

Lysine hydroxylation and O-glycosylation in the globular, C-terminal region of mammalian-expressed, recombinant PrP

Mark A. Ritchie^{a,1}, Lawrence G. Hunt^a, Andrew C. Gill^{a,b,*}

^a The Institute for Animal Health, Compton, Newbury, Berkshire RG20 7NN, UK

^b Neurobiology Division, The Roslin Institute and Royal (Dick) School of Veterinary Sciences, Easter Bush Veterinary Centre, Roslin, Edinburgh EH25 9RG, UK

ARTICLE INFO

Article history:

Received 15 June 2012

Received in revised form 8 August 2012

Accepted 8 August 2012

Available online 18 August 2012

Keywords:

Prion protein

Post-translational modification

Hydroxylysine

O-linked glycosylation

Rigid loop

ABSTRACT

Conversion of PrP^C, the prion protein, to a conformationally altered isoform, PrP^{Sc}, is the major pathogenic event in the transmissible spongiform encephalopathies, a family of neurodegenerative diseases including bovine spongiform encephalopathy, Creutzfeldt-Jakob disease and scrapie. Known post-translational modifications to the protein include disulfide bridge formation, addition of a membrane anchor and N-linked glycosylation. We have previously identified the pro-collagen-like hydroxylation of proline 44 in a murine, recombinant prion protein expressed in Chinese hamster ovary cells and herein report the identification of a second pro-collagen-like modification in this protein. In a proportion of the molecules, Lys193, within the C-terminal, folded domain of the protein, is specifically modified to hydroxylysine with subsequent addition of two hexose units, assumed to be the collagen-like disaccharide modifier Gal-Glu. Proof of the existence of these modifications has been obtained by means of tandem mass spectrometry and Edman degradation. Molecular dynamics simulations show that these modifications lead to a pronounced stabilising effect on the β_2 - α_2 loop, a region of PrP crucial for the disease-associated conversion. If present *in vivo*, these modifications may have important implications in PrP structure, interactions with ligands or may modulate PrP aggregation.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The transmissible spongiform encephalopathies (TSEs) are a group of progressive, neurodegenerative disorders that may be acquired genetically, arise spontaneously or may be transmitted by ingestion or inoculation of diseased tissues. This family of fatal diseases includes bovine spongiform encephalopathy, scrapie of sheep and Creutzfeldt-Jakob disease in humans [1]. The fundamental molecular event in pathogenesis of all TSE diseases is the conversion of a normal, cell-surface glycoprotein (PrP^C – cellular prion protein) to a partially protease-resistant isoform (PrP^{Sc} – prion protein scrapie isoform) [2]. PrP^{Sc} is deposited in insoluble aggregates and this form of the protein is partially resistant to

degradation with proteases. Since there appear to be no differences between the primary structures of PrP^{Sc} and PrP^C, the conversion is believed to be purely conformational [3].

The prion hypothesis suggests that PrP^{Sc} is the only or, at least, major component of the infectious agent and the conformational change of host PrP^C is believed to proceed catalytically through the use of PrP^{Sc} as a template [2]. Such a mechanism requires the interaction of PrP^{Sc} and PrP^C, either directly or through intermediary molecules such as polyanions or lipids, which may also act as cofactors aiding conversion [4–7]. The advent of several seeded misfolding assays, in which PrP^{Sc} can convert recombinant PrP (recPrP) to a protease-resistant isoform *in vitro* [8–11], suggests that a mechanism involving templated, conformational conversion occurs *in vivo* and that disease is indeed passed by ingestion or inoculation of PrP^{Sc}-containing tissues. However, various sporadic forms of TSE disease exist in humans [12], and possibly in cattle [13] and sheep [14], and it is unclear what factors initiate the conformational cascade in such diseases, in the absence of an exogenous infectious agent. Whilst sporadic diseases are believed to be caused by somatic mutations to the *PRNP* gene in dividing cells, which produces PrP^C protein that can spontaneously misfold, it remains possible that a small proportion of PrP^C molecules may undergo covalent post-translational modifications (PTMs) that cause misfolding and drive the disease process.

PrP^C has between 208 and 220 residues, depending on species, and the primary structure is highly homologous across mammals. A

Abbreviations: CAD, collisionally activated dissociation; CHO, Chinese hamster ovary; GPI, glycosylphosphatidylinositol; MS, mass spectrometry; MS/MS, tandem mass spectrometry; *PRNP*, the gene encoding the prion protein; PrP^C, cellular prion protein; PrP^{Sc}, scrapie isoform of the prion protein; PrP GPI–, prion protein lacking the GPI anchor; PTM, post translational modification; recPrP, recombinant prion protein; TSE, transmissible spongiform encephalopathy.

* Corresponding author at: Neurobiology Division, The Roslin Institute and Royal (Dick) School of Veterinary Sciences, Easter Bush Veterinary Centre, Roslin, Edinburgh EH25 9RG, UK. Tel.: +44 0131 651 9121; fax: +44 0131 651 9105.

E-mail address: andrew.gill@roslin.ed.ac.uk (A.C. Gill).

¹ Current address: Waters Pacific Pte Ltd, 1 Science Park Road, #02-08 The Capricorn, Singapore Science Park II, Singapore 117528, Singapore.

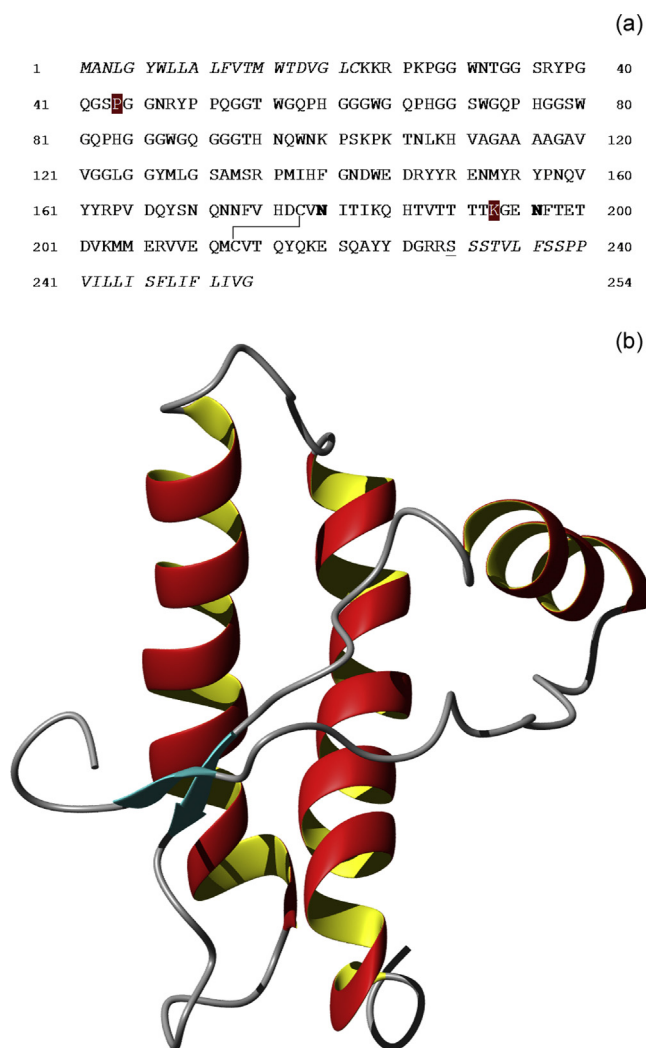


Fig. 1. (a) The full sequence of murine PrP^C. N- and C-terminal signal peptides are shown in italics and the site of attachment of the GPI membrane anchor is underlined. A single disulfide bond links cystines 178 and 213. Sites of N-linked glycosylation are in bold whilst the sites of proline hydroxylation and lysine hydroxylation (this work) are highlighted with a dark background. (b) Ribbon diagram of the structured, C-terminal domain of murine PrP^C (residues 124–226). The structure was based on that of murine PrP determined by Riek et al. [24] (PDB code 1AG2) and rendered using MolMol and PovRay.

number of physiological PTMs have been identified, which appear common to all mammalian prion proteins [15,16]. The sequence of the murine prion protein is shown in Fig. 1(a), on which sites of PTMs are highlighted. These include cleavage of N and C-terminal signal sequences leaving a mature protein spanning residues 23–230 (murine numbering, which is used throughout) followed by attachment of a glycosylphosphatidylinositol (GPI) membrane anchor at the mature C-terminus. In the ER/golgi complex, further modifications involve addition of two N-linked carbohydrate chains to Asn180 and Asn196 [17,18] and formation of a single, internal disulfide bond. We have also identified the pro-collagen-like hydroxylation of proline 44 in both recPrP expressed in Chinese hamster ovary (CHO) cells and, at least at low levels, in PrP^{Sc} purified from the brains of scrapie infected mice [19], a modification that demonstrates that the N-terminal region of PrP can adopt polyproline II structure [20]. There have been reports of modifications to arginine residues in the N-terminal region, although initial suggestions that these are converted to citrulline remain to be confirmed [21,22]. There are further reports of modifications to arginine or lysine residues in the N-terminal region by advanced glycation end

products [23]. The effects of N-terminal amino acid modifications on the structure of PrP have largely not been studied.

The N-terminal region of PrP is rather flexible but, by contrast, the C-terminal domain of PrP^C possesses typical globular structure. NMR studies of recombinant PrP expressed in bacteria have shown that the C-terminal region is composed of a 3-helix bundle with two short β -strands, depicted in Fig. 1(b) [24]. Conversion of PrP^C to PrP^{Sc} is associated with an increase in the β -sheet content of the protein involving at least part [3], or potentially all of the helical region of the protein [25]. Thus, post translational modifications in this region of the protein may be more likely to impact on PrP folding/misfolding. There have been reports of oxidation of specific methionine residues [26] and a suggestion that asparagine/aspartate residues may be altered in PrP^{Sc} [27,28]. In proof of principle studies, it has also been shown that conditions that cause oxidative or nitrative modifications, or that deaminate PrP, can lead to misfolding [29,30] and it is clear that any PTMs that can be detected in the C-terminal domain of PrP are candidates to impact adversely on protein folding.

Following our previous identification of 4-hydroxyproline in the N-terminal region of mammalian-expressed recPrP and in PrP^{Sc} purified from mouse brains [19], herein we report the discovery of a second novel modification to the recPrP. Lys193 was found to be specifically converted to δ -hydroxylysine with subsequent addition of a di-hexose unit. Both Edman degradation and mass spectrometry have been used to characterise these modifications by comparison to complement component C1q, a protein known to contain similar modifications in its collagenous domain, which have previously been shown to constitute the addition of the disaccharide galactose-glucose to hydroxylysines. Molecular dynamics simulations show that disaccharide modification of Lys193 of PrP leads to little effect on the motion in the loop around the modification, but that the loop joining β -sheet2 with α -helix2 becomes more stable. The β_2 - α_2 loop region has been postulated to be implicitly involved in modulating the susceptibility of PrP to misfold, hence modification to Lys193 may be a candidate post-translational modification initiating PrP misfolding.

2. Methods

2.1. Protein production, purification and digestion

The expression, production and purification of aglycosyl recombinant PrP lacking its GPI anchor is as described previously [31]. Briefly, CHO cells were grown on Cytodex 1 microcarriers (Pharmacia) in Techne spinner flasks for 3 days in serum-containing medium, followed by a recPrP production phase in serum-free medium. The protein accumulated in the culture medium and, after 3 days production, soluble protein was purified by means of cation exchange chromatography eluting with a sodium chloride gradient. Fractions containing recPrP, as determined by Western blotting techniques, were pooled and further purified by immobilised metal affinity chromatography, eluting by means of an imidazole gradient. The final fractions contained mature length recPrP, in addition to a 9 kDa N-terminal fragment resulting from a single proteolytic cleavage that has previously been characterised [31].

CHO cell recPrP in 6 M urea, pH 8.5, was reduced with 5 mM dithiothreitol and alkylated with 20 mM iodoacetic acid then was dialysed against 3 M urea, 200 mM ammonium bicarbonate pH 9.0. 10% (w/w) trypsin was added and the recPrP was digested overnight. For selected analyses, detailed in the results section, recPrP was partially digested by the addition of 1% (w/w) trypsin without prior reduction or alkylation. Complement component C1q from human serum (Sigma) was dissolved to 1 mg/ml in 0.2 M ammonium bicarbonate pH 8.5, 2 M urea and

was digested overnight with 5% (w/w) trypsin without prior reduction/alkylation.

2.2. Mass spectrometry

Capillary HPLC columns of 180 μm internal diameter were constructed using a method similar to that reported by Tong et al. [32]. Simply, fused silica (180 μm i.d.), was slurry packed with Hichrom C₁₈ reversed phase packing (3.5 μm beads, 150 Å pore size) and was terminated with a stainless steel screen (16th inch. diameter, 2 μm pore size, Valco). Samples were injected, *via* an external loop, low dispersion injector (Valco) onto a home-made desalting trap and were washed, eluted onto the capillary column and separated by linear gradient of increasing solvent B (where solvent A was 95:5 H₂O:acetonitrile with 0.05% (v/v) trifluoroacetic acid and solvent B was 5:95 H₂O:acetonitrile with 0.05% (v/v) trifluoroacetic acid). A dual syringe Microgradient HPLC pump (Applied Biosystems) provided a flow rate of 50 $\mu\text{l}/\text{min}$, which was split to approximately 1 $\mu\text{l}/\text{min}$. The outlet of the column was connected, *via* a UV detector (214 nm) equipped with a U-Z View flow cell (LC Packings, Amsterdam, Netherlands), to the Z-spray source of a 'Quattro II' tandem quadrupole mass spectrometer (Waters, Altrincham, UK) operated in continuous flow nanospray mode. The mass spectrometer was scanned from m/z 350 to 2200 (4 s per scan) and the cone voltage was ramped from 30 to 75 V over this range to aid detection of ions having high mass-to-charge ratios.

For tandem mass spectrometry experiments, peptides isolated by semi-preparative HPLC column (2.1 mm id, 15 cm length, 3.5 μm bead size, 150 Å pore size, Hichrom, UK) were lyophilised and reconstituted in 1:1 methanol:water with 0.1% (v/v) formic acid. 1 μl aliquots were loaded into the end of a platinum coated, pulled, borosilicate capillary to which a voltage of approximately 1 kV was applied, resulting in spraying flow rates of around 20 nl/min. The hexapole collision cell of the Quattro II mass spectrometer contained argon gas at a pressure of 4×10^{-3} mbar and collision energies were varied between 20 and 30 V to give optimum daughter ion abundance. MS2 was scanned between m/z 50 and 2500 (20 s scan time) and at least 50 scans were summed to give the final spectra.

2.3. Edman degradation

Fractions containing peptides of interest were identified by ESI-MS by use of platinum-coated borosilicate glass capillaries (see above). Edman degradation was performed by means of an Applied Biosystems automated protein sequencer. PTH amino acids were separated using the manufacturer's PTH C₁₈ column and a 140C HPLC system.

2.4. Molecular dynamics simulations

The NMR structure of the murine prion protein, residues 124–226, was downloaded from the protein data bank (PDB ID 1AG2) [24]. This structure was used as a starting point for simulations by use of the AMBER suite of molecular dynamics software (version 7) [33]. We used the Amber99 all atom force field supplemented with Glycam04 parameters for carbohydrates [34]. A pdb file for the Gal-Glu carbohydrate was constructed by use of the on-line glycoprotein builder at the Glycam web site (<http://glycam.ccruc.uga.edu/ccrc/>) and a merged pdb file was created using Swiss PDB Viewer (DeepView) [35]. During creation of molecular topology the carbohydrate was attached to the δ -carbon of Lys193 of the murine PrP by use of the AMBER software, as either 5R or 5S stereoisomer. This resulted in artificially large bond lengths between the protein and the carbohydrate which were removed during energy minimisation and equilibration. Wildtype murine

PrP was also subjected to molecular dynamics simulations starting from the 1AG2 pdb coordinate file with no adjustments.

Starting structures were extensively energy minimised to remove energy hot spots and, in the case of the modified structure, to reduce the bond length between the carbohydrate and the protein; bonds to hydrogen atoms were fixed during this procedure. Structures were then equilibrated *in silico* to 37 °C and molecular dynamics was carried out at 37 °C for 10 ns for each structure. A Born approximation of implicit solvent was used to expedite simulations [36] whilst a simulation of 200 mM salt was included to maintain buffering. Coordinates were output every 10 ps during each simulation and every structure was analysed for root-mean-squared (RMS) deviation from starting and mean structures, for hydrogen bonds and for change in phi-psi angles. Molecular trajectories were analysed using AMBER software or with VMD analysis software [37] and figures were built using VMD or MOLMOL [38].

3. Results

3.1. Mammalian-expressed PrP contains a post-translational modification in the C-terminal region

In our laboratory, a long-term strategy for the production of prion protein for structural analysis has been the over-expression of murine recPrP by stably transfected CHO cell lines. The recPrP produced has been mutated, at the DNA level, to remove the sites of N-linked glycosylation (Asn180Thr and Asn196Thr) and to prevent addition of the glycosylphosphatidylinositol membrane anchor by removing the C-terminal signal sequence. The N-terminal signal sequence is retained and the protein is thus secreted from the cells and it accumulates in the culture medium, from which it can be purified to allow both structural and functional analyses.

After expression of murine recPrP in CHO cells, several proteins can be detected by anti-PrP antibodies by Western blotting [31]. A proportion of the full length protein molecules, mass ~23 kDa, are cleaved specifically by the action of an unknown protease to result in polypeptides of mass ~9 kDa and ~14 kDa equating to the N-terminal and C-terminal portions of the protein respectively. Previously published mass spectrometric analyses showed that this cleavage corresponds exactly to that taking place constitutively *in vivo* and occurs between residues 109 and 110 [31]. During these studies, the N-terminal, 9 kDa fragment co-purified with full length recPrP by Ni-IMAC and cation exchange chromatography, owing to similar charge and metal ion binding properties. Our previous analysis of the N-terminal fragment demonstrated the presence of a novel post-translational modification [19]: proline 44 can be specifically modified to 4-hydroxyproline in an enzymatic reaction requiring both a consensus sequence and polyproline II structure in the substrate (see Fig. 1a). The C-terminal, 14 kDa fragment was lost during our purification protocol, however, limited trypsin digestion of full length recPrP produced an analogous polypeptide spanning residues 110–228, *i.e.*, from the *in vivo* cleavage site to within 2 residues of the C-terminus of the mature length protein. This allows the characterisation of the C-terminal domain of CHO cell-expressed recPrP by online HPLC-MS.

Fig. 2(a) shows the mass spectrum of mature length CHO cell recPrP and Fig. 2(b) shows the mass spectrum of the C-terminal fragment of approximately 13.5 kDa generated by partial tryptic digestion. In addition to species having the expected mass, both spectra indicate the presence of various forms of the proteins with masses higher than those predicted from the amino acid sequences. In the mature length molecule, some of this heterogeneity may be rationalised by the formation of 4-hydroxyproline at codon 44 in a proportion of the molecules (~70%) [19]. The mass spectrum of the 13.5 kDa, C-terminal fragment, spanning residues 110–228,

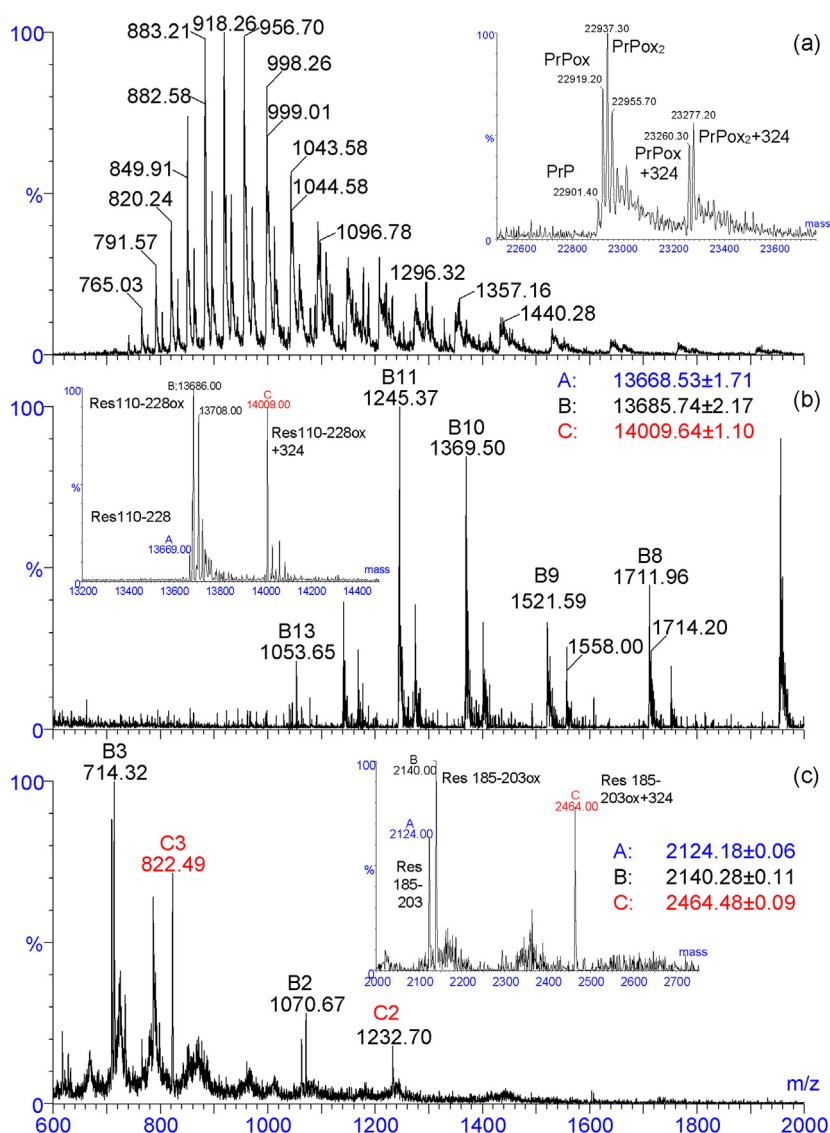


Fig. 2. (a) Mass spectrum and, inset, deconvoluted mass spectrum of mature length, CHO cell-expressed recPrP. The theoretical mass of the unmodified protein is 22904.2 Da. (b) Mass spectrum and, inset, deconvoluted mass spectrum of the polypeptide spanning residues 110–228 generated by partial trypsin digestion of recPrP. The theoretical mass of this protein is 13668.2 Da. In addition to a weak signal corresponding to the correct mass, proteins heavier by 16 and 340 Da are evident. The peak in the deconvoluted spectrum of mass 13,708 Da corresponds to a sodium adduct of the hydroxylated protein. (c) Mass spectrum and, inset, deconvoluted mass spectrum of the tryptic peptide from recPrP spanning residues 185–203. The theoretical mass of this peptide is 2124.29 Da. Close eluting peptides a further 16 and 340 Da higher in mass can also be seen.

however, also reveals the presence of other post-translational modifications. In addition to signals corresponding to a protein of the correct molecular weight for this fragment, a species is present that is 16 Da higher in mass suggesting oxidation of another amino acid. A third species is also evident that is 340 Da higher in mass than that predicted. The absence of a fourth species a further 16 Da higher suggests that this third species may arise through the addition of 324 Da to the protein having the oxidised amino acid, or that addition of a species of 340 Da may occur instead of oxidation at the same site. We noted these modifications in our previous mass spectrometric investigations, without being able fully to characterise them [31]. Other higher molecular weight species that are evident on the deconvoluted mass spectra of Fig. 2(a) and (b) correspond to sodium or potassium adducts.

In order to locate the site(s) of the novel C-terminal modifications, mature length recPrP was reduced and alkylated, dialysed to remove excess reagents and was extensively digested with trypsin. A sample of the resulting peptides was analysed by capillary HPLC–MS. We observed various peptides having masses

corresponding to those predicted from the theoretical structure, in addition to 2 peptides that had co-eluting species 16 Da higher in mass of significant abundance. One of these corresponded to the peptide containing the previously identified 4-hydroxyproline residue (data not shown). The other had a mass of 2124.18 Da, in close agreement to the predicted tryptic peptide spanning residues 185–203, of mass 2124.29 Da, which is generated through a missed tryptic cleavage at Lys193. Fig. 2(c) shows the summed mass spectrum containing signals corresponding to the unmodified peptide, its oxidised equivalent at 16 Da higher in mass, as well as a third peptide eluting from the HPLC column 1 min earlier, a further 324 Da higher in mass. Thus, the modified amino acid occurs within the sequence of 185–203 of recPrP.

3.2. Lys193 is specifically hydroxylated with subsequent disaccharide addition

The recPrP tryptic digest was separated on a semi-preparative reversed phase column and fractions were collected. Fractions

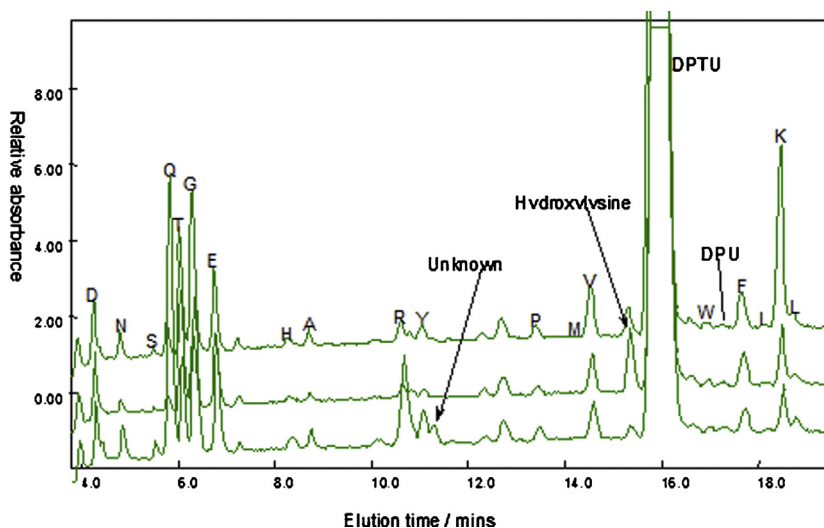


Fig. 3. Edman sequencing chromatograms of cycle 9 from the recPrP tryptic peptide spanning residues 185–203 (upper trace) and the peptides a further 16 Da (middle trace) and 340 Da (lower trace) higher in mass. In the upper trace a major peak corresponding to PTH–lysine is evident. In the middle trace a peak eluting at around 15.5 min, and corresponding to PTH– δ -hydroxylysine can be seen. In the lower trace, an unknown peak is evident eluting just after PTH–tyrosine.

containing the three peptides of interest – Gln185–Lys203 and the two modified equivalents – were identified by offline nanospray mass spectrometry. The peptides were analysed by N-terminal Edman sequencing, which demonstrated that each had the same N-terminal peptide sequence and confirmed the identities of the peptides. At cycle 9, corresponding to Lys193, however, there were significant differences in the Edman cycles. The chromatograms for this cycle in each of the peptides are shown in Fig. 3. The upper trace shows cycle 9 from the unmodified peptide where a major peak corresponding to phenylthiohydantoin (PTH)–lysine can be seen. Although there are significant background peaks from other amino acids, subtraction of data from the previous cycle clearly indicates the presence of lysine in this cycle (data not shown). In the middle trace in Fig. 3, the peak for PTH–lysine is significantly reduced in size whilst a different peak, eluting at 15.5 min, just prior to DPTU, is present. This chromatogram is from the sequencing experiment on the oxidised peptide ($M + 16$ Da) and, by comparison to literature values provided by the PTH column manufacture, the new peak is seen to correspond to the PTH derivative of δ -hydroxylysine. This peak is almost coincident with an unknown peak, present in all cycles, presumably the result of side reactions during the labelling step. The lower trace in Fig. 3 shows the Edman chromatogram from cycle 9 of the species 340 Da heavier in mass than the unmodified peptide. A different peak of low intensity, eluting at around 11 min just after PTH–tyrosine, is present in this chromatogram but does not correspond to any known PTH derivative of modified lysine by comparison with literature values [39]. In collagen and proteins possessing collagenous domains, a common modification to δ -hydroxylysine residues is addition of the disaccharide Gal–Glu [40], which would cause the required mass shift of 324 Da. However, the elution position of PTH– δ -hydroxylysine with covalent addition of the disaccharide Gal–Glu, is not known since it is often not seen during Edman degradation [41,42].

To provide further proof for the 2 different modifications being present on Lys193, the 3 isolated peptides were subjected to collisionally activated dissociation (CAD) experiments by means of tandem mass spectrometry (MS/MS). The three CAD spectra, corresponding to peptides that were unmodified, 16 Da and 340 Da higher in mass than predicted, are shown in Fig. 4(a)–(c) respectively. In each case the triply charged ion was used as the precursor ion and is indicated on the spectra with an asterisk. The CAD spectrum of the unmodified peptide has multiple neutral losses of water from the precursor ion, presumably from the threonine

sidechains. Also evident are both singly and doubly charged y-type fragment ions allowing partial sequence information to be derived and confirming the identity of the peptide. The CAD spectrum of the hydroxylated equivalent is similar, but crucially peaks

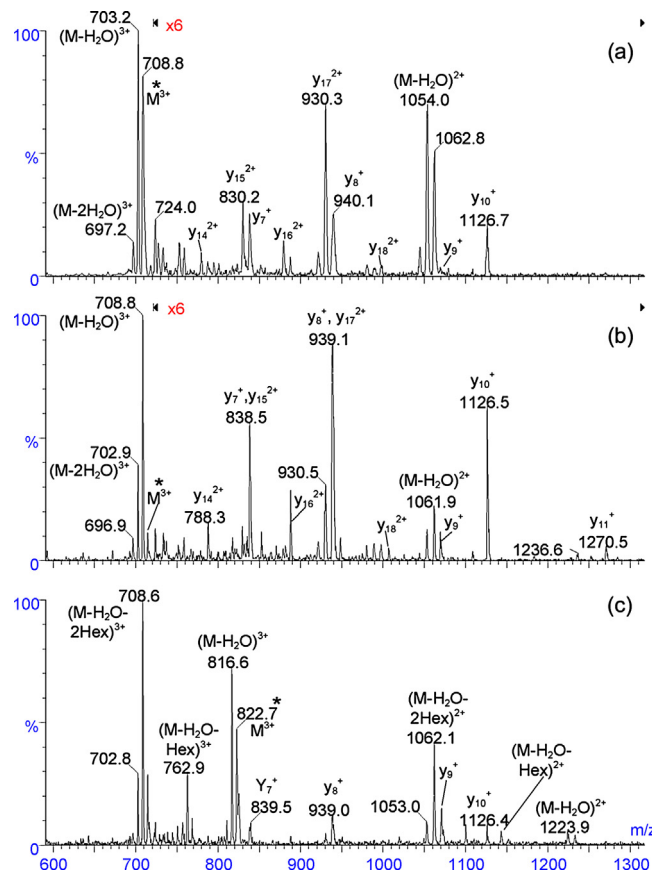


Fig. 4. CAD mass spectra of peptides of sequence QHTVTTTTJGETFTETDVK. (a) J = lysine, precursor ion is 708.0 (b) J = hydroxylysine, precursor ion is 713.4 (c) J = O-glycosylated hydroxylysine, precursor ion is 821.4. The precursor ion is labelled with an asterisk on each spectrum. Y-type fragment ions are seen in all spectra, in addition to loss of water. In (c) loss of 2 consecutive hexose residues can be seen in both doubly and triply charge states.

corresponding to $y_{18}^{2+} - y_{14}^{2+}$ have shifted upwards in mass by 16 Da. y_{17}^{2+} and y_{15}^{2+} now appear in the same place as y_8^+ and y_7^+ respectively, accounting for the relative increase in the size of these signals. A peak at m/z 1270 is assigned as y_{11}^+ and corresponds to fragmentation at the N-terminal side of the hydroxylysine residue. This peak, in conjunction with the peak at 1126.5, assigned as y_{10}^+ , proves that the hydroxylation is at Lys193.

The CAD spectrum of the peptide with an additional 340 Da shows only limited sequence-specific fragmentation. Instead, the spectrum is dominated by the loss of both water and 2 sequential losses of mass 162 Da. This is the mass of a hexose unit, and these signals, occurring as both neutral losses ($3+$ charge state) and as singly charged fragments, resulting in $2+$ peptide fragment ions, demonstrate the presence of a disaccharide on this peptide. The position of peaks assigned as $y_7^+ - y_{10}^+$ remain unchanged relative to the CAD spectrum of the unmodified peptide and indicate that the disaccharide is attached N-terminally to Glycine 194. The absence of the y_{11}^+ peak at m/z 1270 suggests that Lys193 is also the site of attachment of the disaccharide.

3.3. Comparison with complement-component C1q demonstrates lysine glycosides are present in CHO-cell recPrP

In order to confirm our Edman degradation and MS assignments for lysine modifications, we sought a standard protein that carries similar modifications. Complement component C1q exists *in vivo* as a multimeric protein complex held together by triple helices formed by the N-terminal collagenous regions of the protein. These regions have previously been shown to contain a number of hydroxylysine residues with covalently attached disaccharides [43,44]. Thus, to provide a control sample for comparison of mass spectrometry and Edman degradation data, C1q from human serum was digested with trypsin and peptides containing lysine glycosides were isolated. The C1q subunit A-derived peptide spanning residues 95–110 contains 2 known sites of hydroxylysine formation and measurement of the mass of this peptide by mass spectrometry demonstrated that hydroxylation and glycosylation had occurred at both sites. This peptide was subjected to both Edman degradation and MS/MS experiments. Fig. 5(a) shows the chromatogram of cycle 9, the second glycosylated hydroxylysine residue, of the Edman degradation and shows clearly a peak eluting at around 11 min, just after PTH tyrosine. No other peak is seen that cannot be accounted for by contamination of the chromatogram by background levels of amino acids. This peak has the same retention time as that of the 'unknown residue' in the PrP-derived peptide (Fig. 3(c)) and strongly suggests that this peak is caused by the PTH derivative of hydroxylysine modified with the disaccharide galactose-glucose. The relatively weak signals in both the C1q peptide and the PrP peptide is probably caused by chemical degradation of the amino acid under the acidic conditions of the cleavage reactions and, as mentioned previously, glycosylated hydroxylysine is known to be poorly covered by Edman degradation. We also analysed the C1q-derived peptide by mass spectrometric methods and Fig. 5(b) shows the CAD spectrum of the C1q peptide. As with the PrP peptide (Fig. 4(c)), losses corresponding to hexose residues can clearly be seen. In this case, consecutive losses of 4 hexose units are seen, rationalised by the presence of 2 glycosylated hydroxylysine residues. In any case, both Edman degradation and mass spectrometric analysis of this C1q peptide is analogous to the analyses of the PrP-derived peptide providing conclusive evidence that Lys193 of recPrP carries both δ -hydroxylation in a proportion of molecules and subsequent addition of a disaccharide, probably galactose-glucose, in a further proportion.

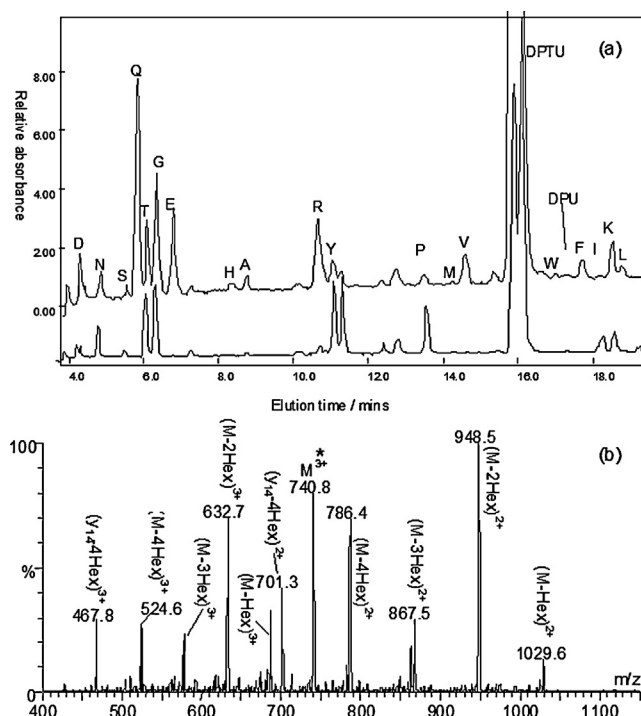


Fig. 5. (a) Overlay of cycle 9 from the putative glycosylated recPrP tryptic peptide (upper trace) and cycle 9 from the known O-linked glycosylated peptide from tryptic digestion of complement component C1q. Despite contamination from various amino acids, including tyrosine, a peak corresponding to the 'unknown' peak in the recPrP peptide cycle is evident. (b) CAD spectrum of the tryptic peptide from human complement component C1q of sequence G₉₅IBGIJGTJGSPGNK₁₁₀ where B=4-hydroxyproline and J=O-glycosylated δ -hydroxylysine. Loss of up to 4 consecutive hexose residues can be seen, in both doubly and triply charged states demonstrating the presence of 2 sites having hydroxylysine linked to disaccharides.

3.4. Molecular modelling of PrP-Hyl-Gal-Glu

In order to probe the effect that glycosylation of Lys193 may have on the structure of PrP, we performed molecular dynamics simulations by use of the AMBER suite of programmes. Starting from the coordinates of murine PrP (PDB code 1AG2), we attached the disaccharide Gal-Glu in the *delta* position of Lys193 by use of the online tool available in the Glycam suite (<http://glycam.cccr.uga.edu/AMBER/index.html>) and the molecular preparation (LEaP) module of AMBER. In principle, delta-hydroxylation of lysine could result in one of two stereoisomers, designated 5R and 5S. In collagen, δ -hydroxylysine is of the 5R type, whereas the lysine hydroxylation that is carried out by the enzyme JMJD6 results in the 5S stereoisomer [45]; thus the hydroxylysine could theoretically be of either type, although the 5R isomer (as found in collagen) is most likely. Nevertheless, we made models of both types of modified PrP and the resulting coordinates were extensively minimised to optimise the lysine-disaccharide linking bonds and to remove any steric clashes introduced during the modification process. The structure of the gal-glu- δ -hydroxylysine residue is shown in Fig. 6(a) whilst Fig. 6(b) and 6(c) show ribbon structures of the post translationally modified and energy minimised protein structures used as the starting point for simulations of the 5R and 5S stereoisomers respectively.

Molecular dynamics were performed over 10 ns for each stereoisomer of modified PrP and a similar simulation was performed on unmodified, wildtype, murine PrP. For all 3 proteins, we calculated the RMS deviation of the backbone atoms compared to the starting structure and also compared to a structure that was averaged over the course of each simulation. The results of these analyses are shown in Fig. 7. In line with previously published

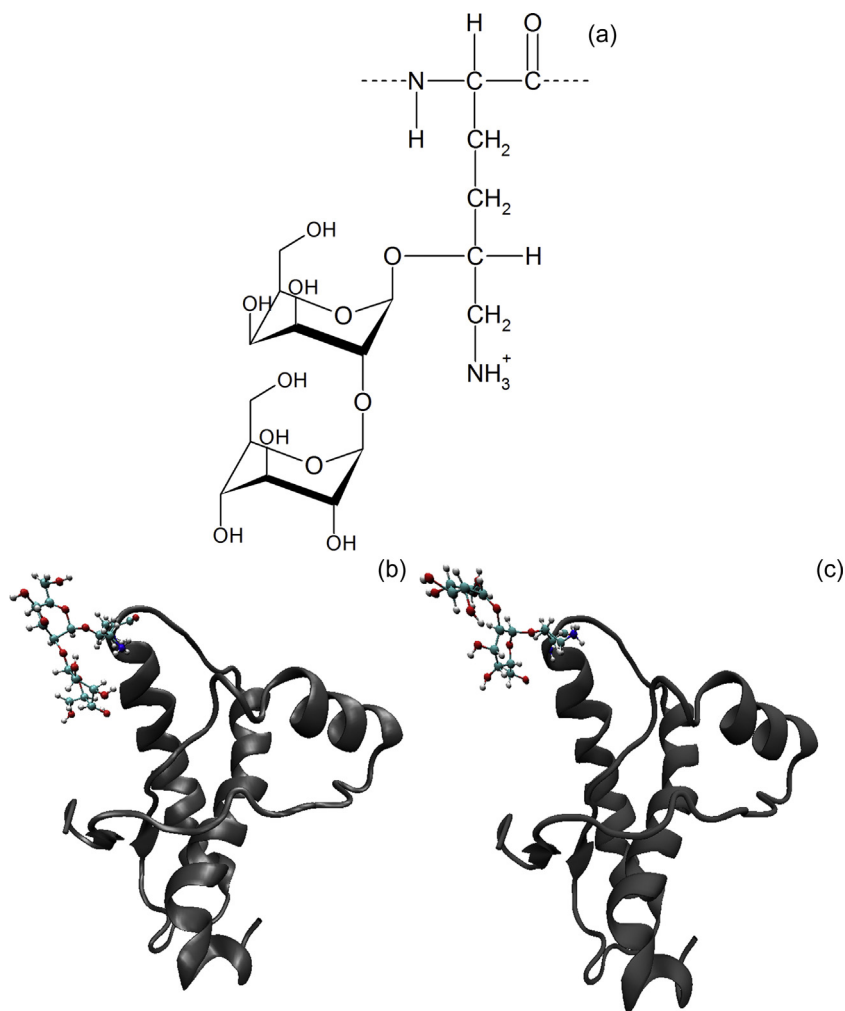


Fig. 6. (a) Schematic diagram of a δ -hydroxylysine residue with the disaccharide Gal-Glu covalently attached to the δ -oxygen. (b) and (c) Ribbon diagrams of human PrP with lys193 and the disaccharide Gal-Glu shown in ball and stick form for 5R and 5S stereoisomers respectively. The NMR structure of murine prion protein (PDB code 1AG2) was used. Lys193 was modified to attach a δ -oxygen atom followed by the 2 hexose units and the structure was energy minimised using the AMBER force field.

results [46,47], wildtype murine PrP shows relatively minor RMS fluctuations from the average structure and these dynamic excursions are most pronounced in the loop regions of the protein linking the regions of globular secondary structure. The loop region linking the second beta strand (β_2) to the second alpha helix (α_2) is particularly mobile relative to other regions of the protein.

Compared with wildtype murine PrP, simulations of PrP carrying the disaccharide at lys193 shows similar mobility in the region of the protein around the modified amino acid. The α_2 - α_3 loop appears more flexible when the disaccharide is attached as a 5S stereoisomer whilst the 5R stereoisomer, usually found in collagen, results in slight stabilisation of this loop region. In the simulations of the 5S stereoisomer, the first α -helix (α_1) also shows increased dynamic excursions although the helical structure is preserved. In the 3-dimensional structure, this helix is located next to the α_2 - α_3 loop region and the site of the modification, indicating that the flexibility of this loop region in this model also causes increased mobility in helix α_1 . Conversely, the β_2 - α_2 loop region shows substantially reduced mobility in the simulations of both modified proteins, as evidenced by small RMS deviations around residues 163–173 relative to the averaged structure for each simulation (Fig. 6(b)). The β_2 - α_2 loop is stabilised regardless of the stereoisomeric nature of the hydroxylysine, suggesting a long range stabilisation of this region of the protein as a result of the presence of O-glycosylation on Lys193. Recent studies have also shown long

range stabilisation of this loop by changes to amino acids in other regions of PrP [48,49].

The differences in mobility of the β_2 - α_2 loop regions of the three proteins can be seen clearly in Fig. 8, where these regions of the average structure of each protein are shown in detail alongside the loop region of the starting structure. For the average structures from the 3 simulations, the diameter of the backbone trace indicates the RMS deviation calculated relative to the average structure over the course of each simulation. In the simulations of all three proteins a short 3–10 helix is formed between residues 165 and 168 and this structure is substantially more stable when hydroxylysine glycosylation is present at Lys193. The stability of the β_2 - α_2 loop region of the prion protein is believed to be critical for the structural conversion to PrP^{Sc}. Indeed, transgenic mice expressing protein in which the mobility of this loop has been constrained developed spontaneous prion disease [50,51], strongly suggesting that a rigid β_2 - α_2 loop is a factor governing susceptibility to disease.

4. Discussion

Hydroxylysine and its O-glycosylated derivative commonly occur in collagen but have been found in a variety of other proteins that possess collagenous domains, such as mannose binding protein (MBP), pulmonary surfactant protein D (SP-D) and complement component C1q. Hydroxylation takes place within consensus

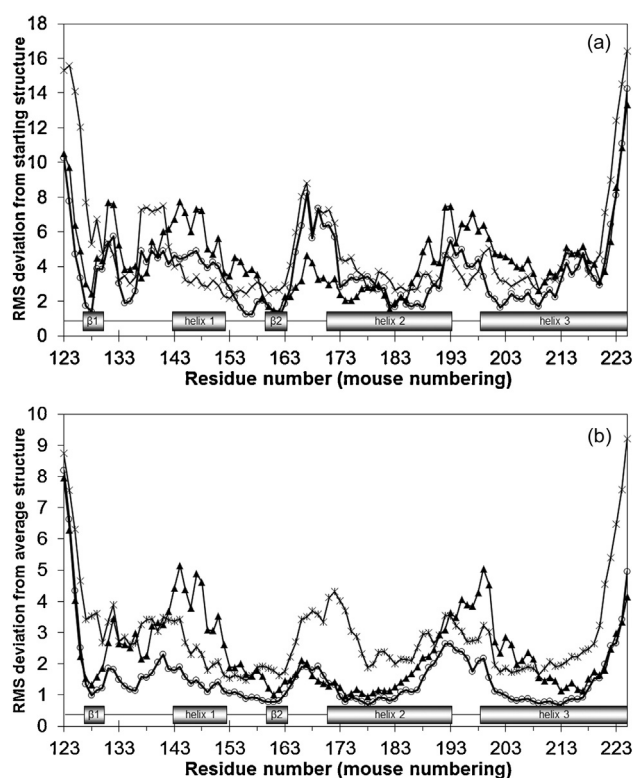


Fig. 7. Graphs showing per residue RMS deviations of backbone atoms from (a) the starting structures and (b) the average structures calculated over the course of the simulations for wildtype, murine PrP (×) the 5R stereoisomer of disaccharide-modified protein (▲) and the 5S stereoisomer disaccharide-modified protein (○).

sequences of Xxx-Lys-Gly by the enzyme lysyl hydroxylase [40], although this motif appears in many proteins without action, so presumably some other indicator, such as the correct conformation, is required for hydroxylation. This is certainly the case with

the previously identified hydroxylation of proline 44 of recPrP, where action of the enzyme prolyl 4-hydroxylase demonstrated the presence of poly-L-proline II structure in the N-terminal region of the protein [19,20]. In PrP, the only site that fulfils the consensus sequence for lysyl hydroxylation is Thr¹⁹²-Lys¹⁹³-Gly¹⁹⁴.

In pro-collagen, addition of monosaccharides, specifically galactose followed by glucose (Fig. 6(a)), to δ -hydroxylysine by carbohydrate transferases results in collagenous hydroxylysine residues having no, one or two hexose units attached. Hydroxylysines in proteins other than collagen tend to have only the entire disaccharide attached implying a separate transferase may act on these proteins. Within collagen the exact role of hydroxylysine is not clear, but these residues have been shown to promote fibril strength through the formation of inter-chain, covalent bonds [52]. The role of the carbohydrates appears to be to prevent these inter-chain links. In other proteins the role of hydroxylysine residues may also be to mediate multimerisation of protein chains.

Detection of carbohydrate attachment to hydroxylysine has previously been based on a range of techniques such as resistance to trypsin digestion, absent residues during Edman sequencing or peptide mass measurement by mass spectrometry. Peptide mass fingerprinting of a trypsin digest of SP-D showed a variety of adducts of mass 340 Da on certain peptides, suggesting hydroxylation and disaccharide addition in positions that this modification was expected to occur [53]. To our knowledge, our work represents the first identification of the elution time of PTH-Glu-Gal-hydroxylysine during Edman sequencing but other studies have also made use of MS/MS techniques to identify disaccharide attachment to hydroxylysine residues [41].

The subcellular site of action of lysyl hydroxylation has been studied extensively and it is generally accepted that both prolyl and lysyl hydroxylases work co-translationally in the ER with the growing or nascent polypeptide chain [54] whilst glycosylation of Hyl residues may occur in either the ER or the golgi. Walmsley et al., found that recPrP lacking the GPI anchor and C-terminal signal sequence, expressed in mammalian cell lines, experienced delayed transit to the cell surface, consistent with increased processing time

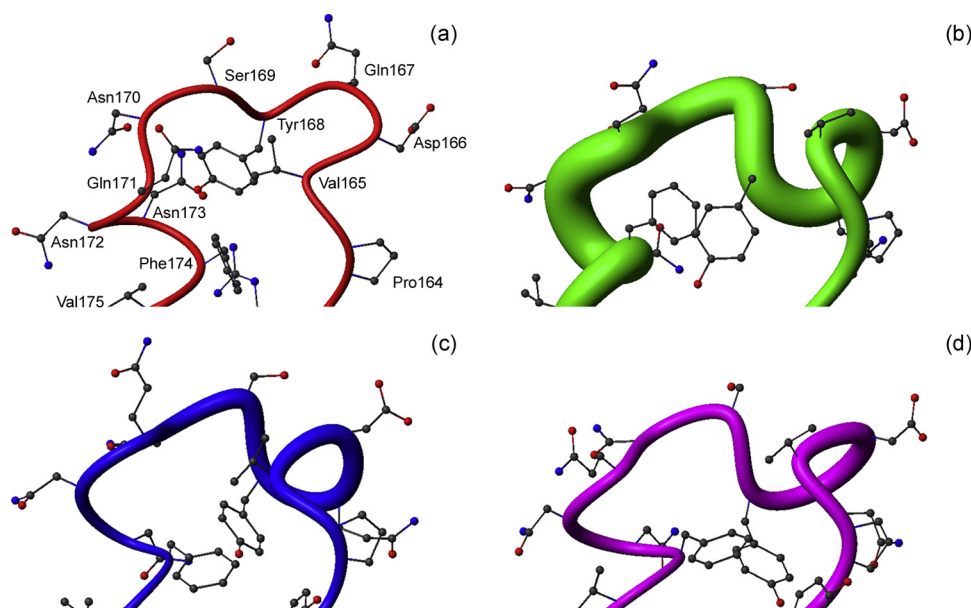


Fig. 8. Close up of the β_2 - α_2 loop region of PrP from (a) the NMR structure of murine PrP determined by Riek et al. [24] used as a starting point for our simulations, PDB code 1AG2 (b) the averaged structure from molecular dynamics simulation of wildtype murine PrP residues 124–226 (c) the averaged structure from molecular dynamics simulations of murine PrP modified to incorporate hydroxylation of lysine 193 with the disaccharide gal-glu attached as a 5R stereoisomer (d) the averaged structure from molecular dynamics simulations of murine PrP modified to incorporate hydroxylation of lysine 193 with the disaccharide Gal-Glu attached as a 5S stereoisomer. For the structures shown in panels (b)–(d) the loop region has been correctly oriented by fitted to the starting structure and the diameter of the backbone spline is proportional to the per-residue RMS deviations from the averaged structures calculated over the course of each simulation. Amino acid side chains are shown as ball and stick models and are labelled in panel (a).

in the ER or golgi [55]. The delay in protein transit is likely to account for the extent of the modifications in our recombinant PrP, which is similarly mutated to remove the C-terminal signal sequence. However, the finding that a proportion of the molecules of PrP^{Sc} purified from the brains of scrapie infected mice were hydroxylated at Pro44 indicated that this modification can occur *in vivo* in the absence of non-physiological mutations. It seems likely that a similarly small percentage of molecules will also contain hydroxylysine but establishing this is problematic due to the close proximity of Lys193 to the N-linked glycosylation site, Asn196. However, transgenic mice have recently been created that express PrP from a gene in which the C-terminal signal sequence has been ablated (PrP GPI⁻) and we expect this protein to be trafficked similarly to that in our CHO cell line. In mice, PrP GPI⁻ is secreted from cells and infection of these animals with mouse-passaged scrapie produces unusual results [56,57]. PrP GPI⁻ is certainly trafficked abnormally, as evidenced by substantially reduced N-linked glycosylation, and it is possible that this allows aberrant modifications, such as proline or lysine hydroxylation, to take place in the secretory pathway. As yet, we do not know whether PrP^C (or PrP^{Sc}) produced *in vivo* can be hydroxylated and glycosylated at Lys193.

An important question is how glycosylation at Lys193 might affect the structure of PrP^C. Previous molecular dynamics simulations of PrP with N-linked glycans attached suggests that the carbohydrates are relatively flexible and can perform large dynamic excursions without greatly influencing the tertiary structure of the protein [46,58]. In the current work, we find that the glycosylation attached to Lys193 is also flexible but that its presence does not greatly influence the motion in the protein backbone of the loop immediately adjacent to Lys193. In one of the two stereoisomeric models, slightly enhanced motion in this region led to some instability in helix 1 but the motion was not sufficient to cause unravelling of helices 2 and 3. Conversely, there appeared to be long range stabilisation effects on the β_2 - α_2 loop of PrP in both 5R and 5S stereoisomer models, presumably as a direct result of the presence of O-glycosylation at Lys193. A rigid β_2 - α_2 loop has been suggested to be important for PrP conversion and we have previously found an inverse correlation between the lability of this loop and the ability of PrP to convert *in vitro* [59]. This raises the possibility that hydroxylation and subsequent glycosylation of Lys193 may predispose PrP to disease-specific conversion. Thus, conditions that one may expect to increase these modifications, such as lack of the C-terminal signal sequence or ER stress, may also result in an enhanced ability of PrP to convert. We are currently following up these ideas in our laboratories.

Finally, it is interesting to note that Lys193 is only 3 residues N-terminal to the second N-linked glycosylation site within PrP, Asn196. The presence of a disaccharide attached to Lys193 may hinder carbohydrate transferases or subsequent glycan trimming reactions at Asn196 leading to abnormal glycosylation in a percentage of PrP molecules. Conversely, the presence of an N-linked glycan at Asn196, may prevent disaccharide addition to any hydroxylysine residue formed at codon 193. Which event is likely to predominate, if either, will depend greatly on the order of modification events in the ER/golgi complex.

Acknowledgements

We thank Sharon Bocking, Sarah Steane, Sureeta Saini, Kenneth Davis and Alan Bennett for CHO-cell protein production in addition to Guy Dodson, Peter Bayley and James Hope for their support. This work was funded, in part, by grants from the BBSRC under the BSEP IV research programme and from a BBSRC Institute Strategic Programme Grant to the Roslin Institute. The funders had no role in study design, in the collection, analysis and interpretation of data, in

the writing of the report nor in the decision to submit the article for publication. ACG would also particularly like to gratefully acknowledge Prof. K.R. Jennings, whose support during his PhD studies was instrumental to his chosen career path (split infinitive and hanging participle deliberate!) – many thanks Keith.

References

- [1] A.C. Gill, F.M. Lane, J. Alibhai, J.C. Manson, S. McCutcheon, The mechanisms of prion disease, CAB reviews: perspectives in agriculture, veterinary science, Nutrition and Natural Resources 6 (2011) 1–34, p. 055.
- [2] S.B. Prusiner, Prions, Proceedings of the National Academy of Sciences of the United States of America 95 (1998) 13363–13383.
- [3] K.M. Pan, M. Baldwin, J. Nguyen, M. Gasset, A. Serban, D. Groth, I. Mehlhorn, Z.W. Huang, R.J. Fletterick, F.E. Cohen, S.B. Prusiner, Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins, Proceedings of the National Academy of Sciences of the United States of America 90 (1993) 10962–10966.
- [4] A.C. Gill, S. Agarwal, T.J.T. Pinheiro, J.F. Graham, Structural requirements for efficient prion protein conversion cofactors may promote a conversion-competent structure for PrP^C, Prion 4 (2010) 235–243.
- [5] J.F. Graham, S. Agarwal, D. Kurian, L. Kirby, T.J.T. Pinheiro, A.C. Gill, Low density subcellular fractions enhance disease-specific prion protein misfolding, Journal of Biological Chemistry 285 (2010) 9868–9880.
- [6] N.R. Deleault, B.T. Harris, J.R. Rees, S. Supattapone, Formation of native prions from minimal components *in vitro*, Proceedings of the National Academy of Sciences of the United States of America 104 (2007) 9741–9746.
- [7] N.R. Deleault, R. Kascasak, J.C. Geoghegan, S. Supattapone, Species-dependent differences in cofactor utilization for formation of the protease-resistant prion protein *in vitro*, Biochemistry – US 49 (2010) 3928–3934.
- [8] G.P. Saborio, B. Permanne, C. Soto, Sensitive detection of pathological prion protein by cyclic amplification of protein misfolding, Nature 411 (2001) 810–813.
- [9] D.A. Kocsisko, J.H. Come, S.A. Priola, B. Chesebro, G.J. Raymond, P.T. Lansbury, B. Caughey, Cell-free formation of protease-resistant prion protein, Nature 370 (1994) 471–474.
- [10] R. Atarashi, J.M. Wilham, L. Christensen, A.G. Hughson, R.A. Moore, L.M. Johnson, H.A. Onwubiko, S.A. Priola, B. Caughey, Simplified ultrasensitive prion detection by recombinant PrP conversion with shaking, Nature Methods 5 (2008) 211–212.
- [11] L. Kirby, C.R. Birkett, H. Rudyk, I.H. Gilbert, J. Hope, *In vitro* cell-free conversion of bacterial recombinant PrP to PrP^{res} as a model for conversion, Journal of General Virology 84 (2003) 1013–1020.
- [12] M.T. Bishop, R.G. Will, J.C. Manson, Defining sporadic Creutzfeldt–Jakob disease strains and their transmission properties, Proceedings of the National Academy of Sciences of the United States of America 107 (2010) 12005–12010.
- [13] S. Dudas, J.M. Yang, C. Graham, M. Czub, T.A. McAllister, M.B. Coulthart, S. Czub, Molecular, biochemical and genetic characteristics of BSE in Canada, PLoS ONE 5 (2010).
- [14] H.A. Simmons, M.M. Simmons, Y.I. Spencer, M.J. Chaplin, G. Povey, A. Davis, A. Ortiz-Pelaez, N. Hunter, D. Matthews, A.E. Wrathall, Atypical scrapie in sheep from a UK research flock which is free from classical scrapie, BMC Veterinary Research 5 (2009).
- [15] N. Stahl, M.A. Baldwin, D. Teplow, L. Hood, R. Beavis, B. Chait, B.W. Gibson, A.L. Burlingame, S.B. Prusiner, Cataloging posttranslational modifications of the scrapie prion protein by mass-spectrometry, in: EHS Neuro, 1992, pp. 361–379.
- [16] N. Stahl, M.A. Baldwin, D.B. Teplow, L. Hood, B.W. Gibson, A.L. Burlingame, S.B. Prusiner, Structural studies of the scrapie prion protein using mass-spectrometry and amino-acid sequencing, Biochemistry – US 32 (1993) 1991–2002.
- [17] M.A. Ritchie, A.C. Gill, M.J. Deery, K. Lilley, Precursor ion scanning for detection and structural characterization of heterogeneous glycopeptide mixtures, Journal of the American Society for Mass Spectrometry 13 (2002) 1065–1077.
- [18] E. Stimson, J. Hope, A.L. Burlingame, Structures of the N-linked glycans of the prion protein isolated from murine models of scrapie, FASEB Journal 12 (1998) A1349.
- [19] A.C. Gill, M.A. Ritchie, L.G. Hunt, S.E. Steane, K.G. Davies, S.P. Bocking, A.G.O. Rhie, A.D. Bennett, J. Hope, Post-translational hydroxylation at the N-terminus of the prion protein reveals presence of PPH structure *in vivo*, EMBO Journal 19 (2000) 5324–5331.
- [20] E.W. Blanch, A.C. Gill, A.G.O. Rhie, J. Hope, L. Hecht, K. Nielsen, L.D. Barron, Raman optical activity demonstrates poly(L-proline) II helix in the N-terminal region of the ovine prion protein: Implications for function and misfunction, Journal of Molecular Biology 343 (2004) 467–476.
- [21] J. Hope, L.J.D. Morton, C.F. Farquhar, G. Multhaup, K. Beyreuther, R.H. Kimberlin, The major polypeptide of scrapie-associated fibrils (Saf) has the same size, charge-distribution and N-terminal protein-sequence as predicted for the normal brain protein (Prp), EMBO Journal 5 (1986) 2591–2597.
- [22] J. Hope, G. Multhaup, L.J.D. Reekie, R.H. Kimberlin, K. Beyreuther, Molecular pathology of scrapie-associated fibril protein (Prp) in mouse-brain affected by the Me7 strain of scrapie, European Journal of Biochemistry 172 (1988) 271–277.
- [23] Y.G. Choi, J.I. Kim, Y.C. Jeon, S.J. Park, E.K. Choi, R. Rubenstein, R.J. Kascasak, R.I. Carp, Y.S. Kim, Nonenzymatic glycation at the N terminus of pathogenic prion

- protein in transmissible spongiform encephalopathies, *Journal of Biological Chemistry* 279 (2004) 30402–30409.
- [24] R. Riek, S. Hornemann, G. Wider, M. Billeter, R. Glockshuber, K. Wuthrich, NMR structure of the mouse prion protein domain PrP(121–231), *Nature* 382 (1996) 180–182.
- [25] V. Smirnovas, G.S. Baron, D.K. Offerdahl, G.J. Raymond, B. Caughey, W.K. Surewicz, Structural organization of brain-derived mammalian prions examined by hydrogen-deuterium exchange, *Nature Structural & Molecular Biology* 18 (2011) 504–506.
- [26] G. Colombo, M. Meli, G. Morra, R. Gabizon, M. Gasset, Methionine sulfoxides on prion protein helix-3 switch on the alpha-fold destabilization required for conversion, *PLoS ONE* 4 (2009).
- [27] D.J. Weber, P.N. McFadden, B. Caughey, Measurement of altered aspartyl residues in the scrapie associated form of prion protein, *Biochemical and Biophysical Research Communications* 246 (1998) 606–608.
- [28] E. Sandmeier, P. Hunziker, B. Kunz, R. Sack, P. Christen, Spontaneous deamidation and isomerization of Asn108 in prion peptide 106–123 and in full-length prion protein, *Biochemical and Biophysical Research Communications* 261 (1999) 578–583.
- [29] D.V. Dear, D.S. Young, J. Kazlauskaitė, F. Meersman, D. Oxley, J. Webster, T.J.T. Pinheiro, A.C. Gill, I. Bronstein, C.R. Lowe, Effects of post-translational modifications on prion protein aggregation and the propagation of scrapie-like characteristics in vitro, *BBA – Proteins Proteomics* 1774 (2007) 792–802.
- [30] D.S. Young, F. Meersman, D. Oxley, J. Webster, A.C. Gill, I. Bronstein, C.R. Lowe, D.V. Dear, Effect of enzymatic deimination on the conformation of recombinant prion protein, *BBA – Proteins Proteomics* 1794 (2009) 1123–1133.
- [31] D.B. Brimacombe, A.D. Bennett, F.S. Wusteman, A.C. Gill, J.C. Dann, C.J. Bostock, Characterization and polyanion-binding properties of purified recombinant recombinant protein, *Biochemical Journal* 342 (1999) 605–613.
- [32] D. Tong, R.L. Moritz, J.S. Eddes, G.E. Reid, R.K. Rasmussen, D.S. Dorow, R.J. Simpson, Fabrication of stable packed capillary reversed-phase columns for protein structural analysis, *Journal of Protein Chemistry* 16 (1997) 425–431.
- [33] D.A. Case, T.E. Cheatham, T. Darden 3rd, H. Gohlke, R. Luo, K.M. Merz, A. Onufriev Jr., C. Simmerling, B. Wang, R.J. Woods, The amber biomolecular simulation programs, *Journal of Computational Chemistry* 26 (2005) 1668–1688.
- [34] M. Basma, S. Sundara, D. Calgan, T. Vernali, R.J. Woods, Solvated ensemble averaging in the calculation of partial atomic charges, *Journal of Computational Chemistry* 22 (2001) 1125–1137.
- [35] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, *Electrophoresis* 18 (1997) 2714–2723.
- [36] A. Onufriev, D. Bashford, D.A. Case, Exploring protein native states and large-scale conformational changes with a modified generalized born model, *Proteins* 55 (2004) 383–394.
- [37] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *Journal of Molecular Graphics* 14 (33–38) (1996) 27–38.
- [38] R. Koradi, M. Billeter, K. Wuthrich, MOLMOL: a program for display and analysis of macromolecular structures, *Journal of Molecular Graphics* 14 (51–55) (1996) 29–32.
- [39] G.A. Grant, M.W. Crankshaw, Identification of PTH-amino acids by HPLC, *Methods in Molecular Biology* 211 (2003) 247–268.
- [40] J. Myllyharju, Intracellular post-translational modifications of collagens, *Topics in Current Chemistry* 247 (2005) 115–147.
- [41] Y. Taga, M. Kusubata, K. Ogawa-Goto, S. Hattori, Development of a novel method for analyzing collagen O-glycosylations by hydrazide chemistry, *Molecular & Cellular Proteomics* 11 (2012) M111.010397.
- [42] R.M. Vanacore, D.B. Friedman, A.J. Ham, M. Sundaramoorthy, B.G. Hudson, Identification of S-hydroxylsyl-methionine as the covalent cross-link of the noncollagenous (NC1) hexamer of the alpha1alpha1alpha2 collagen IV network: a role for the post-translational modification of lysine 211 to hydroxyllysine 211 in hexamer assembly, *Journal of Biological Chemistry* 280 (2005) 29300–29310.
- [43] D. Pflieger, C. Przybylski, F. Gonnet, J.P. Le Caer, T. Lunardi, G.J. Arlaud, R. Daniel, Analysis of human C1q by combined bottom-up and top-down mass spectrometry: detailed mapping of post-translational modifications and insights into the C1r/C1s binding sites, *Molecular and Cellular Proteomics* 9 (2010) 593–610.
- [44] K. Yonemasu, H. Shinkai, T. Sasaki, Comparable content of hydroxyllysine-linked glycosides in subcomponents C1q of the first component of human, bovine and mouse complement, *Collagen and Related Research* 1 (1981) 385–390.
- [45] M. Mantri, N.D. Loik, R.B. Hamed, T.D. Claridge, J.S. McCullagh, C.J. Schofield, The 2-oxoglutarate-dependent oxygenase JMJD6 catalyses oxidation of lysine residues to give 5S-hydroxyllysine residues, *Chembiochem* 12 (2011) 531–534.
- [46] M.L. DeMarco, V. Daggett, All-atom molecular dynamics simulations of the diglycosylated and membrane-bound prion protein, *Glycobiology* 15 (2005) 1235.
- [47] M.L. DeMarco, V. Daggett, Molecular mechanism for low pH triggered misfolding of the human prion protein, *Biochemistry – US* 46 (2007) 3045–3054.
- [48] B. Christen, S. Hornemann, F.F. Damberger, K. Wuthrich, Prion protein nmr structure from Tammar Wallaby (*Macropus eugenii*) shows that the beta 2-alpha 2 loop is modulated by long-range sequence effects, *Journal of Molecular Biology* 389 (2009) 833–845.
- [49] A.D. Scouras, V. Daggett, Disruption of the X-loop turn of the prion protein linked to scrapie resistance, *Protein Engineering, Design & Selection* 25 (2012) 243–249.
- [50] C.J. Sigurdson, S. Joshi-Barr, C. Bett, O. Winson, G. Manco, P. Schwarz, T. Rulicke, K.P.R. Nilsson, I. Margalith, A. Raeber, D. Peretz, S. Hornemann, K. Wuthrich, A. Aguzzi, Spongiform encephalopathy in transgenic mice expressing a point mutation in the beta 2-alpha 2 loop of the prion protein, *Journal of Neuroscience* 31 (2011) 13840–13847.
- [51] C.J. Sigurdson, K.P.R. Nilsson, S. Hornemann, M. Heikenwalder, G. Manco, P. Schwarz, D. Ott, T. Rulicke, P.P. Liberski, C. Julius, J. Falsig, L. Stitz, K. Wuthrich, A. Aguzzi, De novo generation of a transmissible spongiform encephalopathy by mouse transgenesis, *Proceedings of the National Academy of Sciences of the United States of America* 106 (2009) 304–309.
- [52] D.R. Eyre, M.A. Weis, J.J. Wu, Advances in collagen cross-link analysis, *Methods* 45 (2008) 65–74.
- [53] R. Leth-Larsen, U. Holmskov, P. Hojrup, Structural characterization of human and bovine lung surfactant protein D, *Biochemistry Journal* 343 (Pt 3) (1999) 645–652.
- [54] K.I. Kivirikko, T. Pihlajaniemi, Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases, *Advances in Enzymology and Related Areas of Molecular Biology* 72 (1998) 325–398.
- [55] A.R. Walmsley, F.N. Zeng, N.M. Hooper, Membrane topology influences N-glycosylation of the prion protein, *EMBO Journal* 20 (2001) 703–712.
- [56] B. Chesebro, M. Trifilo, R. Race, K. Meade-White, C. Teng, R. LaCasse, L. Raymond, C. Favara, G. Baron, S. Priola, B. Caughey, E. Masliah, M. Oldstone, Anchorless prion protein results in infectious amyloid disease without clinical scrapie, *Science* 308 (2005) 1435–1439.
- [57] B. Chesebro, B. Race, K. Meade-White, R. LaCasse, R. Race, M. Klingeborn, J. Striebel, D. Dorward, G. McGovern, M. Jeffrey, Fatal transmissible amyloid encephalopathy: a new type of prion disease associated with lack of prion protein membrane anchoring, *PLoS Pathogen* 6 (2010).
- [58] J. Zuegg, J.E. Gready, Molecular dynamics simulation of human prion protein including both N-linked oligosaccharides and the GPI anchor, *Glycobiology* 10 (2000) 959–974.
- [59] L. Kirby, S. Agarwal, J.F. Graham, W. Goldmann, A.C. Gill, Inverse correlation of thermal lability and conversion efficiency for five prion protein polymorphic variants, *Biochemistry – US* 49 (2010) 1448–1459.