can misfold along multiple pathways.[66]

The Pro residues appear to restrict the boundary of the structural core of the oligomers

Misfolded moPrP is found in different types of aggregates: oligomers[49] whose formation has been studied here, worm-like fibrils[67] and straight fibrils.[68] All three aggregates possess β-sheet structure,[69] some of which would confer protection against hydrogen exchange.[49,69–70] Previous hydrogen

exchange studies had shown that the core of the aggregate, defined as the region most protected against HX, extended from residues 168 to 223 in worm-like fibrils,[71] as well as in straight fibrils formed *in vitro*.[70] The previous studies had shown that the core in oligomers is weakly protected against HX, compared to in worm-like fibrils and fibrils. Hence, in the current study, the core was defined as the sequence stretch comprising sequence segments that have a protection factor against HX of greater than 50 (see Results). By this criterion, the current study suggests that the protective core in the oligomers extends from residue 168 to residue 212 in wt moPrP, but gets expanded to extend from residues 144 to 231 in 3PA moPrP (Figure 7).

Pro residues are rarely found in β-sheets in proteins, but they are often found flanking β-strands because of the local twist they impart to the peptide backbone.[16] Hence, it would appear that in 3PA moPrP, the absence of Pro residues in the sequence stretch 144–168, has allowed expansion of the protective β-sheet core. But it is more likely that β sheet structure spans the same length of sequence stretch in both the wt moPrP and 3PA moPrP oligomers, as both oligomers show the same amount of β-sheet structure in the far-UV CD spectra (Figure S1a). The presence of Pro residues in the wt moPrP oligomers might have made the structure present in sequence stretch 144–168 dynamic and hence, not protective against HX.

Recent cryo-EM studies have revealed the diversity inherent in prion aggregate structure. The β-sheet structure extends from residues169 to 224 for straight fibrils formed *in vitro* by the human prion protein,[68] and from residues 94 to 225, residues 94 to 226 and residues 79 to 140 in the case of prion aggregates derived from diseased mouse,[72] hamster[73] and human[74] brain, respectively. The β-sheet core comprises of 15, 11 and 8 β-strands for prion aggregates derived from diseased mouse,[72] hamster[73] and human[74] brain. Pro136 is present at one end of β5 of mouse,[72] β4 of hamster,[73] and β8 of human[74] brain-derived aggregates, whereas Pro157 and Pro164 flank the ends of β8 of mouse[72] and β5 of hamster[73] prion aggregates. It appears that the Pro residues might be playing a role in restricting the lengths of the β-strands in these fibrillar aggregates, as they could be doing in the case of the oligomers. It is clear from a previous HX study of fibrillar prion aggregate isolated from diseased brain,[75] that all the β-strands of the fibril core seen in the cryo-EM studies, are not protective against hydrogen exchange. Unfortunately, detailed structural characterization of prion oligomers is not yet possible, and it remains to be seen how similar the β-sheet structure in prion oligomers is to that in prion aggregate derived from diseased brain.

Materials and Methods

Buffers and reagents

All the reagents used in this study were of high purity grade, and were purchased from Sigma. Ultra-pure GdnHCl was obtained from United States Biochemicals, and was of the highest purity grade.

Site-directed mutagenesis

The mutant variants of full-length moPrP were generated using the Quickgene site-directed mutagenesis kit (Stratagene). Primers containing 1 or 2 nucleotide changes were obtained from Sigma. Four mutant variants were prepared: P136A, P157A, P164A and 3PA (P136A P157A P164A). The mutations were confirmed by DNA sequencing.

Protein expression and purification

Wt moPrP and the mutant variants were expressed in *Eschercihia coli* BL21 (DE3) codon plus cells (Stratagene) transformed with a pET17b plasmid containing the full-length sequence (23–231) of the moPrP gene. All the moPrP variants were purified as described previously.[67,76] The purity of each moPrP variant preparation was confirmed by mass spectrometry using a Synapt G2 HD mass spectrometer (Waters Corporation). Each moPrP variant had its expected mass.

Far-UV CD measurements

Far-UV CD spectra were collected using a Jasco J-815 spectropolarimeter. They were recorded using a protein concentration of 10 μM, in a 1 mm cuvette, using a scan speed of 50 nm/min, a digital integration time of 2 s, and a bandwidth of 1 nm. The wavelength was scanned from 200 to 250 nm, and a total of 15 spectra were averaged. Far-UV CD spectra under native conditions were acquired in 10 mM sodium acetate, pH 4, at 25 °C.

Urea induced equilibrium unfolding studies

Urea-induced equilibrium unfolding was carried out at pH 4, at 25 °C in 10 mM sodium acetate buffer. For these studies, 10 μM protein was incubated in different denaturant concentrations for 1 h at 25 °C before the far-UV CD signal at 222 nm was monitored. The data were fit to a two-state (N U) equilibrium unfolding model, and the thermodynamic parameters were obtained.[77]

Misfolding studies at pH 4

Misfolding was monitored by the measurement of ellipticityat 222 nm (θ222). The protein in 10 mM sodium acetate buffer (pH 4) was diluted twofold with 2x misfolding buffer (10 mM sodium acetate buffer containing 300 mM NaCl, pH 4) so that it was finally in 1x misfolding buffer (10 mM sodium acetate buffer containing 150 mM NaCl, pH 4). The protein concentration used for most of the experiments was 10 μM, except for the temperature-dependent kinetic studies of the misfolding of wt moPrP where it was 100 μM. The samples were incubated at 37 °C. At different times of misfolding, aliquots of the protein sample were withdrawn for analysis by measurement of θ222.

To study the misfolding kinetics of P157A and 3PA moPrP, a 1 mm quartz cuvette maintained at 37 °C was used. The kinetics were monitored by measurement of θ222. Both the protein sample and 2x misfolding buffer were incubated at 37 °C before starting the reaction. The reaction was started by mixing 20 μM moPrP solution with an equal volume of 2x misfolding buffer. The time from the mixing of the protein solution with 2x misfolding buffer to the first reading was 20 s. The fraction misfolded form (f_m) was calculated using the equation:

$$f_m = \frac{\theta_t - -\theta_0}{\theta_\infty - -\theta_0}$$

θ_t and θ_0 are the ellipticities at 222 nm at time t and at time t = 0 of the misfolding process, respectively and θ_∞ is the ellipticity at 222 nm of the completely misfolded form (Figure S1a).

Dynamic light scattering (DLS) measurements

DLS measurements were carried out on a DynaPro-99 instrument (Wyatt Technology Corp.). All the buffers were filtered through 0.02 μm filters (Whatman). The scattering intensity at 90°, and its autocorrelation, were acquired simultaneously using a laser at 829.4 nm to illuminate the sample. 50 acquisitions were collected in each experiment. The acquisition time was set at 5 s, the signal-to-noise ratio threshold at 2.5, the temperature at 25 °C, and the sensitivity at 80%. Data with uneven autocorrelation functions were excluded. The data were resolved into a Gaussian distribution using DynaLS (Protein Solutions Ltd.).

Peptide mapping

To generate a peptide map of moPrP, the protein, in 10 mM sodium acetate buffer at pH 4, was subjected to online pepsin digestion in 0.05% formic acid using an immobilized pepsin cartridge (Applied Biosystems) at a flow rate of 40 μl/min on a nanoAcquity UPLC (Waters). Fragments were collected in a peptide trap column (C18 reversed-phase chromatography column), washed to remove salt, and eluted on an analytical C18 reversed-phase chromatography column using a gradient of 3 to 40% acetonitrile (0.1% formic acid) at a flow rate of 40 μl/min. The peptides were analyzed using the coupled Synapt G2 HD

mass spectrometer (Waters Corporation). Peptide sequencing was done using the MS/tandem MS method, followed by analysis with the Protein Lynx Global Server software (Waters Corporation), and manual inspection.

HDX-MS measurements

(a) For native monomers

To initiate deuterium labeling, 100 μM protein was diluted into a labeling buffer [10 mM sodium acetate buffer in D2O (pH 4), corrected for the isotope effect], so that HDX occurred in 95% D2O and at 25 °C. At different times, a 50 μl aliquot was mixed with 50 μl ice-cold 20 mM glycine-HCl buffer (pH 2.5) to quench the labeling. These samples were then immediately injected into the HDX module (Waters Corporation) coupled to a nanoAcquity UPLC for online pepsin digestion using an immobilized pepsin cartridge (AppliedBiosystems) at a flow rate of 40 μl/ min of water (0.05% formic acid). The peptides eluting from the pepsin column were collected using the trap column, washed to remove salt, and eluted on an analytical C18 reversed-phase chromatography column with a gradient of 3 to 40% acetonitrile (0.1% formic acid) at a flow rate of 40 μl/ min (total elution time of 12 min). All columns were kept at 4 °C in the cold chamber of the HDX module to minimize back exchange. The mass spectrometer parameters were as follows: source temperature, 35 °C; desolvation temperature, 100 °C; capillary voltage, 3.0 kV.

Peptide masses were calculated from the centroid of the isotopic envelope using the MassLynx software, and the shift in mass of labeled peptides relative to unlabeled peptides was used to determine the extent of deuterium incorporation at each time point of HDX. Since the sample was in 95% D2O during labeling, and exposed to H2O during quenching, control experiments were carried out to correct for back exchange. To this end, the protein was incubated in 10 mM sodium acetate, pH 4 (in 95% D2O), and deuterated by unfolding at 65 °C for 15 min, followed by refolding on ice. Refolded moPrP was shown to behave like native moPrP, using CD spectroscopy and thermal equilibrium unfolding studies. The fully (95%) deuterated moPrP samples were processed in the same way as the labeling reaction samples. The extent of deuterium incorporation in each peptide fragment, % D was calculated using the equation:

$$\%D = \frac{m_{(t)} - m_{(0)}}{m_{(95)} - m_{(0)}} \times 100$$

where $m_{(t)}$ is the measured centroid mass at time t, $m_{(0)}$ is the measured mass of the undeuterated reference peptide, and $m_{(95)}$ is the measured mass of a fully deuterated reference peptide (in 95% D2O). For calculation of the free energy of opening (ΔG_{op}), the observed HDX rate constant (k_{obs}) for a peptide fragment, which was obtained by fitting the kinetic curve of HDX to either a single exponential or a two-exponential equation depending upon the goodness of fit, was compared to the intrinsic rate constant (k_{int}) expected for the same peptide fragment in a random coil state.[78–79] The HDX protection factor ($P_f = k_{int}/k_{obs}$) for the peptide was used to calculate ΔG_{op} for the sequence segment corresponding to the peptide ($\Delta G_{op} = -RT \ln P_f$). ΔG_{op} is the free energy of unfolding of the N state to the partially unfolded form in which the sequence segment is locally unfolded.

(b) For oligomers

Oligomers prepared by incubation of 10 μM protein in 150 mM NaCl, 10 mM sodium acetate, pH 4, at 37 °C, (for a time corresponding to 5 time constants of the misfolding reaction) were concentrated to 500 μM by centrifugation (10000g) using a 10 kDa molecular mass cut-off filter. Deuterium labeling was initiated by diluting 5 μl of oligomers with 45 μl of 1x misfolding buffer prepared in D2O (pH 4, corrected for isotope effect). Labeling was allowed to occur at 25 °C for 300 s or 3000 s. After that, 450 μl of ice-cold quenching buffer (8 M guanidine hydrochloride (GdnHCl), 100 mM glycine, pH 2.5) were added to stop the labeling reaction and dissolve the oligomers. After 2 min of incubation on ice, the samples were desalted using a Sephadex G-25 HiTrap desalting column pre-equilibrated with water at pH 2.5, in conjunction with an AKTA Basic HPLC. The desalted samples were injected into the HDX module (Waters) coupled with the nanoAcquity UPLC, for online pepsin digestion (see above).

Exchange experiments of native moPrP were also carried out in a similar way after 10-fold dilution of the protein in 10 mM sodium acetate into D2O containing buffer (pH 4, corrected for the isotope effect). To mimic the digestion conditions used for oligomers, 50 μl of deuterium labeled sample of native moPrP were mixed with 450 μl of ice-cold quenching buffer containing 8 M guanidine hydrochloride (GdnHCl), 100 mM glycine, pH 2.5, and processed in the same way as were the oligomers.

Peptide masses were calculated from the centroids of the isotopic envelopes, using the MassLynx software, and the shift in the mass of labeled peptides relative to that of the unlabeled peptides was used to determine the extent of deuterium labeling after 3000 s of HDX. As the sample was in 90% D2O during labeling and exposed to H2O after dissolution in GdnHCl, control experiments were carried out to correct for back exchange. To this end, the protein was incubated in 10 mM sodium acetate, pH 4 (in 90% D2O), and deuterated by unfolding at 65 °C for 15 min, followed by refolding on ice. The fully deuterated (90%) moPrP samples were processed in the same way as the labeling

reaction samples. The extent of deuterium incorporation in each peptide, % D was calculated using the equation:

$$\%D = \frac{m_{(t)} - m_{(0)}}{m_{(90)} - m_{(0)}} \times 100$$

where $m_{(t)}$ is the measured centroid mass at time t, $m_{(0)}$ is the measured mass of the undeuterated reference peptide, and $m_{(90)}$ is the measured mass of the fully deuterated reference peptide (in 90% D_2O).

The extent of deuterium incorporation (% D) into each sequence segment upon labelling of the oligomers by a 3000 s HDX pulse at pH 4, 25 °C, was used to provide an approximate estimate of k_{obs}, using the following equation:

$$\%D = 1 - e^{(-3000)(k_{obs})}$$

The HDX protection factor (P_f) for each sequence segment of the oligomers was determined as described above.

Unfortunately, the resolution of the peptide map was not sufficient to provide values of k_{obs} and hence, P_f, at the individual residue-level. Hence, the values of both k_{int} and k_{obs} used in the determination of P_f for a sequence segment were values averaged over all the residues of the peptide fragment corresponding to the sequence segment.

CRediT authorship contribution statement

Suman Pal: Conceptualization, Methodology, Software, Writing – original draft, Validation, Formal analysis, Investigation. **Jayant B. Udgaonkar:** Conceptualization, Writing – review & editing, Validation, Visualization, Resources, Supervision, Funding acquisition.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary Data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2022. 167854.

References

1. Kretzschmar, H., Prusiner, S., Stowring, L., DeArmond, (1986). Scrapie prion proteins are synthesized in neurons. *Am. J. Pathol.* **122**, 1.
2. Stahl, N., Borchelt, D.R., Hsiao, K., Prusiner, S.B., (1987). Scrapie prion protein contains a phosphatidylinositol glycolipid. *J Cell* **51**, 229–240.
3. Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R., Wüthrich, K., (2021). NMR structure of the mouse prion protein domain PrP (121–231). *In: NMR with Biological Macromolecules in Solution: A Selection of Papers Published from 1996 to 2020 by Kurt Wüthrich.* World Scientific, pp. 117–119.
4. Ning, L., Guo, J., Jin, N., Liu, H., Yao, X., (2014). The role of Cys179–Cys214 disulfide bond in the stability and folding of prion protein: insights from molecular dynamics simulations. *J. Mol. Model.* **20**, 1–8.
5. Steele, A.D., Lindquist, S., Aguzzi, A., (2007). The prion protein knockout mouse: a phenotype under challenge. *J. Prion.* **1**, 83–93.
6. Yu, G., Chen, J., Xu, Y., Zhu, C., Yu, H., Liu, S., Sha, H., Chen, J., et al., (2009). Generation of goats lacking prion protein. *J. Mol. Reprod. Develop..* 76.
7. Richt, J.A., Kasinathan, P., Hamir, A.N., Castilla, J., Sathiyaseelan, T., Vargas, F., Sathiyaseelan, J., Wu, H., et al., (2007). Production of cattle lacking prion protein. *J. Nat. Biotechnol.* **25**, 132–138.
8. Castle, A.R., Gill, A.C., (2017). Physiological functions of the cellular prion protein. *J. Front. Mol. Biosci.* **4**, 19.
9. Moore, R.A., Taubner, L.M., Priola, S.A., (2009). Prion protein misfolding and disease. *J Curr. Opin. Struct. Biol.* **19**, 14–22.
10. Thody, S.A., Mathew, M., Udgaonkar, J.B., (2018). Mechanism of aggregation and membrane interactions of mammalian prion protein. *J. Biochim. Biophys. Acta - Biomembranes* **1860**, 1927–1935.
11. Borchelt, D., Taraboulos, A., Prusiner, S., (1992). Evidence for synthesis of scrapie prion proteins in the endocytic pathway. *J. Biol. Chem.* **267**, 16188–16199.
12. Tixador, P., Herzog, L., Reine, F., Jaumain, E., Chapuis, J. M., Le Dur, A., Laude, H., Beringue, V., (2010). The physical relationship between infectivity and prion protein

aggregates is strain-dependent. *J. PLoS Pathogens* **6**, e1000859.

13. Khan, M.Q., Sweeting, B., Mulligan, V.K., Arslan, P.E., Cashman, N.R., Pai, E.F., Chakrabartty, A., (2010). Prion disease susceptibility is affected by β-structure folding propensity and local side-chain interactions in PrP. *Proc. Nat. Acad. Sci.* **107** (46), 19808–19813.

14. Rousseau, F., Serrano, L., Schymkowitz, J.W., (2006). How evolutionary pressure against protein aggregation shaped chaperone specificity. *J. Mol. Biol.* **355**, 1037–1047.

15. Reumers, J., Maurer-Stroh, S., Schymkowitz, J., Rousseau, F., (2009). Protein sequences encode safeguards against aggregation. *J. Human Mutation* **30**, 431–437.

16. Richardson, J.S., Richardson, D.C., (2002). Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc. Nat. Acad. Sci.* **99**, 2754–2759.

17. Levitt, M., (1981). Effect of proline residues on protein folding. *J. Mol. Biol.* **145**, 251–263.

18. Osváth, S., Gruebele, M., (2003). Proline can have opposite effects on fast and slow protein folding phases. *Biophys. J.* **85**, 1215–1222.

19. Theillet, F.-X., Kalmar, L., Tompa, P., Han, K.-H., Selenko, P., Dunker, A.K., Daughdrill, G.W., Uversky, V.N., (2013). The alphabet of intrinsic disorder: I. Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disordered Proteins* **1**, e24360.

20. Steward, A., Adhya, S., Clarke, J., (2002). Sequence conservation in Ig-like domains: the role of highly conserved proline residues in the fibronectin type III superfamily. *J. Mol. Biol.* **318**, 935–940.

21. Meuvis, J., Gerard, M., Desender, L., Baekelandt, V., Engelborghs, Y., (2010). The conformation and the aggregation kinetics of α-synuclein depend on the proline residues in its C-terminal region. *Biochemistry* **49**, 9345–9352.

22. Wedemeyer, W.J., Welker, E., Scheraga, H.A., (2002). Proline cis– trans isomerization and protein folding. *Biochemistry* **41**, 14637–14644.

23. Maschio, M.C., Fregoni, J., Molteni, C., Corni, S., (2021). Proline isomerization effects in the amyloidogenic protein β 2-microglobulin. *PCCP* **23**, 356–367.

24. Zhou, X.Z., Kops, O., Werner, A., Lu, P.-J., Shen, M., Stoller, G., Küllertz, G., Stark, M., et al., (2000). Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and tau proteins. *Mol. Cell* **6**, 873–883.

25. Burns, C.S., Aronoff-Spencer, E., Dunham, C.M., Lario, P., Avdievich, N.I., Antholine, W.E., Olmstead, M.M., Vrielink, A., et al., (2002). Molecular features of the copper binding sites in the octarepeat domain of the prion protein. *Biochemistry* **41**, 3991–4001.

26. Moore, R.A., Herzog, C., Errett, J., Kocisko, D.A., Arnold, K.M., Hayes, S.F., Priola, S.A., (2006). Octapeptide repeat insertions increase the rate of protease-resistant prion protein formation. *Protein Sci.* **15**, 609–619.

27. Gielnik, M., Pietralik, Z., Zhukov, I., Szymańska, A., Kwiatek, W.M., Kozak, M., (2019). PrP (58–93) peptide from unstructured N-terminal domain of human prion protein forms amyloid-like fibrillar structures in the presence of Zn 2+ ions. *RSC Adv.* **9**, 22211–22219.

28. Kraus, A., Anson, K.J., Raymond, L.D., Martens, C., Groveman, B.R., Dorward, D.W., Caughey, B., (2015). Prion protein prolines 102 and 105 and the surrounding lysine cluster impede amyloid formation. *J. Biol. Chem.* **290**, 21510–21522.

29. Shamsir, M.S., Dalby, A.R., (2007). β-Sheet containment by flanking prolines: Molecular dynamic simulations of the inhibition of β-sheet elongation by proline residues in human prion protein. *Biophys. J.* **92**, 2080–2089.

30. Popova, S.N., Tarvainen, I., Capellari, S., Parchi, P., Hannikainen, P., Pirinen, E., Haapasalo, H., Alafuzoff, I., (2012). Divergent clinical and neuropathological phenotype in a G erstmann–S träussler–S cheinker P 102 L family. *Acta Neurol. Scand.* **126**, 315–323.

31. Bernardi, L., Cupidi, C., Bruni, A.C., (2017). Pathogenic mechanisms of the prion protein gene mutations: a review and speculative hypotheses for pathogenic potential of the Pro39Leu mutation in the associated FTD-like phenotype. *J. Neurol. Neurosci.* **8**

32. Singh, J., Udgaonkar, J.B., (2016). Unraveling the molecular mechanism of pH-inducedy misfolding and oligomerization of the prion protein. *J. Mol. Biol.* **428**, 1345–1355.

33. Van der Kamp, M.W., Daggett, V., (2010). Influence of pH on the human prion protein: insights into the early steps of misfolding. *Biophys. J.* **99**, 2289–2298.

34. Sengupta, I., Bhate, S.H., Das, R., Udgaonkar, J.B., (2017). Salt-mediated oligomerization of the mouse prion protein monitored by real-time NMR. *J. Mol. Biol.* **429**, 1852–1872.

35. Uversky, V.N., Li, J., Fink, A.L., (2001). Evidence for a partially folded intermediate in α-synuclein fibril formation. *J. Biol. Chem.* **276**, 10737–10744.

36. Kusumoto, Y., Lomakin, A., Teplow, D.B., Benedek, G.B., (1998). Temperature dependence of amyloid β-protein fibrillization. *Proc. Natl. Acad. Sci.* **95**, 12277–12282.

37. Manno, M., Craparo, E.F., Podestà, A., Bulone, D., Carrotta, R., Martorana, V., Tiana, G., San Biagio, P.L., (2007). Kinetics of different processes in human insulin amyloid formation. *J. Mol. Biol.* **366**, 258–274.

38. Kumar, S., Mohanty, S.K., Udgaonkar, J.B., (2007). Mechanism of formation of amyloid protofibrils of barstar from soluble oligomers: evidence for multiple steps and lateral association coupled to conformational conversion. *J. Mol. Biol.* **367**, 1186–1204.

39. Choi, E.J., Mayo, S.L., (2006). Generation and analysis of proline mutants in protein G. *Protein Eng., Design Select.* **19**, 285–289.

40. Rose, G.D., (2021). Reframing the protein folding problem: Entropy as organizer. *Biochemistry* **60**, 3753–3761.

41. MacArthur, M.W., Thornton, J.M., (1991). Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**, 397–412.

42. Lee, J., Chang, I., Yu, W., (2019). Atomic insights into the effects of pathological mutants through the disruption of hydrophobic core in the prion protein. *Sci. Rep.* **9**, 1–14.

43. Morrissey, M., Shakhnovich, E., (1999). Evidence for the role of PrPC helix 1 in the hydrophilic seeding of prion aggregates. *Proc. Natl. Acad. Sci.* **96**, 11293–11298.

44. Moulick, R., Das, R., Udgaonkar, J.B., (2015). Partially unfolded forms of the prion protein populated under misfolding-promoting conditions: characterization by hydrogen exchange mass spectrometry and NMR. *J. Biol. Chem.* **290**, 25227–25240.

45. Bajaj, K., Madhusudhan, M.S., Adkar, B.V., Chakrabarti, P., Ramakrishnan, C., Sali, A., Varadarajan, R., (2007). Stereochemical criteria for prediction of the effects of proline mutations on protein stability. *PLoS Comput. Biol.* **3**, e241.

46. Jain, R., Dani, V., Mitra, A., Srivastava, S., Sarma, S.P., Varadarajan, R., Ramakumar, S., (2002). Structural consequences of replacement of an α-helical Pro residue in Escherichia coli thioredoxin. *Protein Eng.* **15**, 627–633.

47. Moulick, R., Udgaonkar, J.B., (2017). Identification and structural characterization of the precursor conformation of the prion protein which directly initiates misfolding and oligomerization. *J. Mol. Biol.* **429**, 886–899.

48. Hadži, S., Ondračka, A., Jerala, R., Hafner-Bratkovič, I., (2015). Pathological mutations H187R and E196K facilitate subdomain separation and prion protein conversion by destabilization of the native structure. *FASEB J.* **29**, 882–893.

49. Singh, J., Udgaonkar, J.B., (2015). Structural effects of multiple pathogenic mutations suggest a model for the initiation of misfolding of the prion protein. *Angew. Chem.* **127**, 7639–7643.

50. Kuwata, K., Nishida, N., Matsumoto, T., Kamatari, Y.O., Hosokawa-Muto, J., Kodama, K., Nakamura, H.K., Kimura, K., et al., (2007). Hot spots in prion protein for pathogenic conversion. *Proc. Natl. Acad. Sci.* **104**, 11921–11926.

51. Kamatari, Y.O., Hayano, Y., Yamaguchi, K.I., Hosokawa-Muto, J., Kuwata, K., (2013). Characterizing antiprion compounds based on their binding properties to prion proteins: implications as medical chaperones. *Protein Sci.* **22**, 22–34.

52. Lee, J., Chang, I., (2019). Structural insight into conformational change in prion protein by breakage of electrostatic network around H187 due to its protonation. *Sci. Rep.* **9**, 1–11.

53. Goluguri, R.R., Sen, S., Udgaonkar, J., (2019). Microsecond sub-domain motions and the folding and misfolding of the mouse prion protein. *Elife* **8**, e44766.

54. Bhate, S.H., Udgaonkar, J.B., Das, R., (2021). Destabilization of polar interactions in the prion protein triggers misfolding and oligomerization. *Protein Sci.* **30**, 2258–2271.

55. Sengupta, I., Udgaonkar, J., (2019). Monitoring site-specific conformational changes in real-time reveals a misfolding mechanism of the prion protein. *Elife* **8**, e44698.

56. Fooks, H., Martin, A., Woolfson, D., Sessions, R., Hutchinson, E., (2006). Amino acid pairing preferences in parallel β-sheets in proteins. *J. Mol. Biol.* **356**, 32–44.

57. Schmid, F.X., (1993). Prolyl isomerase: enzymatic catalysis of slow protein-folding reactions. *Ann. Rev. Biophys. Chem.* **22**, 123–143.

58. Honda, R.P., Xu, M., Yamaguchi, K.-I., Roder, H., Kuwata, K., (2015). A native-like intermediate serves as a branching point between the folding and aggregation pathways of the mouse prion protein. *Structure* **23**, 1735–1742.

59. Honda, R.P., Yamaguchi, K.-I., Kuwata, K., (2014). Acid-induced molten globule state of a prion protein. *J. Biol. Chem.* **289**, 30355–30363.

60. Moulick, R., Goluguri, R.R., Udgaonkar, J.B., (2019). Ruggedness in the free energy landscape dictates misfolding of the prion protein. *J. Mol. Biol.* **431**, 807–824.

61. Hu, W., Kan, Z.-Y., Mayne, L., Englander, S.W., (2016). Cytochrome c folds through foldon-dependent native-like intermediates in an ordered pathway. *Proc. Nat. Acad. Sci.* **113**, 3809–3814.

62. Hughson, F.M., Wright, P.E., Baldwin, R.L., (1990). Structural characterization of a partly folded apomyoglobin intermediate. *Science* **249**, 1544–1548.

63. Bai, Y., Sosnick, T.R., Mayne, L., Englander, S.W., (1995). Protein folding intermediates: native-state hydrogen exchange. *Science* **269**, 192–197.

64. Bollen, Y.J., Kamphuis, M.B., van Mierlo, C.P., (2006). The folding energy landscape of apoflavodoxin is rugged: Hydrogen exchange reveals nonproductive misfolded intermediates. *Proc. Nat. Acad. Sci.* **103**, 4095–4100.

65. Chamachi, N.G., Chakrabarty, S., (2017). Temperature-induced misfolding in prion protein: evidence of multiple partially disordered states stabilized by non-native hydrogen bonds. *Biochemistry* **56**, 833–844.

66. Yu, H., Liu, X., Neupane, K., Gupta, A.N., Brigley, A.M., Solanki, A., Sosova, I., Woodside, M.T., (2012). Direct observation of multiple misfolding pathways in a single prion protein molecule. *Proc. Nat. Acad. Sci.* **109**, 5283–5288.

67. Jain, S., Udgaonkar, J.B., (2008). Evidence for stepwise formation of amyloid fibrils by the mouse prion protein. *J. Mol. Biol.* **382**, 1228–1241.

68. Wang, L.-Q., Zhao, K., Yuan, H.-Y., Wang, Q., Guan, Z., Tao, J., Li, X.-N., Sun, Y., et al., (2020). Cryo-EM structure of an amyloid fibril formed by full-length human prion protein. *Nature Struct. Mol. Biol.* **27**, 598–602.

69. Sengupta, I., Udgaonkar, J.B., (2018). Structural mechanisms of oligomer and amyloid fibril formation by the prion protein. *Chem. Commun.* **54**, 6230–6242.

70. Singh, J., Udgaonkar, J.B., (2013). Dissection of conformational conversion events during prion amyloid fibril formation using hydrogen exchange and mass spectrometry. *J. Mol. Biol.* **425**, 3510–3521.

71. Singh, J., Sabareesan, A., Mathew, M., Udgaonkar, J.B., (2012). Development of the structural core and of conformational heterogeneity during the conversion of oligomers of the mouse prion protein to worm-like amyloid fibrils. *J. Mol. Biol.* **423**, 217–231.

72. Manka, S.W., Zhang, W., Wenborn, A., Betts, J., Joiner, S., Saibil, H.R., Collinge, J., Wadsworth, J.D., (2022). 2.7 Å cryo-EM structure of ex vivo RML prion fibrils. *Nat. Commun.* **13**, 1–11.

73. Kraus, A., Hoyt, F., Schwartz, C.L., Hansen, B., Artikis, E., Hughson, A.G., Raymond, G.J., Race, B., et al., (2021). High-resolution structure and strain comparison of infectious mammalian prions. *Mol. Cell* **81** (4540–4551), e6.

74. Hallinan, G.I., Ozcan, K.A., Hoq, M.R., Cracco, L., Vago, F. S., Bharath, S.R., Li, D., Jacobsen, M., et al., (2022). Cryo-EM structures of prion protein filaments from Gerstmann–Sträussler–Scheinker disease. *Acta Neuropathol.*, 1–12.

75. Smirnovas, V., Baron, G.S., Offerdahl, D.K., Raymond, G. J., Caughey, B., Surewicz, W.K.J.N.S., (2011). Structural organization of brain-derived mammalian prions examined by hydrogen-deuterium exchange. *Nat. Struct. Mol. Biol.* **18**, 504–506.

76. Singh, J., Kumar, H., Sabareesan, A.T., Udgaonkar, J.B., (2014). Rational stabilization of helix 2 of the prion protein prevents its misfolding and oligomerization. *J. Am. Chem. Soc.* **136**, 16704–16707.

77. Agashe, V.R., Udgaonkar, J.B., (1995). Thermodynamics of denaturation of barstar: evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride. *Biochemistry* **34**, 3286–3299. 107(46): 19808–19813.

78. Bai, Y., Milne, J.S., Mayne, L., Englander, S.W., (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins: Struct. Function, Bioinformat.* **17**, 75–86.

79. Nguyen, D., Mayne, L., Phillips, M.C., Walter Englander, S., (2018). Reference parameters for protein hydrogen exchange rates. *J. Am. Soc. Mass Spectrom.* **29**, 1936–1939.

80. Singh, J., Udgaonkar, J.B., (2016). The pathogenic mutation T182A converts the prion protein into a molten globule-like conformation whose misfolding to oligomers but not to fibrils is drastically accelerated. *Biochemistry* **55**, 459–469.