Journal Pre-proofs

Research Article

Loss of residues 119 – 136, including the first β-strand of human prion protein, generates an aggregation-competent partially "open" form

Laszlo L. P. Hosszu, Daljit Sangar, Mark Batchelor, Emmanuel Risse, Andrea M. Hounslow, John Collinge, Jonathan P. Waltho, Jan Bieschke

Please cite this article as: L. L. P. Hosszu, D. Sangar, M. Batchelor, E. Risse, A.M. Hounslow, J. Collinge, J.P. Waltho, J. Bieschke, Loss of residues 119 – 136, including the first β-strand of human prion protein, generates an aggregation-competent partially "open" form, *Journal of Molecular Biology* (2023), doi: https://doi.org/10.1016/j.jmb.2023.168158

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Classification: Biological Sciences; Biochemistry

Loss of residues 119 – 136, including the first β-strand of human prion protein, generates an aggregation-competent partially "open" form

Laszlo L. P. Hosszu[1], Daljit Sangar[1], Mark Batchelor[1], Emmanuel Risse[1], Andrea M. Hounslow[2], John Collinge[1], Jonathan P. Waltho[2, 3], and Jan Bieschke[1]

[1] MRC Prion Unit at UCL, UCL Institute of Prion Diseases, 33 Cleveland Street, London, W1W 7FF, UK

[2] Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK

[3] Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester, M1 7DN, UK

Author Contributions: LLPH, AMH, ER, JB designed research; LLPH, DS, MB, AMH, ER performed research, LLPH, AMH, ER, JB, AMH, JPW, analysed data; LLPH, JB, JPW wrote the paper; JB, JC provided funding and research infrastructure.

Corresponding author: Jan Bieschke; E-mail: j.bieschke@prion.ucl.ac.uk

Competing interest statement: J.C. is a Director and shareholder of D-Gen Limited, an academic spin-out company working in the field of prion disease diagnosis, decontamination, and therapeutics. The other authors declare no competing interests.

Data availability: The data that support the findings of this study are available from the corresponding author upon reasonable request. Assigned NMR chemical shifts have been deposited with the Biological Magnetic Resonance Bank (BMRB), under ID #51897.

This is an open access article

Abstract

In prion replication, the cellular form of prion protein (PrP^C) must undergo a full conformational transition to its disease-associated fibrillar form. Transmembrane forms of PrP have been implicated in this structural conversion. The cooperative unfolding of a structural core in PrP^C presents a substantial energy barrier to prion formation, with membrane insertion and detachment of parts of PrP presenting a plausible route to its reduction. Here, we examined the removal of residues 119 - 136 of PrP, a region which includes the first β-strand and a substantial portion of the conserved hydrophobic region of PrP, a region which associates with the ER membrane, on the structure, stability and self-association of the folded domain of PrP^C. We see an "open" native-like conformer with increased solvent exposure which fibrilises more readily than the native state. These data suggest a stepwise folding transition, which is initiated by the conformational switch to this "open" form of PrP^C.

Keywords

Prion protein, Neurodegeneration, Biophysics, Hydrogen-deuterium exchange

Abbreviations

AFM, Atomic force microscopy; CD, Circular dichroism; CJD, Creutzfeldt-Jakob disease; EM, Electron microscopy; FFI, Fatal familial insomnia; GuHCl, Guanidine hydrochloride; MRE, Mean residue ellipticity; PrP, prion protein; PrP^C, cellular PrP isoform; PrP^{Sc}, pathogenic (scrapie) PrP isoform; HuPrP, human PrP; PrP$^{23/91/119/137}$ residues 23/91/119/137–231 of human PrP; HSQC, Heteronuclear single quantum coherence; SV-AUC, Sedimentation velocity analytical ultracentrifugation; TSP, Sodium 3-trimethylsilyl-2,2,3,3-(^{2}H$_4$) propionate.

Introduction

Prion diseases, such as bovine spongiform encephalopathy in cattle, scrapie in sheep, and Creutzfeldt-Jakob disease (CJD), kuru, and Gerstmann-Sträussler-Scheinker syndrome in humans, are a group of neurodegenerative disorders caused by prions, self-replicating β-sheet-rich infectious polymeric assemblies of misfolded host-encoded cellular prion protein (PrP^C)[1-6]. Whilst rare, prion diseases are an area of intense research interest, as it is increasingly recognised that other degenerative brain diseases, such as Alzheimer's and Parkinson's diseases, also involve the accumulation and spread of aggregates of misfolded host proteins through an analogous process of seeded protein polymerisation[2, 7-10]. Consequently, study of 'prion-like' mechanisms has been recognised to have much a wider relevance to the understanding of neurodegenerative disorders[11-13].

PrP^C is a widely-expressed cell surface, glycosylphosphatidylinositol (GPI)-anchored glycoprotein that is sensitive to protease treatment and is soluble in detergents. The precise cellular function of PrP remains ill-defined, with studies suggesting many dissimilar roles, such as metal binding, binding to misfolded oligomeric proteins, signalling in complex with other membrane receptors, and cell adhesion, consistent with it being a multi-functional protein[14]. Such multi-functional proteins have an increased incidence in human disease[15]. PrP^C consists of two structural domains, an unstructured N-terminus, spanning approximately residues 23-125, which contains five repeats consisting of a single nonapeptide and four octapeptides, and a structured, mainly α-helical C-terminal domain, which includes a single disulphide bond and two glycosylation sites[1, 16, 17]. A highly conserved section, (approximately residues 110-136) referred to as the conserved hydrophobic region (CHR) spans both domains. Prions, in contrast, are dominated by β-sheet structure and are insoluble in detergents, with some displaying marked protease-resistance. These have been classically designated as PrP "Scrapie" (PrP^Sc) and are found only in prion-infected tissue[18].

Prion diseases may be acquired (transmitted between animals or humans), inherited or sporadic (of unknown cause). Inherited prion diseases comprise around 10-15% of total prion disease cases, with over 30 different pathogenic somatic mutations and approximately 10 polymorphic variants having been identified in the human PrP gene (*PRNP*)[19]. Amyloid formation in a range of proteins is associated with destabilisation of the protein native state through somatic mutations[20-23]. Although

destabilisation of PrP[C] does not correlate with specific disease phenotypes and no definitive link between PrP stability and disease has been established, the majority of *PRNP* pathogenic mutations reside within the structured C-terminal domain, with all fully penetrant pathogenic mutations showing significant destabilisation[24]. *In vitro* experiments also show that destabilisation of native PrP[C] through the use of chemical denaturants, temperature, redox conditions, pH or nucleic acid-binding favours the formation of protease-resistant amyloid or fibrillar structures, some of which have been claimed to be associated with disease[25-27]. The most general model proposed thus far for the process of nucleated protein polymerisation is that PrP[C] fluctuates between its dominant native state and other minor isoforms. These can self-associate in an ordered manner to produce a stable supramolecular structure composed of misfolded PrP monomers, which can convert other isoforms to the infectious isomer in an autocatalytic manner[2].

Many studies have thus focused on characterising partially folded intermediate states which may be involved in PrP self-association. A key observation in this regard is that the three α-helices and the second strand of the PrP[C] β-sheet form a stable assembly, which may be considered to be the core of the protein. This is because these structural elements have stabilities equivalent to the overall fold stability of the protein and the backbone amides of these structural elements exchange with solvent solely when PrP[C] completely unfolds[28]. Minor isoforms which retain this structural core are thus energetically favoured over those in which this core region is disrupted. Notably, the first strand of the PrP β-sheet, which is also the most N-terminal structural element in the folded C-terminal domain, lies outside of this core region and displays markedly reduced stability in comparison to the other PrP secondary structure elements (approximately 30-fold less)[28]. This raises the possibility that this secondary structure element may act independently of the core of PrP[C], and indeed NMR chemical shift data indicate that the PrP[C] native state ensemble contains a population of conformers where the PrP N-terminus, including the first strand of the PrP[C] β-sheet, is detached from the PrP[C] core[29].

The detached region comprises up to residue 146, and includes the conserved hydrophobic region of PrP (CHR), which spans residues 110 – 136. The CHR displays exceptionally high conservation across a wide range of species, indicative of an essential role in the endogenous function of PrP[C] [30, 31]. The CHR controls PrP co-translational translocation at the endoplasmic reticulum (ER) during the biosynthesis of the protein[32]. It can associate with the ER membrane in various topologies, with

mutations in this region upregulating particular PrP isoforms, which can lead to neurodegenerative disease in mice and some heritable prion diseases[33]. Its amino acid sequence also displays many of the characteristics of a transmembrane helix[34], and NMR studies show that residues within this region interact with membrane analogues (dodecylphosphocholine micelles) in aqueous solution both as part of PrP^C [35] and as a distinct peptide[31]. Here, it adopts a membrane-spanning helical conformation when associated with lipid micelles[31] and a radically different conformation to the reported solution structures of PrP^C. In addition, antibody binding to the CHR, which would inhibit membrane insertion, also inhibits the propagation of proteinase K-resistant PrP^{Sc} and prion infectivity, further implicating the CHR and its membrane association in prion disease[36]. Competitive antibody binding to the CHR has also been proposed to block interactions between PrP and cofactor molecules thereby inhibiting PrP conversion and prion propagation[36].

Insertion of the CHR into the endoplasmic reticulum (ER) membrane necessarily involves detachment of residues 110-136 from the folded C-terminal domain, including the first strand of the PrP β-sheet, from the PrP^C core. Whilst the proportion of PrP molecules with a detached CHR region is low in solution[29], a high local concentration of membrane-binding sites for membrane-incorporated PrP would be available, likely resulting in a significant bound population. Thus, it is reasonable to consider whether these membrane-associated forms of PrP^C may destabilise the PrP^C core region and facilitate its self-association. Our aim in this study was to investigate the effect of the removal of the CHR on the structure, stability and self-association of the remaining folded domain of PrP, through the biophysical characterisation of truncated PrP^C lacking the CHR. We find that this form of the protein whilst retaining many of the characteristics and structure elements of native PrP^C, has substantially increased solvent exposure relative to the PrP folded C-terminal domain lacking amino acids 119 - 136 and thus a substantial portion of the CHR, and is significantly more susceptible to fibrillisation than PrP retaining a partial CHR, the full CHR or the whole N-terminus, raising the possibility that this partially-folded state may be a principal precursor in the formation of ordered fibrillar structures.

Results

Choice of PrP constructs studied

In order to assess the effects of removal of the CHR on the structured domain of PrP^C we generated a recombinant PrP construct spanning residues 137 - 231, which we termed PrP^{137} (Fig. 1). Data for this construct was primarily compared with that of PrP^{119} (residues 119 - 231 of PrP), which contains the structured elements of PrP^C, but lacks the unstructured N-terminus and a substantial portion of the CHR, which complicates the structural analyses (Fig. 1), PrP^{91} (residues 91 - 231 of PrP), which spans the sequence incorporated into prions, and retains the full CHR, and full-length PrP (PrP^{23}) where appropriate.

PrP^{137} retains native-like secondary and tertiary structure

Circular dichroism (CD) and NMR spectra of PrP^{137} in solution were acquired to provide information on the secondary and tertiary structure of the protein. The far-UV CD spectrum of PrP^{137} (Fig. 2 (a)) is broadly similar to that of PrP^{119}, with two minima at approximately 208 and 222 nm, characteristic of a predominantly α-helical protein. The loss of β-sheet signal resulting from the loss of residues 119-136 is not apparent as it is masked by the significantly more intense α-helical signal. PrP^{137} thus appears to retain the mainly α-helical structure of the untruncated C-terminal domain. There is some reduction in the mean residue ellipticity (MRE) for the minimum at 208 nm, which is suggestive of some alteration in the packing of the α-helices however.

Both 1D and 2D NMR spectra of PrP^{137} display a broad range of dispersed NMR signals (Fig. 2 (b/c)). The wide chemical shift dispersion indicates extensive tertiary organisation, and tight packing of amino acid side chains in the protein core characteristic of a fully folded, native-like protein. This is significant, as in so-called 'molten globule' intermediate states, truncated or destabilised proteins can display native-like secondary structure content, and compact hydrodynamic radius, despite the absence of well-defined tertiary structure[37]. The characteristics found in molten globule states have been found in the transient intermediate states found during the folding of certain proteins[38]. However, PrP^{137} NMR spectra are markedly different to PrP^{119}, indicating considerable perturbation of the folded domain, in particular residues in close proximity to the N-terminal deletion (Fig. 2 (c)).

7

Structural perturbations in PrP[137] resulting from loss of the CHR

The full extent of the structural perturbations resulting from the loss of residues 119 - 136 was determined through analysis of backbone NMR chemical shifts. Differences between observed chemical shifts and their corresponding random coil (unstructured) values, in particular in Cα and carbonyl resonances, are highly correlated with protein secondary structure[39, 40]. Chemical shift data also provides accurate, quantitative and site-specific mapping of protein backbone mobility as well as backbone order parameters[41].

NMR resonance chemical shifts (Cα/C′/Cβ/N/H^N) show that the backbone conformation of PrP[137] is very similar to PrP[119], with all three α-helices retained (Fig. 3)[39, 40]. Remarkably, although some loss of structure is observed, the second β-strand also retains clearly observable extended structure despite the loss of the complementary β-strand and its stabilising hydrogen bonding contacts (Fig. 3). This may be due to its tight association with helices 2 and 3, comprising close packing and burial of the hydrophobic side-chain of V161, and also the need to accommodate the bulky aromatic side-chains of Y162 and Y163.

Some change in protein dynamics is observed, as indicated by changes in the backbone order parameters (S^2). S^2 reports on internal sub-nanosecond (ns) motions, and values range from 0 for highly flexible to 1 for rigid systems. The helical regions of PrP variants exhibit S^2 values of 0.8 – 0.9, typical of structured regions of folded proteins. Residues of the α2–α3 loop (residues 194–199) and the C-terminus of helix 3 display reduced S^2 values, reflecting increased flexibility, commonly observed in loop regions of globular proteins[42, 43]. A reduction in S^2 for residues 164-168, which are part of the so-called β2-α2 loop linking the second strand of the β-sheet and helix 2 (residues 165–172; Fig. 1, Fig. 3 (c)) implies that the β2-α2 loop is more flexible in PrP[137] than PrP[119]. This loop, which has been shown to affect prion cross-species transmissibility[44-48], is adjacent to the β-sheet, packing against residues N-terminal to the first β-strand, the β-sheet itself, and the C-terminus of helix 3 (Fig. 1). The reduced order parameters would appear to be at odds with the predicted increase in α-helicity, however, this loop is subject to millisecond (ms) timescale motions proposed to be associated with a large-scale co-operative conformational change between a 3_{10}-helix and a type I β-turn[49, 50]. Millisecond timescale motions can result in line-broadening of NMR signals and anomalous order parameter values[42, 43]. In addition, as observed in the majority of prion proteins,

8

backbone assignments for residues 167, 169 - 171 could not be determined, as their NMR signals were line-broadened beyond detection in PrP[119] and PrP[137] (Fig. S1). [13]C backbone shifts could be determined from their succeeding residues, however as a result of this line-broadening chemical shift data for this region is more limited, and is reflected in the lower level of confidence of the secondary structure prediction for this region. However the C-terminus of helix 3 (residues 219 – 230) also displays reduced S^2 values in PrP[137] (Fig. 3 (c)). A number of residues in this region of the protein interact closely with the β2-α2 loop. For example, Y218 and S222 with M166, and I215 and Y218 with Q172. Perturbation of the dynamics of the β2 - α2 loop has previously been shown to affect the dynamics of the C-terminus of helix 3 [43], so alteration on the order parameters for these regions are consistent with loss of the N-terminal residues affecting regions beyond the immediate site of the truncation. In addition, the C-terminus of helix 1 and following residues (residues 153-158) also display an increased level of flexibility (Fig. 3 (c)).

PrP[137] stability

Despite the relative lack of secondary structure perturbations observed, removal of residues 119 - 136, including the first β-strand resulted in a marked reduction of approximately 2 kcal mol[-1] in the thermodynamic stability of PrP[137] (Fig. 4 (b)). This equates to an approximately 20-fold reduction in the equilibrium between the native and unfolded states ($K_{(N/U)}$ ~2000) when compared to PrP[119]. As with PrP[119], loss of secondary structure upon equilibrium denaturation in PrP[137] occurs in a single co-operative transition, without the formation of any populated intermediate species (Fig. 4 (a)). This unfolding transition is markedly less cooperative than in PrP[119], but distinct from unfolding transitions demonstrated by molten globule states[29]. The calculated *m*-values, which describe the sensitivity of the folded / unfolded state equilibrium to denaturant, and reflect the increase in solvent exposure of the hydrophobic core as the protein unfolds show that that the degree of hydrophobic exposure in PrP[137] on unfolding is significantly reduced in comparison to PrP[119] (Fig. 4 (c)).

In common with the denaturant-induced unfolding, the thermal unfolding of PrP[137] consisted of a single co-operative transition, without the formation of any populated intermediate species (Fig. 4 (d)). Both equilibrium unfolding transitions (GuHCl- or thermally-induced), were found to be completely reversible (Fig. 4 (d/f)). The thermal stability of PrP[137] was also markedly reduced in

comparison with PrP[119], with a significant (≈ 10°C) reduction in the mid-point for thermal unfolding (Fig. 4 (e)).

Stability of secondary structure elements and solvent accessibility.

An even more sensitive measure of local stability is hydrogen/deuterium exchange of backbone amides, observed by the decay of NMR amide signals. As with PrP[119], protected amides in PrP[137] were located within the structured elements of the protein, in this case predominantly within helices 2 and 3 (Fig. 5). Protection factors (PF) for amides in these regions were equivalent to the free energy change for unfolding (*i.e.* PF = $K_{(N/U)}$). This behaviour is observed in PrP[119], similar length PrP constructs such as PrP[121] [51] and also a number of other proteins (*e.g.* Barnase/Staphylococcal Nuclease)[52], where a substantial proportion of core residues can exchange only in the fully unfolded state (Fig. 5). Protection factors for the second strand of the PrP β-sheet of PrP[137] were however below detectable levels. This is consistent with the loss of the first strand of the β-sheet and its hydrogen-bonding interactions with the second β-strand. However, it was also observed that the first α-helix in PrP[137] displayed markedly less protection than the remaining α-helices. In PrP[137] it does not display measurable protection (the lower limit for detection of protection factors was approximately 200), whereas in PrP[119] and PrP[121] it has amide protection factors comparable to the other secondary structure elements, and equivalent to the equilibrium constant for unfolding (Fig. 5 (a)). The stability of helix 1 relative to the other secondary structure elements is maintained in untruncated PrP constructs both at acidic (PrP[119] at pH 5.5) and neutral pH (PrP[121] at pH 7.0), despite the marked loss of stability associated with the protonation of key residues in PrP at lower pH[53, 54]. This loss of stability precludes measurement of hydrogen exchange rates under conditions where the intrinsic exchange rates are lower[55]. In contrast, loss of residues 119 -136, including the PrP β-sheet, selectively destabilises α-helix1 relative to the other α-helices in PrP[137], as a general loss of stability (for example through reduced pH), is associated with significantly reduced protection factors in all secondary structure elements (Fig. 5).

Propensity of PrP[137] to oligomerise

Given the increased exposure of hydrophobic residues, as indicated by the increased *m* value derived from the equilibrium denaturation data and the reduced hydrogen protection for the first α-helix and 2nd β-strand, we sought to determine whether this increased the propensity of PrP[137] to self-associate. This was initially done using sedimentation velocity analytical ultracentrifugation (SV-

AUC) (Fig. 6). A single species with a sedimentation coefficient of 1.6S, characterised by a frictional ratio (*f/f0*) of 1.25, giving a molecular mass of 11.8 kDa was observed for PrP[137] at pH 7.5. Corresponding values at pH 6.5 were 1.67S, 1.23 and 11.95 kDa respectively. These molecular masses match closely the expected molecular mass for the PrP[137] monomeric protein (Fig. 6). For PrP[119] a single species was also observed at pH 6.5 and 7.5 with a molecular mass corresponding with those of the monomeric protein (13.6 kDa). PrP[137] displays a slightly reduced frictional ratio relative to PrP[119] indicating a more spherical character, consistent with removal of the relatively unstructured N-terminus (residues 119-125) of PrP[119]. Thus it appeared that under the conditions used, PrP[137] is no more susceptible to oligomerisation than PrP[119], despite the increased exposure of hydrophobic residues. This conclusion was tested further over a range of protein concentrations (15 – 135 µM) at pH 6.5 and 7.5, with no multimeric species observed (Fig. 6).

Propensity of PrP[137] to fibrilise

When agitated at 42°C under native conditions, PrP can be induced to form amyloid[56]. Binding of the fluorescent thiazole dye thioflavin T to these β-sheet-rich fibrillar structures reports their formation, allowing a quantitative analysis of the kinetics of fibril formation[57]. Although no association of native PrP[137] monomers was observed in solution, we found that PrP[137] can be induced to fibrilise much more readily than full-length WT PrP (PrP[23]), PrP[91] (which contains the CHR) and PrP[119] (which retains a portion of the CHR), with sigmoid kinetics typical of amyloid formation, but with significantly shorter half- and lag-times (Fig. 7 (a/b) & Fig. S2). Notably we find that shortening the unstructured N-terminus, in particular deletion of a portion of the CHR up to residue 119, inhibits fibrillisation. This observation is supported by other studies which show that the unstructured PrP[C] N-terminus can direct its oligomeric association[58] and contrasts with the increased fibrillisation propensity seen with PrP[137]. This highlights the importance of removal of the CHR and first β-strand on the fibrillisation characteristics of PrP[137]. Atomic Force Microscopy (AFM) and negative stain electron microscopy confirmed PrP[137] fibril formation (Fig. 7 (c/d)). Both narrow straight (~7 nm diameter) and twisted fibrils (~18 nm diameter) typical of amyloid fibrils were observed (Fig. 7 (d)).

11

Discussion

In this study, we have examined the effect of the removal of the N-terminus of the folded domain of PrP, on the structure, stability and fibrillisation characteristics of PrP^C. This truncated form of PrP^C mimics the structural effect of the association of the conserved hydrophobic region with the ER membrane on the folded C-terminal domain of PrP^C.

We find that the truncated protein retains the native secondary structure elements of PrP^C, including, remarkably, the second strand of the PrP β-sheet, but with increased conformational flexibility at the C-termini of α-helices 1 and 3, relative to PrP lacking the first β-strand and a substantial portion of the CHR (PrP^{119}). It also retains a well-defined tertiary structure, with strong evidence for co-operative folding. There is a marked reduction in stability and increased exposure of hydrophobic residues however, which would appear to facilitate its conversion to fibrillar states significantly. The structural elements of PrP^C most affected by the deletion are the second strand of the β-sheet and the first α-helix, both of which display markedly increased solvent exposure and reduced stability relative to the other structure elements.

Previous hydrogen exchange data show that the core of PrP^C, comprising the three α-helices and the second β-strand, is a stable entity, which has to completely unfold for exchange with the solvent to occur[28]. These data suggest that for fully folded PrP^C, the most likely route to PrP^{Sc} is through the completely unfolded state, and that any partially-folded intermediate states involved in this conversion would most likely retain this core region.

Under partially denaturing, acidic conditions however, various regions of mouse PrP^C have been shown to undergo sub-global unfolding, forming at least two distinct partially unfolded forms (PUFs)[59]. These are in equilibrium with the native state and display increased solvent exposure relative to it. A key characteristic of the second of these PUFs is that the first α-helix, second strand of the β-sheet and loop regions in between are disordered and solvent accessible. In this study we find that that removal of the PrP N-terminus to residue 137, mimicking the insertion of the CHR into the ER membrane, also markedly reduces the stability of α-helix 1 and the second β-strand, generating a partially unfolded form of PrP^C which displays many of the characteristics of the PUF

generated under partially-denaturing acidic conditions. Both of these open forms fibrillise much more readily than the native form of the protein, and resemble a conformation implicated to be an initial intermediate in the conversion of monomeric PrP into misfolded oligomer at pH 4 [60]. α-helix 1 in particular has been implicated in the conversion to PrP[Sc] [61, 62], with the anti-PrP therapeutic monoclonal antibody ICSM18 effectively curing prion-infected cells through binding to and stabilising α-helix 1 of PrP[C] [63]. Indeed its humanised version has been used in the first in-human treatment programme using an anti-PrP[C] monoclonal antibody[64].

The majority of the CHR itself is largely unprotected from amide exchange in PrP[C] [16, 28, 51, 65], including the first PrP β-strand, which displays anomalously low protection factors in PrP[C]. This is intriguing given the full protection seen in the other paired β-strand. The detachment of the CHR from the PrP core thus raises two possibilities, firstly that the PrP β-sheet is maintained under native conditions and the loop regions not including the first β-strand detaches, or that secondly, the residues within the β-sheet that are involved in hydrogen bonding within the PrP β-sheet are involved in hydrogen bonding with other hydrogen bond acceptors. The marked reduction in protection for the second β-strand in PrP[137] would suggest that the first possibility is more likely, at least in solution.

Overall, our results are concordant with other *in vitro* experiments that show that destabilisation of PrP[C] favours the formation of fibrillar and/or protease resistant structures[66, 67], a number of which have been proposed to be associated with the disease process[25-27]. PrP[C]/PrP[Sc] conversion may thus be initiated or facilitated by an increased population of this aggregation competent state, formed by insertion of the CHR into the ER membrane. This proposal is consistent with the observation that antibody sequestration of the CHR inhibits mouse prion propagation[36]. This may be through sequestration of the CHR from membrane insertion and/or stabilisation of the PrP[C] native state.

Our data suggest an alternative pathway to structural conversion of PrP to the prion conformer, in which the detached CHR residues destabilise the folded core of the protein, thus lowering the energy barrier to its structural conversion. The CHR could interact with the active surface of the prion fibril first, thus destabilising the rest of the protein core and generating the "open" form characterised here, which refolds afterwards. Alternatively, destabilisation by the membrane-bound CHR could facilitate interaction of α-helix 1 with the prion, initiating the structural conversion process. The data presented thus highlight the key role of the PrP first β-strand and N-terminus on

13

the stability and association of the folded domain of PrP^C, and provide an intriguing explanation as to how this may facilitate the conversion of PrP^C to PrP^{Sc}.

Materials and Methods

Protein Expression and Purification

The open reading frame of the human PrP gene (PRNP) (residues 137-231), containing methionine at residue 129, was synthesised *de novo* by *Eurofins MWG Operon*, with a thrombin-cleavable His-Tag added to the PrP N-terminus. The ligated pTrcHisB/PRNP construct was used to transform the *Escherichia coli* host strain BL21(DE3) (*Novagen*), genotype *F' ompT hsdSB (rB- mB-) gal dcm* (DE3), which was then plated onto Luria-Bertoni (LB) agar plates containing 100 μg/ml carbenicillin. Cultures were grown for purification using a modification of protocols previously described[29]. Briefly, following harvesting, cells were sonicated and their inclusion bodies containing PrP resolubilised in 6M Guanidine Hydrochloride (GuHCl), 50 mM Tris.Cl, 0.8% β-mercaptoethanol, pH 8.0. These were loaded onto a Ni-NTA column equilibrated in 6 M GuHCl, 10 mM Tris-HCl, 100 mM Na_2PO_4, 10mM glutathione pH 8.0, and eluted from the column using 10mM Tris-HCl, 100 mM Na_2PO_4, 1M Imidazole pH 7.0. Residual GuHCl was removed through dialysis against 25 mM Tris.HCl pH 8.4, $CaCl_2$ added to a final concentration of 2.5 mM, and the N-terminal His-tag cleaved by thrombin for 16 h at room temperature (0.1 U thrombin (*Novagene*)/1 mg of PrP added). The cleaved protein was loaded onto Q-Sepharose column equilibrated with 10 mM HEPES pH 8.2, and the PrP eluted by the same buffer containing 1M NaCl. PrP^{137} was dialysed against 10 mM HEPES pH 7.5, and stored at -80°C. Protein concentrations were determined by UV absorption using a calculated molar extinction of 15025 M^{-1} cm^{-1} at 280 nm (*https://web.expasy.org/protparam*).

Circular Dichroism (CD) spectroscopy

CD spectra were measured with a Jasco J-715 spectropolarimeter. Far-UV (amide) CD spectra (300 – 180 nm) were acquired on 22.5 μM protein at 25°C using 0.5 mm pathlength quartz cuvettes. The sample temperature was controlled with a circulating water bath. 45 spectra were averaged.

Equilibrium unfolding experiments

6.5 μM PrP^{137} and PrP^{119} in 10 mM sodium phosphate, 1 mM sodium azide, pH 7.0, were incubated in increasing concentrations of GuHCl denaturant at 25°C. Molecular ellipticity ([θ], degree M^{-1} cm^{-1}) was recorded at 222 nm (5 nm bandwidth; 20 s integration time) in the Jasco J-715

15

spectropolarimeter. The denaturation profile for each protein was measured in 3 separate experiments.

Calculating the equilibrium constant between folded and unfolded states (K & K_w)

For the two-state equilibrium unfolding transitions, data were fitted to the following equation, where K and $K_{(W)}$ are equilibrium constants between the folded and unfolded states at a given denaturant activity (D) and in water, respectively, and *m* describes the sensitivity of the equilibrium to denaturant activity[68].

$$K = K_{(W)} \exp (m . D) \qquad \text{(Eqn. 1)}$$

For visual representation of the data shown, data were converted to proportion folded, α_F, using the following, $\alpha_F = (K/(1+K))$. Data fitting was carried out using *GNUplot* v. 5.2. The significance of the differences in free energy for folding and *m* values between PrP^{119} and PrP^{137} were determined by two-tailed student's t-Test.

Calculation of Denaturant Activity

Due to the non-linear relationship between denaturant concentration and the free energy for folding, GuHCl concentration ([GuHCl]) was converted to molar denaturant activity to obtain a more reliable extrapolation of data at high denaturant concentrations. The equation used is

$$D = [GuHCl](7.5M/(7.5M + [GuHCl])) \qquad \text{(Eqn. 2)}$$

where D is the molar denaturant activity[68].

Equilibrium Thermal Denaturation Monitored by CD

The amide CD absorption of 6.5 μM PrP^{137} in 10 mM HEPES, 25 mM NaCl, 1 mM sodium azide, pH 7.0, was recorded at increasing temperature (1 °C/minute change). The ellipticity signal (α) was converted to the proportion of molecules in the native state α_N according to the relationship α_N = (θ - θ_U)/(θ_N - θ_U), where θ_U and θ_N are the ellipticity signals for the unfolded and native states, respectively. The Van't Hoff enthalpy (H), temperature mid-point for thermal folding (T_m), and equilibrium constant for folding ($K_{(FU)}$)were calculated as follows.

$$K_{(FU)} = \exp((H/R)*(1/T_m-1/T)) \quad \text{(Eqn. 3)}$$

where R is the gas constant (kcal mol^{-1}K^{-1}). All equilibrium unfolding transitions, GuHCl- or thermally induced, were reversible.

NMR spectroscopy

Uniformly ^{15}N- and ^{13}C/^{15}N-labeled PrP$^{137–231}$ samples were expressed in *E. coli* using an EMBL minimal medium recipe with ^{13}C$_6$-glucose and (^{15}NH$_4$)$_2$SO$_4$ as the sole carbon and nitrogen sources, respectively, and purified as described above. NMR spectra were acquired at 293K on 1.0 mM ^{15}N- and ^{13}C/^{15}N-labeled samples in 10 mM sodium phosphate, 2 mM sodium azide, 1mM TSP, pH 7.0 (in either 90% H$_2$O plus 10% D$_2$O *(v/v)* or 50% D$_2$O *(v/v)*) using Bruker DRX-500 and DRX-600 spectrometers equipped with 5-mm ^{13}C/^{15}N/^{1}H triple resonance probes. NMR samples were placed in Sigma FEP NMR sample tube liners (Z286397-1EA) held within Wilmad PP-528 NMR tubes for NMR data acquisition. Proton chemical shifts were referenced to TSP, with ^{15}N and ^{13}C chemical shifts calculated relative to TSP, using the gyromagnetic ratios of ^{15}N, ^{13}C, and ^{1}H (^{15}N/^{1}H = 0.101329118, ^{13}C/^{1}H = 0.251449530). NMR data were processed and analysed on Linux Workstations using Felix 2007 (Accelrys, San Diego) software.

Backbone NMR assignment

Backbone resonances (^{H}N, N, Cα, C', Cβ) of PrP^{137} were obtained using a standard suite of triple resonance NMR experiments[69], and assigned using the *asstools* set of assignment programs[70]. Almost complete backbone assignments were determined, the exceptions being the amide protons and nitrogens of residues 167, 169–171, and 175, for which no resonances were detected. These

residues occupy a loop region between β-strand 2 and α-helix 2, which is undergoing conformational exchange, resulting in line-broadening of NMR signals[43].

Prediction of secondary structure propensity and protein flexibility

Secondary structure propensity was calculated from H^N, $C\alpha$, $C\beta$, CO, & N random coil chemical shifts and average secondary shifts for protein secondary structure using the program TALOS-N[40]. TALOS-N is an artificial neural network based system for empirical prediction of protein backbone φ/ψ torsion angles, sidechain χ1 torsion angles and secondary structure using chemical shifts. The order parameters S^2 for the backbone amide groups were predicted from backbone (H^N, $C\alpha$, CO, & N) and $C\beta$ chemical shifts using the Random Coil Index (RCI) approach as implemented within TALOS-N[40, 41].

Amide Hydrogen-Deuterium Exchange

Hydrogen-deuterium exchange rates (k_{ex}) were determined by adding 140 μl 10 mM sodium phosphate, 1 mM sodium azide, pH 7.0, dissolved in 100% (*v/v*) D_2O to the same volume of 1 mM PrP[137] in the equivalent protonated buffer. A series of 2D sensitivity-enhanced ^1H-^{15}N HSQC spectra[71, 72] were acquired at 293K on a Bruker DRX-800 spectrometer. The decay curves of the ^1H-^{15}N HSQC cross-peaks were fitted to single exponential decays with offset, using a *Numerical Recipes in Fortran 77* Levenberg–Marquardt algorithm (Vol. 1., Cambridge University Press) and protection factors (k_{ex}/k_{int}) for observable amides were determined using intrinsic amide exchange rates (k_{int})[55]. Parameters determined in Bai et al (1993)[55] were used to obtain the intrinsic amide exchange rates (k_{int}) rates expected in unstructured oligo and poly-peptides. This was implemented using an *awk* script, with pH, temperature, protein sequence, and Bai parameters as input. Acquisition of the first experiment began approximately 5 minutes after mixing.

Analytical Ultracentrifugation

Sedimentation velocity ultracentrifugation experiments (SV-AUC) were carried out using a Beckman Optima XL-I analytical ultracentrifuge. Samples were loaded into Beckman AUC sample cells with 12 mm optical path two-channel centrepieces, with matched buffer in the reference sector. Cells were spun at 50,000 rpm in an AnTi-50 rotor and scans were acquired using both interference and

absorbance optics (at 280 nm) at 10 minute intervals over 16 hours. The sedimentation profiles were analysed using the software SEDFIT (v13b)[73]. Partial specific volumes (ρ) for PrP[137] were calculated from the amino acid sequence using SEDNTERP software[74]. Buffer densities and viscosities were measured using an Anton Paar DMA5000 density meter and an Anton Paar AMVn automated microviscometer, respectively. Sedimentation velocity data were analysed using the *c(s)* method of distribution[73] to characterise the sedimentation coefficient distribution of all species present in solution. The proportions of each sample occupying the main peaks in the distribution were calculated by integration of the peaks.

Quantitative analysis of the kinetics of PrP fibril formation

PrP[137], PrP[23], PrP[91] and PrP[119] were diluted to 5µM in 125mM Sodium Phosphate, 150mM NaCl, 7.5mM Bis-Tris pH 6.7, 1 µM Thioflavin T. All solutions were filtered through a 0.22 µm filter to remove particulates. 100 µl aliquots were placed in Greiner or Corning 96-well flat-bottomed plates (#655077 or 3651)_containing 3 0.5 mm diameter zirconium ceramic beads in each well to assist agitation, and seeded with 0.05% (v/v) homogenous fibril preparations. The plates were incubated at 42 °C with constant agitation in Tecan Infinite F200 Microplate or BMG Fluorostar Omega Fluorimeters. Fibril formation was monitored through the increase in ThT fluorescence (excitation 430 nm, emission 485 nm, 20 nm bandwidths), with readings acquired every 300 or 600 seconds. 3 – 5 replicates were used for each PrP sample.

To determine the half- and lag-times for fibril formation, data were fitted to an empirical function described by Nielsen et al (2001)[75].

$$F_i + F_f/\{1 + exp[-(t - t_m)/\tau]\} \qquad \text{(Eqn. 4)}$$

where F_i is the initial fluorescence reading, F_f is the final fluorescence reading, t is time, t_m is the time taken to half maximal fluorescence and τ is the reciprocal of the propagation rate during the rise phase [$1/k_{(apparent)}$]. Lag-time is defined as t_m - 2τ.

Electron Microscopy

End-point samples (5 - 6 µL) from aggregation assays were sonicated in silanised Eppendorf tubes using a Grant XUBA1 water-bath for 10 - 15 seconds, and loaded onto carbon-coated 300 mesh copper grids (Electron Microscopy Sciences) that had been glow discharged for 40 seconds using a PELCO easiGLOW glow discharge unit (Ted Pella Inc., USA). Samples were left to bind for 30 minutes, blotted dry, washed in water (1 x 50 µL), blotted, and then stained with 10 µL Nano-W (methylamine tungstate) stain (*Nanoprobes*) for 1 minute, then 30 seconds blotting in-between stains. Images were acquired on a TALOS electron microscope (*FEI*, Eindhoven, NL now *Thermofisher*).

Atomic Force Microscopy

End-point samples (20 µL) from aggregation assays were washed by pelleting in a benchtop microfuge (30 minutes, 16,100 x g, 25°C), followed by removal of supernatant, and resuspension of pellets in 20 µL de-ionised water. These were sonicated in silanised Eppendorf tubes using a Grant XUBA1 water-bath for 10 - 15 seconds, and 20 µL placed on clean, freshly cleaved grade V-1 mica (Cat#: AGG250-7, *Agar Scientific Ltd.*, UK). After 10 minutes, the solvent was wicked off by filter paper and the mica washed 4 times with 20 µL of de-ionised water to remove residual salts and buffer from the sample. Samples were dried overnight, and AFM images acquired in contact mode on a *Nanosurf* C3000i Atomic Force Microscope (*Nanosurf AG*, Switzerland) using TAP 190-A1-G tips. Images were visualised using the *Nanosurf* Imaging software (v. 3.10.2), and height profiles analysed using *Gwyddion* 2.61 software (*http://gwyddion.net*).

Statistics and reproducibility

In the reported experiments, each protein sample was identically-engineered. The sample size (n) of each experiment is provided in the corresponding figure captions in the main manuscript and supplementary information files. Sample sizes were chosen to support meaningful conclusions. All *in vitro* folding experiments were replicated at least three times. *In vitro* fibrillisation assays were replicated three - five times. Investigators were not blinded during experimental measurements or data analysis.

Acknowledgements

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. Assigned NMR chemical shifts have been deposited with the Biological Magnetic Resonance Bank (BMRB), under ID #51897.

References

1. Prusiner, S. B. (1998). Prions. *Proc Natl Acad Sci USA* **95**, 13363-13383.

2. Collinge, J. & Clarke, A. (2007). A general model of prion strains and their pathogenicity. *Science* **318**, 930-936.

3. Terry, C., Wenborn, A., Gros, N., Sells, J., Joiner, S., Hosszu, L. L., Tattum, M. H., Panico, S., Clare, D. K., Collinge, J., Saibil, H. R., & Wadsworth, J. D. (2016). Ex vivo mammalian prions are formed of paired double helical prion protein fibrils. *Open Biol* **6**,

4. Terry, C., Harniman, R. L., Sells, J., Wenborn, A., Joiner, S., Saibil, H. R., Miles, M. J., Collinge, J., & Wadsworth, J. D. F. (2019). Structural features distinguishing infectious ex vivo mammalian prions from non-infectious fibrillar assemblies generated in vitro. *Sci Rep* **9**, 376.

5. Kraus, A., Hoyt, F., Schwartz, C. L., Hansen, B., Artikis, E., Hughson, A. G., Raymond, G. J., Race, B., Baron, G. S., & Caughey, B. (2021). High-resolution structure and strain comparison of infectious mammalian prions. *Mol Cell* **81**, 4540-4551.

6. Manka, S. W., Zhang, W., Wenborn, A., Betts, J., Joiner, S., Saibil, H. R., Collinge, J., & Wadsworth, J. D. F. (2022). 2.7Å cryo-EM structure of ex vivo RML prion fibrils. *Nat Commun* **13**, 4004.

7. Griffith, J. S. (1967). Self Replication and scrapie. *Nature* **215**, 1043-1044.

8. Eisenberg, D. & Jucker, M. (2012). The amyloid state of proteins in human diseases. *Cell* **148**, 1188-1203.

9. Jaunmuktane, Z., Mead, S., Ellis, M., Wadsworth, J. D., Nicoll, A. J., Kenny, J., Launchbury, F., Linehan, J., Richard-Loendt, A., Walker, A. S., Rudge, P., Collinge, J., & Brandner, S. (2015). Evidence for human transmission of amyloid-beta pathology and cerebral amyloid angiopathy. *Nature* **525**, 247-250.

10. Purro, S. A., Farrow, M. A., Linehan, J., Nazari, T., Thomas, D. X., Chen, Z., Mengel, D., Saito, T., Saido, T., Rudge, P., Brandner, S., Walsh, D. M., & Collinge, J. (2018). Transmission of amyloid-beta protein pathology from cadaveric pituitary growth hormone. *Nature* **564**, 415-419.

11. Jucker, M. & Walker, L. C. (2013). Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45-51.

12. Walker, L. C. & Jucker, M. (2015). Neurodegenerative Diseases: Expanding the Prion Concept. *Annu Rev Neurosci* **38**, 87-103.

13. Collinge, J. (2016). Mammalian prions and their wider relevance in neurodegenerative diseases. *Nature* **539**, 217-226.

14. Hirsch, T. Z., Martin-Lanneree, S., & Mouillet-Richard, S. (2017). Functions of the Prion Protein. *Prog Mol Biol Transl Sci* **150**, 1-34.

15. Espinosa-Cantu, A., Cruz-Bonilla, E., Noda-Garcia, L., & DeLuna, A. (2020). Multiple Forms of Multifunctional Proteins in Health and Disease. *Front Cell Dev Biol* **8**, 1-10.

16. Zahn, R., Liu, A. Z., Lührs, T., Riek, R., Von Schroetter, C., García, F. L., Billeter, M., Calzolai, L., Wider, G., & Wüthrich, K. (2000). NMR solution structure of the human prion protein. *Proc. Natl. Acad. Sci. USA* **97**, 145-150.

17. Hosszu, L. L. P., Jackson, G. S., Trevitt, C. R., Jones, S., Batchelor, M., Bhelt, D., Prodromidou, K., Clarke, A. R., Waltho, J. P., & Collinge, J. (2004). The residue 129 polymorphism in human prion protein does not confer susceptibility to CJD by altering the structure or global stability of PrP^C. *J. Biol. Chem.* **279**, 28515-28521.

18. Collinge, J. (2001). Prion diseases of humans and animals: their causes and molecular basis. *Annu. Rev. Neurosci.* **24**, 519-550.

19. Beck, J. A., Poulter, M., Campbell, T. A., Adamson, G., Uphill, J. B., Guerreiro, R., Jackson, G. S., Stevens, J. C., Manji, H., Collinge, J., & Mead, S. (2010). PRNP allelic series from 19 years of prion protein gene sequencing at the MRC Prion Unit. *Hum Mutat* **31**, E1551-E1563.

20. Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C. F., & Pepys, M. B. (1997). Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* **385**, 787-793.

21. Ramirez-Alvarado, M., Merkel, J. S., & Regan, L. (2000). A systematic exploration of the influence of the protein stability on amyloid fibril formation in vitro. *Proc Natl Acad Sci U S A* **97**, 8979-8984.

22. Smith, D. P., Jones S, Serpell, L. C., Sunde, M., & Radford, S. E. (2003). A systematic investigation into the effect of protein destabilisation on beta 2-microglobulin amyloid formation. *J Mol Biol* **330**, 943-954.

23. Yee AW, A. U. I. D., Aldeghi, M. A., Blakeley MP, A. U. I. D., Ostermann, A. A., Mas, P. J., Moulin, M., de Sanctis, D., Bowler MW, A. U. I. D., Mueller-Dieckmann, C., Mitchell, E. P., Haertlein, M., de Groot, B. L., Boeri, E. E., & Forsyth, V. T. (2019). A molecular mechanism for transthyretin amyloidogenesis. *Nature Communications* **10**, 925.

24. Liemann, S. & Glockshuber, R. (1999). Influence of amino acid substitutions related to inherited human prion diseases on the thermodynamic stability of the cellular prion protein. *Biochemistry* **38**, 3258-3267.

25. Legname, G., Baskakov, I. V., Nguyen, H. O., Riesner, D., Cohen, F. E., DeArmond, S. J., & Prusiner, S. B. (2004). Synthetic mammalian prions. *Science* **305**, 673-676.

26. Wang, F., Wang, X., Yuan, C. G., & Ma, J. (2010). Generating a prion with bacterially expressed recombinant prion protein. *Science* **327**, 1132-1135.

27. Supattapone, S. (2013). Elucidating the role of cofactors in mammalian prion propagation. *Prion* **8**, 100-105.

28. Hosszu, L. L. P., Baxter, N. J., Jackson, G. S., Power, A., Clarke A, Waltho, J. P., Craven, C. J., & Collinge J (1999). Structural mobility of the human prion protein probed by backbone hydrogen exchange. *Nature Struct. Biol.* **6**, 740-743.

29. Hosszu, L. L. P., Wells, M. A., Jackson, G. S., Jones S, Batchelor, M., Clarke, A. R., Craven, C. J., Waltho, J. P., & Collinge, J. (2005). Definable Equilibrium States in the Folding of Human Prion Protein. *Biochemistry* **44**, 16649-16657.

30. Schatzl, H. M., Da Costa, M., Taylor, L., Cohen, F. E., & Prusiner, S. B. (1995). Prion protein gene variation among primates. *J. Mol. Biol.* **245**, 362-374.

31. Sauve, S., Buijs, D., Gingras, G., & Aubin, Y. (2011). Interactions between the conserved hydrophobic region of the prion protein and dodecylphosphocholine micelles. *J. Biol. Chem.* **287**, 1915-1922.

32. Ott, C. M., Akhavan, A., & Lingappa, V. R. (2007). Specific features of the prion protein transmembrane domain regulate nascent chain orientation. *J Biol. Chem.* **282**, 11163-11171.

33. Hegde, R. S., Mastrianni, J. A., Scott, M. R., DeFea, K. A., Tremblay, P., Torchia, M., DeArmond, S. J., Prusiner, S. B., & Lingappa, V. R. (1998). A transmembrane from of the prion protein in neurodegenerative disease. *Science* **279**, 827-834.

34. Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567-580.

35. Hornemann, S., Von Schroetter, C., Damberger, F. F., & Wuthrich, K. (2009). Prion protein-detergent micelle interactions studied by NMR in solution. *J. Biol. Chem.* **284**, 22713-22721.

36. Abskharon, R., Wang, F., Wohlkonig, A., Ruan, J., Soror, S., Giachin, G., Pardon, E., Zou, W., Legname, G., Ma, J., & Steyaert, J. (2019). Structural evidence for the critical role of the prion protein hydrophobic region in forming an infectious prion. *PLoS Pathog* **15**, e1008139.

37. Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* **11**, 739-756.

38. Ptitsyn, O. B. (1995). Molten globule and protein folding. *Adv. Protein Chem* **47**, 83-229.

39. Wishart, D. S. & Sykes, B. D. (1994). The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J Biomol NMR* **4**, 171-180.

40. Shen, Y. & Bax, A. (2013). Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227-241.

41. Berjanskii, M. V. & Wishart, D. S. (2005). A simple method to predict protein flexibility using secondary chemical shifts. *J Am Chem Soc* **127**, 14970-14971.

42. O'Sullivan, D. B., Jones, C. E., Abdelraheim, S. R., Brazier, M. W., Toms, H., Brown, D. R., & Viles, J. H. (2008). Dynamics of a truncated prion protein, PrP(113-231), from ^{15}N NMR relaxation: Order parameters calculated and slow conformational fluctuations localized to a distinct region. *Protein Sci* **18**, 410-423.

43. Hosszu, L. L. P., Conners, R., Sangar, D., Batchelor, M., Sawyer, E. B., Fisher, S., Cliff, M. J., Hounslow, A. M., McAuley, K., Leo, B. R., Jackson, G. S., Bieschke, J., Waltho, J. P., & Collinge, J. (2020). Structural effects of the highly protective V127 polymorphism on human prion protein. *Communications Biology* **3**, 1-12.

44. Sigurdson, C. J., Nilsson, K. P., Hornemann, S., Manco, G., Fernandez-Borges, N., Schwarz, P., Castilla, J., Wuthrich, K., & Aguzzi, A. (2010). A molecular switch controls interspecies prion disease transmission in mice. *J Clin Invest* **120**, 2590-2599.

45. Christen, B., Damberger, F. F., Perez, D. R., Hornemann, S., & Wuthrich, K. (2013). Structural plasticity of the cellular prion protein and implications in health and disease. *Proc Natl Acad Sci U S A* **110**, 8549-8554.

46. Kurt, T. D., Bett, C., Fernandez-Borges, N., Joshi-Barr, S., Hornemann, S., Rulicke, T., Castilla, J., Wuthrich, K., Aguzzi, A., & Sigurdson, C. J. (2014). Prion Transmission Prevented by Modifying the beta2-alpha2 Loop Structure of Host PrP^C. *J Neurosci* **34**, 1022-1027.

47. Watts, J. C., Giles, K., Patel, S., Oehler, A., DeArmond, S. J., & Prusiner, S. B. (2014). Evidence That Bank Vole PrP Is a Universal Acceptor for Prions. *PLoS Pathog* **10**, e1003990.

48. Kurt, T. D., Jiang, L., Bett, C., Eisenberg, D., & Sigurdson, C. J. (2014). A proposed mechanism for the promotion of prion conversion involving a strictly conserved tyrosine residue in the beta2-alpha2 loop of PrPC. *J. Biol. Chem.* **289**, 10660-10667.

49. Damberger, F. F., Christen, B., Perez, D. R., Hornemann, S., & Wuthrich, K. (2011). Cellular prion protein conformation and function. *Proc Natl Acad Sci U S A* **108**, 17308-17313.

50. Abskharon, R. N., Giachin, G., Wohlkonig, A., Soror, S. H., Pardon, E., Legname, G., & Steyaert, J. (2014). Probing the N-Terminal beta-Sheet Conversion in the Crystal Structure of the Human Prion Protein Bound to a Nanobody. *J Am Chem Soc* **136**, 937-944.

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

52. Freund, S. M., Wong, K. B., & Fersht, A. R. (1996). Initiation sites of protein folding by NMR analysis. *Proc Natl Acad Sci U S A* **93**, 10600-10603.

53. 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.

54. Hosszu, L. L. P., Tattum, M. H., Jones, S., Trevitt, C. R., Wells, M. A., Waltho, J. P., Collinge, J., Jackson, G. S., & Clarke, A. R. (2010). The H187R mutation of the human prion protein induces conversion of recombinant prion protein to the PrP^{Sc}-like form. *Biochemistry* **49**, 8729-8738.

55. Bai, Y., Milne, J. S., Mayne, L., & Englander, S. W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins* **17**, 75-86.

56. Sangar, D., Jack, K., Batchelor, M., Mistry, B., and Bieschke, J. Syntaxin 6 delays prion protein fibril formation and prolongs presence of toxic aggregation intermediates. BioRxiv . 1-1-2022.

57. Baskakov, I., Disterer, P., Breydo, L., Shaw, M., Gill, A., James, W., & Tahiri-Alaoui, A. (2005). The presence of valine at residue 129 in human prion protein accelerates amyloid formation. *FEBS Lett.* **579**, 2589-2596.

58. Trevitt, C. R., Hosszu, L. L. P., Batchelor, M., Panico, S., Terry, C., Nicoll, A. J., Risse, E., Taylor, W. A., Sandberg, M. K., Al Doujaily, H., Linehan, J. M., Saibil, H. R., Scott, D. J., Collinge, J., Waltho, J. P., & Clarke, A. R. (2014). N-terminal Domain of Prion Protein Directs its Oligomeric Association. *J. Biol. Chem.* **289**, 25497-25508.

59. 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.

60. 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 Int Ed Engl* **54**, 7529-7533.

61. Morrissey, M. P. & Shakhnovich, E. I. (1999). Evidence for the role of PrP^{C} helix 1 in the hydrophilic seeding of prion aggregates. *Proc. Natl. Acad. Sci. USA* **96**, 11293-11298.

62. Norstrom, E. M. & Mastrianni, J. A. (2006). The Charge Structure of Helix 1 in the Prion Protein Regulates Conversion to Pathogenic PrPSc. *J Virol.* **80**, 8521-8529.

63. Antonyuk, S. V., Trevitt, C. R., Strange, R. W., Jackson, G. S., Sangar, D., Batchelor, M., Cooper, S., Fraser, C., Jones, S., Georgiou, T., Khalili-Shirazi, A., Clarke, A. R., Hasnain, S. S., & Collinge, J. (2009). Crystal structure of human prion protein bound to a therapeutic antibody. *Proc. Natl. Acad. Sci. USA* **106**, 2554-2558.

64. Mead, S., Khalili-Shirazi, A., Potter, C., Mok, T., Nihat, A., Hyare, H., Canning, S., Schmidt, C., Campbell, T., Darwent, L., Muirhead, N., Ebsworth, N., Hextall, P., Wakeling, M., Linehan, J., Libri, V., Williams, B., Jaunmuktane, Z., Brandner, S., Rudge, P., & Collinge, J. (2022). Prion

protein monoclonal antibody (PRN100) therapy for Creutzfeldt-Jakob disease: evaluation of a first-in-human treatment programme. *Lancet Neurology* **21**, 342-354.

65. Nicholson, E. M., Mo, H., Prusiner, S. B., Cohen, F. E., & Marqusee, S. (2002). Differences between the prion protein and its homolog Doppel: a partially structured state with implications for scrapie formation. *J. Mol. Biol.* **316**, 807-815.

66. Baskakov, I. V., Legname, G., Gryczynski, Z., & Prusiner, S. B. (2004). The peculiar nature of unfolding of the human prion protein. *Protein Sci* **13**, 586-595.

67. Gonzalez-Montalban, N., Makarava, N. F., Savtchenko, R. F., & Baskakov, I. V. (2011). Relationship between conformational stability and amplification efficiency of prions. *Biochemistry* **50**, 7933-7944.

68. Parker, M. J., Spencer, J., & Clarke, A. R. (1995). An integrated kinetic analysis of intermediates and transition states in protein folding reactions. *J Mol Biol* **253**, 771-786.

69. Bax, A. & Grzesiek, S. (1993). Methodological Advances in Protein Nmr. *Accounts of Chemical Research* **26**, 131-138.

70. Reed, M. A., Hounslow, A. M., Sze, K. H., Barsukov, I. G., Hosszu, L. L. P., Clarke, A. R., Craven, C. J., & Waltho, J. P. (2003). Effects of domain dissection on the folding and stability of the 43 kDa protein PGK probed by NMR. *J Mol Biol* **330**, 1189-1201.

71. Bodenhausen, G. & Ruben, D. J. (1980). Natural Abundance N-15 NMR by Enhanced Heteronuclear Spectroscopy. *Chemical Physics Letters* **69**, 185-189.

72. Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S. J., Sorensen, O. W., & Griesinger, C. (1994). A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J Biomol. NMR* **4**, 301-306.

73. Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606-1619.

74. Laue, T. M., Shah, B. D., Ridgeway, T. M., & Pelletier, S. L. (1992). *Computer-aided interpretation of analytical sedimentation data for proteins*. In Analytical Ultracentrifugation in Biochemistry and Polymer Science (Harding, S. E., Horton, J. C., & Rowe, A. J., eds), pp. 90-125, Royal Society of Chemistry, Cambridge.

75. Nielsen, L., Frokjaer, S., Brange, J., Uversky, V. N., & Fink, A. L. (2001). Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry* **40**, 8397-8409.

76. Lipari, G. & Szabo, A. (1982). Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .1. Theory and Range of Validity. *J. Am. Chem. Soc.* **104**, 4546-4559.

Figures and Figure Captions
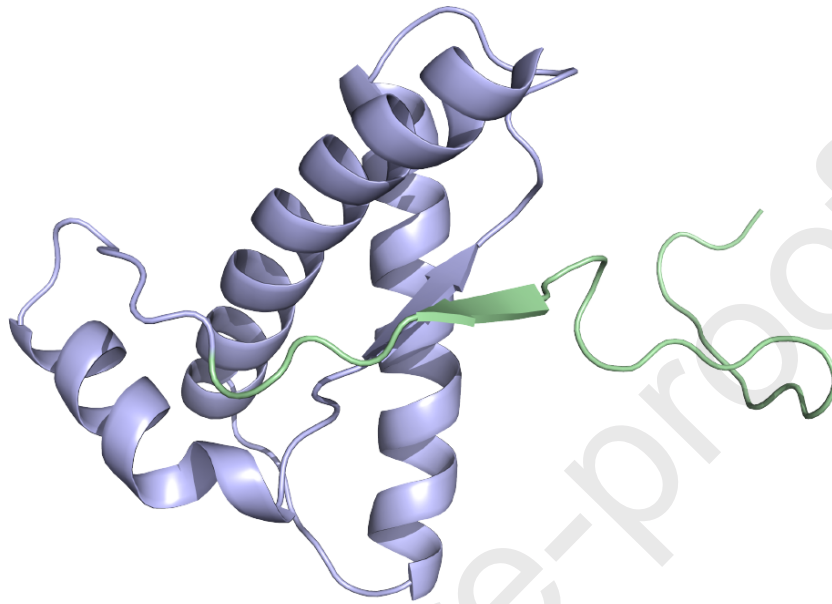
Figure 1

Cartoon representation of PrP[C] secondary structure[16, 17]. Residues 119-226 are displayed, with the N-terminal region deleted in PrP[137] highlighted in green. The first strand of the PrP β-sheet is removed in PrP[137].
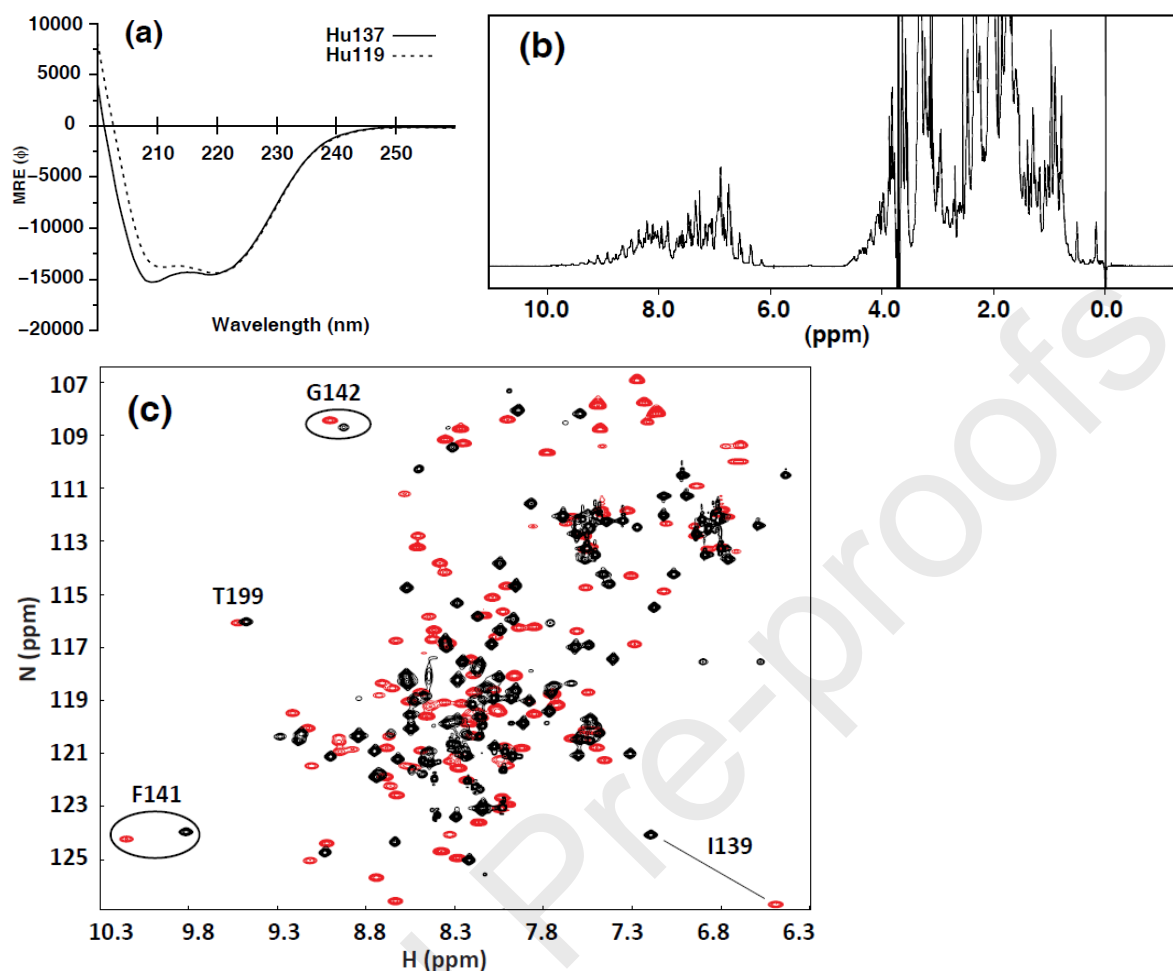
Figure 2

Secondary and tertiary structure of PrP137. *(a)* Far UV-CD spectra of PrP137 and PrP119. Both proteins display CD spectra dominated by α-helical structure, with typical minima at 208 and 222 nm. *(b)* 1D 1H NMR spectrum of PrP137. The NMR spectrum is typical for a fully folded protein, with well-dispersed NMR signals, including resonances at higher field than 0.7 ppm, characteristic of strong tertiary interactions between methyl groups and aromatic rings, indicating extensive tertiary organisation within the protein. *(c)* Overlay of 1H-15N HSQC spectra of PrP137 (black) and PrP119 (red). Signals from regions of the protein in close proximity to the truncation are markedly shifted (I139/F141), in contrast to those remote from the truncation.
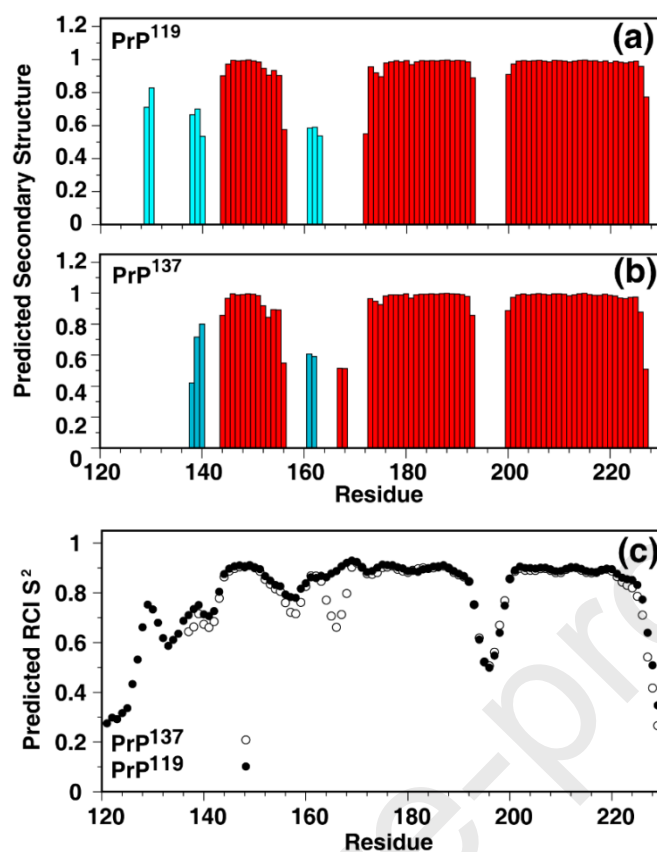
Figure 3

Structural perturbations in PrP^{137} Predicted secondary structure elements for (*a*) PrP^{119} and (*b*) PrP^{137}, as calculated by TALOS-N[40]. Cyan bars indicate extended β-sheet structure, and red bars α-helical regions. The height of the bars reflects the probability of the structure prediction. TALOS-N correctly identifies the PrP^{119} secondary structure elements, and the extended loop spanning residues 138 -140[16, 17, 51, 63] (*c*) RCI (Random Coil Index)[40, 41] prediction of order parameter S^2. S^2 values report on internal sub-nanosecond (ns) motions, and range from 0 for highly flexible to 1 for rigid systems. The RCI method uses secondary backbone shifts (i.e., the difference between observed and reference random coil shifts) to the predict amplitudes of backbone motions[41]. It has certain advantages over the commonly used *Modelfree* analysis[76] of ^{15}N NMR relaxation data in that it does not rely on a model of overall rotation and does not need prior knowledge of the protein's tertiary structure[41]. An increase in mobility at the C-termini of helices 1 and 3 (residues 144-154; 218-230) and in the β2-α2 loop (residues 167-171) of PrP^{137} is predicted, although as the assignments of D167, Y169, S170, N171 could not be determined due to line-broadening of NMR signals beyond detection (Fig. S1), chemical shift data for this region is limited, reducing the probability of the structure prediction.

Figure 4

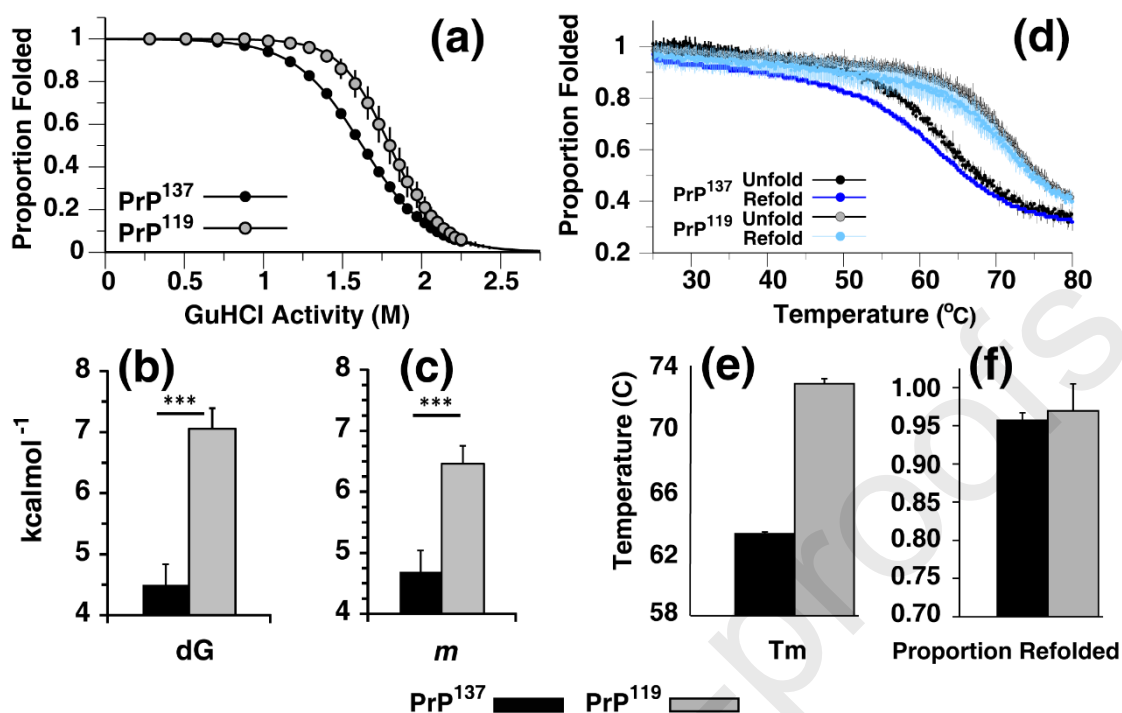*Thermodynamic and thermal stability of PrP^{137} (a)* Denaturant- induced unfolding transition of PrP^{137} and PrP^{119}. The solid line represents the non-linear least-squares fit to a two-state model of folding (Eqn. 1) with pre- and post-transition baseline slopes. The points shown are corrected for the pre- and post-transition baselines. *(b)* Free energy change (ΔG) of the equilibrium unfolding transitions calculated using Eqn. 1. (*c*) Degree of hydrophobic core exposure (*m* values) in PrP^{137} and PrP^{119} derived from the above equilibrium denaturation curve. (*d*) Uncorrected thermally-induced PrP^{137} and PrP^{119} folding transitions, as calculated by Eqn. 3. *(e)* Temperature mid-point (T$_m$) for PrP^{137} and PrP^{119} thermal unfolding transitions. (*f*) Proportion of refolded protein for PrP^{137} and PrP^{119} following thermal denaturation. All folding transitions were monitored by the change in amide CD signal at 222 nm and error bars represent ± standard deviation with ***: P < 0.001 (2 tailed Student's t-test)
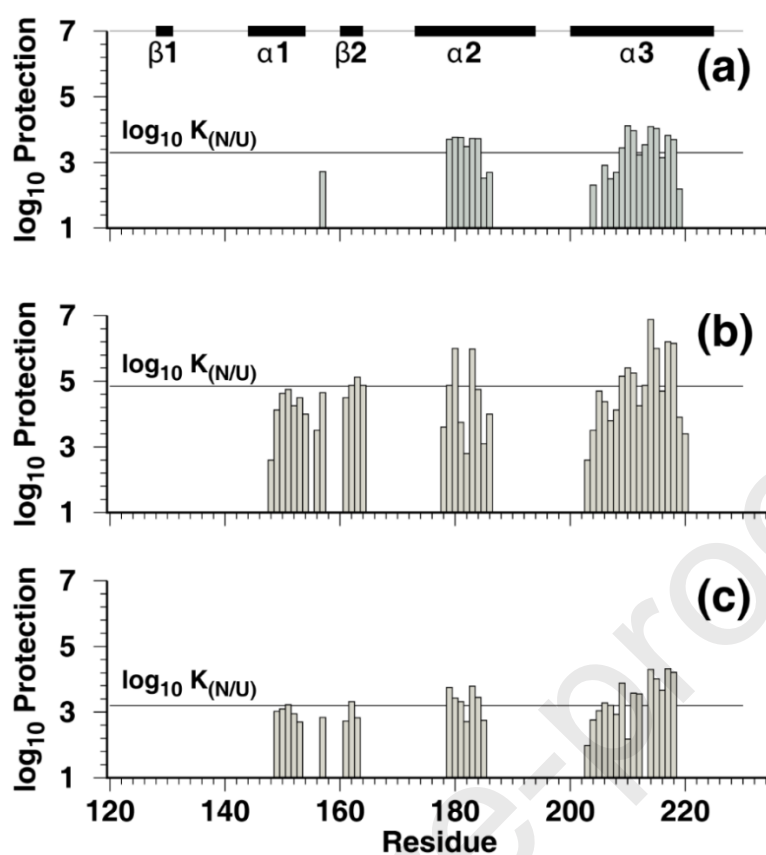
Figure 5

Stability of PrP[137] secondary structure elements as assessed by native-state hydrogen/deuterium exchange. Amide protection factors (k_{ex}/k_{int}) of (a) PrP[137] (at pH 7) (b) PrP spanning residues 121-231 (PrP[121]) (at pH 7)[51] and (c) PrP[119] (at pH 5.5)[17] for those residues with measurable protection determined through H_2O/D_2O exchange. The protection factor corresponding to the equilibrium constant between the native (N) and unfolded (U) states of the respective proteins under each condition is plotted as a line ($\log_{10} K_{(N/U)}$). Regions of highest protection (greater than $K_{(N/U)}$) are associated with a hyper-stable region retained upon unfolding, which consists of a small nucleus surrounding the disulphide bond linking helices 2 and 3[28, 65]. In PrP[119] and PrP[121] the major secondary structure elements (shown at the top of the figure) apart from the first β-strand, have stabilities equivalent to the overall stability of the protein, and therefore the protein has to fully unfold for these regions of the protein to exchange with the solvent. In PrP[137], the second strand of the PrP β-sheet (residues 160-164) does not display measurable protection, consistent with the loss of hydrogen-bonding with the first β-strand. α-helix 1 also displays markedly reduced protection in PrP[137] compared to PrP[119] and PrP[121], indicating that it has been selectively destabilised, relative to the other α-helices. This is contrast to the reduced protection factors for all secondary structure

33

elements associated with a general destabilisation of PrP through the protonation of key residues at acidic pH ((b) and (c))[53, 54].
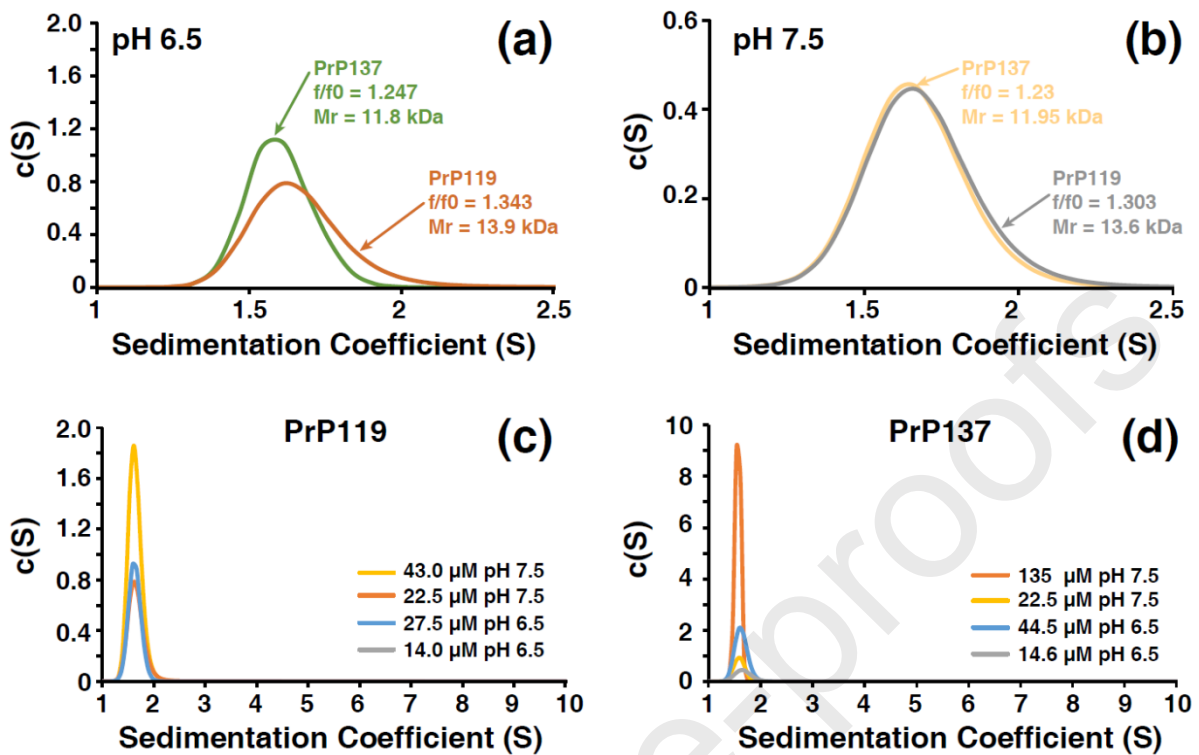
Figure 6

Oligomeric state of PrP[137] and PrP[119] determined by sedimentation velocity AUC. Expanded (*a/b*) and full (*c/d*) sedimentation coefficient distributions of PrP[119] and PrP[137]. (*a/b*) Both PrP constructs sediment with derived molecular masses equivalent to the expected molecular mass of their monomers at both pH 6.5 and pH 7.5. (*c/d*) Protein concentration-dependence for aggregation. Both PrP[119] and PrP[137] remain monomeric at pH 6.5 and pH 7.5 over a range of protein concentrations, with no indication of the population of oligomeric states.
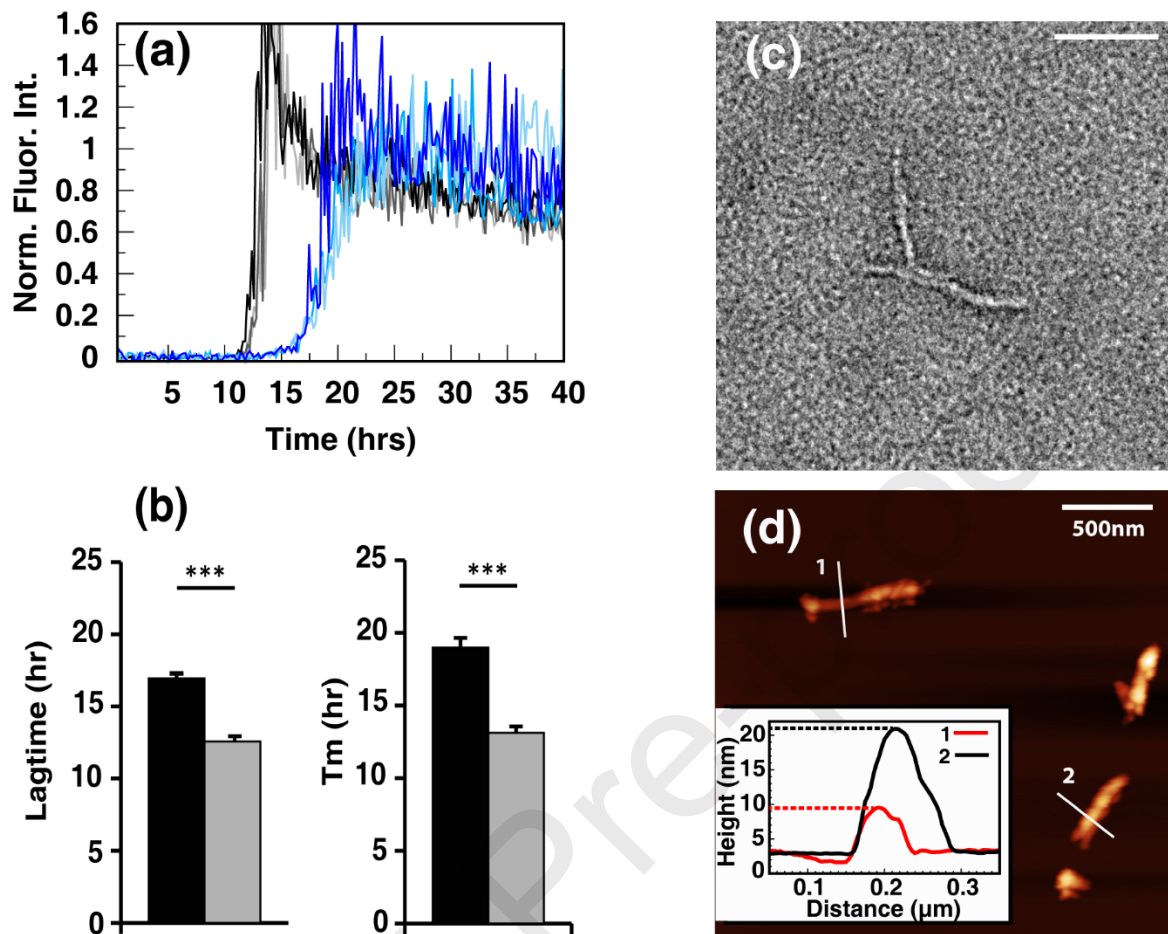
Figure 7

Effect of the deletion of the unstructured PrP N-terminus (including the CHR) and first β-strand on PrP fibril formation (a) Kinetic traces of ThT fluorescence (PrP^{137} – black/grey; PrP^{23} – sky blue/ royal blue). Three replicate traces for each protein are shown. (b) Fibrillogenesis of PrP^{23} (black) occurs with significantly longer mean half- and lagtimes in comparison to PrP^{137} (light grey) (P ≤ 0.001, paired t-test). The data shown are the mean readings from three replicates ±SD. (c) Negative stain transmission electron microscopy images of PrP^{137} fibrils (Scale Bar: 50 nm). (d) AFM images of PrP^{137} fibrils (Scale Bar: 500 nm). (Inset) Height profiles of selected fibrils, as measured by *Gwyddion* 2.6.1.
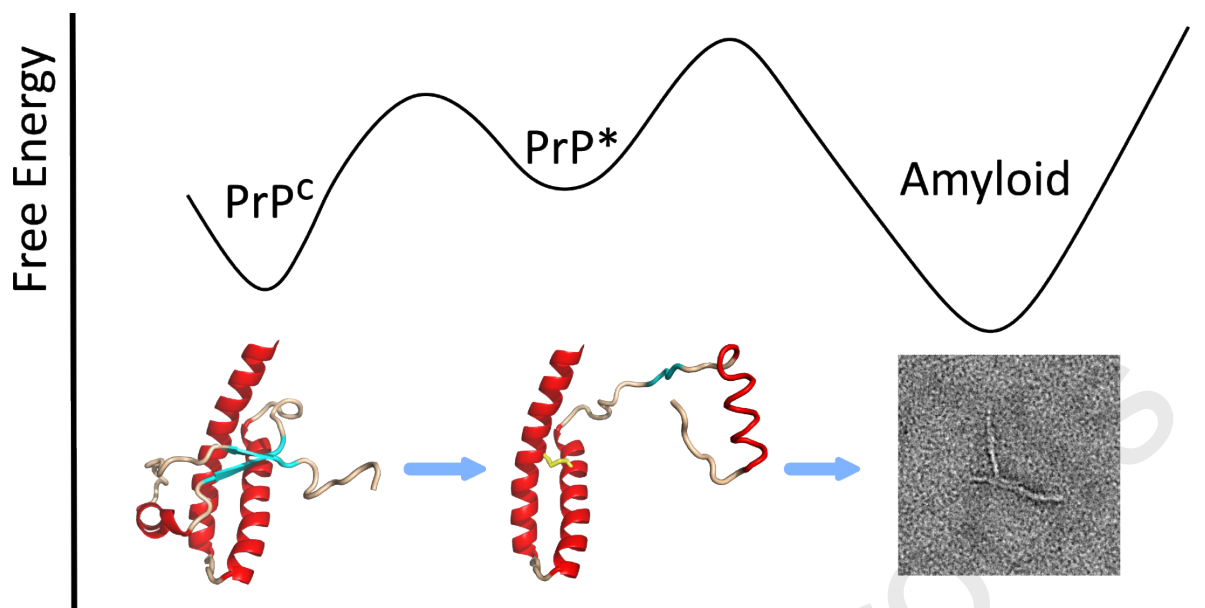
LLPH, JW, JB: Conceptualization; **LLPH, AMH, ER, JW, JB:** Methodology; **LLPH, DS, MB, AMH, ER:** Investigation, **LLPH, JB, JPW, JC:** Writing- Reviewing and Editing; **JB, JC:** Project administration, Funding acquisition.

Highlights

- A distinct "open" form of human PrP, in equilibrium with its native state was identified.
- This native-like conformer has increased solvent exposure and fibrilises more readily than the native state.
- The conserved hydrophobic region of PrP, which associates with the ER membrane determines the stability of this state.
- These data suggest a stepwise folding transition, which is initiated by a conformational switch to this "open" form of PrP.

Declaration of interests

☐ 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.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

J.C. is a Director and shareholder of D-Gen Limited, an academic spin-out company working in the field of prion disease diagnosis, decontamination, and therapeutics. The other authors declare no competing interests.

☒The authors declare the following financial interests/personal relationships which may be