Transmissible spongiform encephalopathies are a group of neural degenerative diseases that may be infectious, sporadic, or hereditary and are associated with an abnormally folded prion protein. Unfortunately at the current time it is not at all clear what the normal structure of the prion protein actually is or how it is toxic to cells.

Extensive research on prion diseases has led to a dramatic increase in understanding of the pathogenesis of prion disorders, which will hopefully lead to the development of effective treatments. The inability to detect the disease in blood using current technology has made screening difficult. While fortunately there has been a decline in the number of clinical cases of transmissible variant CJD, evidence indicates that very long incubation periods can occur in humans so there may be a long slow, gradual epidemic. In particular, clinical cases in genotypes other than those homozygous for methionine at codon 129 of PRNP have not yet occurred, but such cases might be expected to have longer incubation periods and show differences in pathology to those seen to date.

Transgenic animal studies have shown that a large proportion of infected animals develop sub-clinical disease. Moreover, results from a large prevalence study in humans show that several cases test positive but do not develop clinical disease. It is possible therefore that further cases of secondary transmission could occur by iatrogenic spread, which could result in vCJD persisting in the UK at low levels for many years, highlighting the importance of continued vigilance.

**Key words:** Prion diseases, *PRNP* gene, Creutzfeldt Jakob disease (CJD)
and speech abnormalities and seizures. In most patients, these symptoms are followed by involuntary movements and the appearance of a typical diagnostic electroencephalogram tracing (Knight and Will, 2004).

Variant CJD (vCJD) was first described in 1996 in the United Kingdom. Epidemiologic and laboratory evidence indicate that vCJD in humans is caused by transmission of the agent causing BSE in cattle via consumption of BSE-contaminated cattle products (Brown et al., 2001; Belay and Schonberger, 2002). The variant form of CJD has several different characteristics (Brown et al., 2001; Belay and Schonberger, 2002); the median age at death for vCJD patients is 28 years, compared with 68 years for patients with classic CJD. In addition, all vCJD cases have neuropathologic findings distinctly different from those of classic CJD (Irons, 1998), and all have had a particular genetic profile i.e., homozygosity for methionine at codon 129 of the prion protein gene (Will et al., 2004). Thus, cases of vCJD can be distinguished from classic CJD by the characteristic spongiform encephalopathy and normal electroencephalogram tracing (Knight and Will, 2004). Interestingly, 14 of the 15 SINEs (Alu sequences) belong to subfamilies (AluSx, AluJ, and monomers) thought to be >35 million years old (Smit and Riggs, 1995; Kapitonov and Jurka, 1996).

From a comparison of human, sheep and mouse PRNP/PrP genes, it is clear that there are common sites of integration of known classes of transposable elements or processed pseudogenes (Fig. 2). Insertion of these elements since the mammalian radiation has increased the size of the human PRNP locus by 20% and both the mouse and sheep loci by 41% although some large deletions have partially reversed this size increase. In almost all cases, repetitive sequences belonging to subfamilies (AluSx, AluJ, and monomers) thought to be >35 million years old (Smit and Riggs, 1995; Kapitonov and Jurka, 1996).

The exonic structure of the human gene differs from that of other mammals. While the first exon and the beginning of intron 1 are poorly conserved, there is a highly conserved region of 99-bp present in human DNA, 2303 bp after exon 1. This region is 82% and 81% identical to exon 2 of sheep and mouse, respectively (Fig. 3) (Lee et al., 1998). However this gene region does not appear to be encoded in humans as it is not present in any human PRNP mRNAs analyzed to date even though this human exon 2-like sequence is flanked by consensus splice donor and acceptor sites similar (but not identical) to those in the mouse and sheep genes. Also, since the exon 2-like sequence does not contain a translation start codon (ATG) in the correct translational reading frame with human PRNP exon 3, it forms part of the UTR of human RNA.

All human PRNP cDNAs isolated to date have exon 1 spliced directly to exon 3 (using the exon numbering system of the sheep and mouse PrP genes). It is possible that nucleotide substitutions at the donor splice site in the human gene (as indicated in Fig. 3B) are responsible for the absence of exon 2 since the acceptor splice site is typical and highly conserved. The conservation of the

The human prion protein gene (PRNP)

The human gene (PRNP) that encodes the prion protein (PrP) was first identified in scrapie-infected rodents. It is a single-copy chromosomal gene and the protein is encoded from a single exon. In addition to the protein-coding exon (exon 3), PrP genes in mammals contain one (humans) or two (sheep and mice) 5' noncoding exons that are included in the final mRNA (Lee et al., 1998) (Fig. 1). Further upstream, the promoter region of the human PRNP gene is a well-characterized region that is highly GC-rich, lacks a canonical TATA box, contains a CCAAT box, and has a number of putative binding sites for transcription factors SP1, AP1 and AP2 (Mahal et al., 2001).

The structure of the PRNP gene and mRNA

The nucleotide sequence of human PRNP, determined from cosmide clone pGPRP1 (Puckett et al., 1991), is 35,522 bp in length. The gene has two known exons (134 and 2355 nucleotides in length) separated by an intron of 12,696 bases, and a dG + dC content of 44.1%. A 1-kb region around exon 1 forms a "CpG island" (Fig. 1) (Lee et al., 1998), as found 5' of the transcriptional start site of many housekeeping genes.

There are also two repeats in the human PRNP gene locus (MER74 and MER88; Fig. 2) that are related to two consensus sequences in the human repeat database (MER73 and MER54) (characteristic of viral LTRs (Long Terminal Repeats), having TG...CA terminal dinucleotides flanked by 5-bp insertion site duplications (Smit and Riggs, 1995).

In fact, a large percentage of the human PRNP locus appears to be derived from mobile elements; there are several short interspersed repetitive element (SINE), and long interspersed repetitive element (LINE) insertions in the PRNP gene (Smit, 1996; Smit et al., 1996). Interestingly, 14 of the 15 SINEs (Alu sequences) belong to subfamilies (AluSx, AluJ, and monomers) thought to be >35 million years old (Smit and Riggs, 1995; Kapitonov and Jurka, 1996).
The first half of exon 2 in the human sequence is possibly caused by the presence of transcription factor binding sites at this locus. Baybutt and Manson (1997) have shown that the last 550 bp of intron 1 is important for promoter activity of the mouse PrP gene, and it is possible that this functional region extends into exon 2 (Lee et al., 1998).

Alternatively, the human exon 2 could be included within a subset of mRNAs from tissues that have not yet been analyzed. This possibility is compatible with observed tissue specific alternate exon 2 splicing of PrP in other mammals; exon 2-containing transcripts have been described in cattle (Yoshimoto et al., 1992) and in the Syrian hamster (SHa) PrP gene, for which exon 1 originally appeared directly spliced to exon 3 (Basler et al., 1986). Whereas 90% of brain SHaPrP mRNAs exhibit exon 1-3 splicing, 10% include exon 2 sequences similar to those of sheep and mouse (Li and Bolton, 1997). Interestingly, an increased abundance of exon 2-containing mRNAs was observed in brain mRNA of scrapie-infected hamsters, which may reflect preferential expression in astrocytes (which become activated during the course of prion infections).

**The PrPc protein**

The normal prion protein, PrPc, encoded by exon 3 of the prion gene (PRNP), is produced in prion diseases,
Prion protein misfolding and Creutzfeldt-Jakob disease (CJD)

from the normal host gene but undergoes an incompletely understood post-translational conformational change to a disease related form, PrPSc. The function of normal PrPc is not known but it may have roles in anti-oxidant systems and cellular copper metabolism (Knight and Will, 2004) although within the nervous system, it may have a role in synaptic function as PrP-knockout mice show impaired GABA-mediated synaptic inhibition and reduced long-term potentiation (Collinge et al., 1994; Manson et al., 1995).

The normal PrPc protein is a glycoprotein that is attached to the plasma membrane by a glycosyl phosphatidylinositol anchor (Prusiner, 1998). It is present in cells of the central nervous system (Kretzschmar et al., 1986; Stahl et al., 1987, 1990) and in muscle cells (Brown et al., 1998). In neurons prion protein is mainly localized at synaptic terminals, in cholesterol-rich microdomains of the plasma membrane known as caveolae (Prusiner, 1998), PrPc is composed of approximately 250 residues, and is proteolytically processed to remove a 22-residue N-terminal signal peptide and a 23 C-terminal amino acid segment after addition of the glycosyl phosphatidylinositol (GPI) anchor to Ser-231 (Prusiner, 1996, 1997). Molecular modeling studies indicate that PrPc contains several α-helical regions of secondary structure (Gasset et al., 1992; Huang et al., 1994). The processed protein contains a 179-214 disulfide crosslink and glycosylation sites at Asn-181 and Asn-197. A protease resistant core present between residues 90-231, is sufficient to transmit

---

Fig. 2. Comparison of the human, sheep, and mouse PRNP/PrP loci. The complete sheep and human sequences are shown, whereas the Prp mouse cosmid sequence extends 2.5 kb upstream of the aligned regions. Deletions in the genome >250 bp are indicated by brackets. Identified interspersed repeats are represented by four differently shaded arrows for the four major classes of repeats: SINEs, LINEs, LTR elements, and DNA transposons. Wide arrows are used for elements that are absent at the orthologous sites in the other two species and thus probably integrated after the mammalian radiation. Older elements are indicated with thinner arrows. The indicated shared interspersed repeats are usually not detectable in the mouse genome (and sometimes not in the sheep genome) by direct comparison with the repeat consensus sequence but could be inferred from the alignment with human DNA. A measure of similarity between the human and either sheep or mouse sequences, depicted by a graph above and below the human sequence, respectively. Regions outside the coding region that are conserved between all species are indicated with an asterisk (*) (derived from Lee et al., 1998).
infectivity when misfolded.

**Postranslational modifications producing a mature PrPc**

The co- and posttranslational modifications of PrPc are initiated by the cleavage of the N-terminal signal peptide (aa 1-22) and the attachment of core glycans while the nascent chain is still associated with the translocon. Shortly after the protein is fully translocated the GPI anchor is attached to the acceptor amino acid close to the C terminus. The final maturation of PrPc includes the processing of the core glycans into complex type glycans by a series of enzymatic reactions in the endoplasmic reticulum (ER) and Golgi compartment. Experiments in which the N-terminal domain of PrPc is deleted show that this region is not essential and has no impact on the formation of complex glycosylated and GPI-anchored PrPc. By contrast, α-helix 1 is essential for GPI anchoring and complex glycosylation. Moreover, a single amino acid substitution in helix 3 can cause a loss of the GPI anchor possibly due to interactions between helix 1 and helix 3 (Pekari and Schmidt, 2003).

The GPI anchor signal sequence, the attachment of the GPI anchor and membrane anchorage of PrPc are all necessary for the terminal processing of glycans. However, complex glycosylation is not specific for GPI-anchored PrP; PrP-CD4 in which PrP is fixed to the membrane via the heterologous CD4 transmembrane domain is complex-glycosylated as well. In the absence of a membrane anchor, terminal glycosylation does not take place indicating that membrane anchoring provides a specialized environment (i.e. the membrane surface.

**Fig. 3.** Alignments of sequences outside the PrP coding regions conserved between the human, mouse, and sheep. The most conserved regions are blocked. A. There are two conserved sections of intron 1 that may be involved in transcriptional regulation. Both sites are within a conserved LINE1 element (Fig. 2). However, in the sheep gene they all diverge from the consensus binding sites. Baybutt and Manson (1997) identified transcriptional suppressor activity between bases 9744 and 9932 and promoter activity between bases 10,114 and 10,307 of the mouse sequence, suggesting that the second, highly conserved block (bases 10,200-10,233 in mouse) may be involved in transcriptional activation. B. Alignment of the human, sheep, and mouse PrP gene exon 2 regions. Comparison of the exon 2 splice sites shows that the donor site is quite distinct in sheep and mouse but has become indistinct in the human genome because of the presence of C and A at positions 4 and 5. The donor site consensus and that of the other species (including cow, rat, and hamster; data not shown) contain a purine and G. C. A 40-bp block –100 bp before the polyadenylation site in the PpP gene is highly conserved between four mammalian species. D. In the 3' flanking region of the PpP gene are two small blocks of conserved sequences lying within an ancient DNA transposon, called MER5a. There are three other conserved regions: one 500 bp upstream of “exon 2” (Fig. 3A), one at the 3' end of the transcripts (Fig. 3C), and one 3-4 kb downstream of the polyadenylation site (Fig. 3D). (Derived from Smit and Riggs, 1995)
may be necessary for native folding of PrPc (Pekari and Schmidt, 2003).

**PrPc protein misfolding**

The pathogenic form of the protein is referred to as PrPSc (PrP Scrapie) or PrPRes (protease resistant protein). The conversion of PrPc→PrPSc, the pathogenic form, is a late post-translational process, occurring after PrPc has reached its normal cellular location. In contrast to PrPc which is largely α-helical, PrPSc is characterized by a high content of β-sheet structures (Caughey et al., 1991a,b; Gasset et al., 1993), partial resistance to proteolytic digestion, and a propensity to aggregate into insoluble amyloid-like fibrils and plaques (Prusiner, 1998; Torrent et al., 2003, 2004).

The exact mechanism of conversion of PrPc into PrPSc is unclear; however, it is clear that PrPSc autocatalyses the process (Kocisko et al., 1994). Two mechanisms of propagation have been suggested. One model advocates template-directed refolding (Aguzzi et al., 2001), whereby exogenously introduced PrPSc interacts with PrPc to cause a conformational change of PrPc into PrPSc, overcoming a high energy barrier that would normally prevent the spontaneous conversion. The second proposed mechanism, ‘seeding’ or ‘nucleation polymerization’, suggests that PrPSc and PrPc are in a thermodynamic equilibrium from which PrPSc can be recruited to a more stable crystal-like ‘seed’ composed of several precursor PrPSc molecules (Jarrett and Lansbury, 1993). Studies in transgenic mice indicate that a protein chaperone facilitates the conversion of PrPSc from PrPc (Telling et al., 1995).

Recent studies using highly purified PrPSc have demonstrated its ability to transmit disease (Castilla et al., 2005), lending strong support to the ‘protein-only’ hypothesis of transmission. Furthermore the presence of PrPSc correlates with disease transmission and endogenous PrPc is a requirement for the development of the disease (Brandner et al., 1996a,b) which can be decreased with anti-PrP antibodies (Gabizon et al., 1988). Changes in neural electrophysiology seen in PrP-knockout mice (Collinge et al., 1994; Manson et al., 1995) are similar to those seen in scrapie-infected mice (Jefferys et al., 1994), in keeping with a ‘loss of function’ effect. However, in vitro, neural toxicity does not always correlate with conversion to PrPSc and the mechanism by which PrPSc causes neurodegeneration is unclear (Ma et al., 2002).

**Covalent reaction of the intramolecular disulfide bond of PrPc to form PrPSc**

PrPc→PrPSc conversion is proposed to involve disulfide polymerization, by which the terminal thiolate group of a PrPSc polymer attacks the intramolecular disulfide bond of a PrPc monomer, so that an intermolecular disulfide bond is formed, thus lengthening the PrPSc polymer while maintaining a terminal thiolate group (Welker et al., 2001). The initial thiolate attack is facilitated by the experimentally observed association of the PrPc monomer and PrPSc polymer (Horiiuchi and Caughey, 1999; Horiiuchi et al., 2000), which may destabilize the tertiary structure and the intramolecular PrPc disulfide bond. The subsequent conformational transition to the β-enriched scrapie isoform evidently protects the intermolecular disulfide bond (Muramoto et al., 1996; Herrmann and Caughey, 1998), inhibiting the back-reaction of depolymerization. Thus, in this model, the smallest infectious unit is an oligomer of PrPSc that protects its intramolecular disulfide bond(s) from thiolate attack and has at least one free thiolate at which the PrPc→PrPSc conversion can occur (Welker et al., 2001).
A proposed mechanism for protein conformational change

Kelly (1998) proposed that formation of a weak dimer of PrPc (90-231), which can occur according to analytical equilibrium ultracentrifugation studies (James et al., 1997), could be very important in converting PrPc into the prionogenic state through tertiary and quaternary structural changes (Fig. 4). Such an assembly process may be necessary to partially or almost completely unfold PrPc to form the prionogenic intermediate as in the conversion of transhyretin to amyloid fibrils or in the conversion of the lysozyme variants into amyloid respectively (Colon and Kelly, 1991, 1992; Kelly, 1996-1998; Sunde et al., 1997; Torrent et al., 2004). However, while these mechanisms are proposed to produce a PrPSc fibrillar structure similar to amyloid, in fact little is known about the structural details of PrPSc other than that it is richer in β-sheet structure than PrPc (Lansbury, 1992; Kelly, 1997; Prusiner, 1996, 1997).

Kinetic intermediate

A study by Apetri and Surewicz (2002) provided evidence that the prion protein folds into its native state via a rate-limiting, three-state mechanism involving a monomeric intermediate (Fig. 4). This early intermediate appears to be relatively compact and especially stable under acidic conditions. In addition to its role in normal prion protein folding, the folding intermediate is also likely to be of major importance in the PrPc→PrPSc conversion.

The partially folded PrP intermediate is more stable than the fully unfolded protein. This, combined with a relatively high hydrophobicity and aggregation propensity typical of folding intermediates, renders the intermediate state of PrP a better candidate than the unfolded state as a direct monomeric precursor of PrPSc. The role of a folding intermediate in the PrP conversion is in line with the finding that the transition of the recombinant prion protein to a scrapie-like form is strongly promoted in the presence of relatively low concentrations of urea, i.e. under conditions that would increase the population of an intermediate state (Morillas et al., 2001). In contrast, conditions favoring the native state (no denaturant), or those that shift equilibrium toward the fully unfolded state (high concentration of urea), are not conducive to the conversion reaction (Morillas et al., 2001). It should also be noted that the in vitro transition of the recombinant PrP to a scrapie-like form appears to be especially effective at low pH (Swietnicki et al., 1997; Hornemann and Glockshuber, 1998). This correlates well with the observation that acidic pH stabilizes the intermediate during the prion protein folding reaction. The PrPc→PrPSc conversion in mildly acidic conditions has implications for prion disease pathogenesis since it is reported that this conversion may take place in acidic compartments of the cell (Caughey et al., 1991a,b).

Amyloid fibril formation

The reason for the toxicity of PrPSc is unknown. Recent in vitro evidence provides a clue since high-pressure treatment can trigger formation of β-sheet-rich, protease-resistant, insoluble prion protein structures known as amyloid aggregates (Torrent et al., 2004). At a low protein concentration range and pressures greater than 450 MPa, a metastable intermediate structure, probably a soluble oligomer, is formed (Torrent et al., 2003). At a higher protein concentration (typically 2 mg/mL), the intermediate undergoes further conformational changes and aggregates irreversibly, with the hydrophobic parts of the prion protein becoming water exposed. The protein part undergoing this structural change is probably the region containing residues 90-167, which encompasses helix 1 and the flexible unstructured N-terminal part of the molecule. It appears that high pressure does not completely unfold PrP, but instead induces a misfolded conformation containing water-exposed hydrophobic domains. At a sufficiently high protein concentration (>2 mg/mL) intermolecular collapse of these domains eventually provokes protein aggregation (Torrent et al., 2004).

Fluorescence and electron microscopy, suggests that the aggregates, formed after transient high pressure treatment and decompression, can be classified as pre-amyloid structures with a strongly increased β-sheet content with respect to the native form, and an increased resistance to proteolysis (similarly to the infectious PrPSc form). It is therefore possible that the pre-amyloid form contributes to pathogenesis or is itself infectious (Torrent et al., 2004).

Different proteins do not however, follow the same unfolding pathways (Balny et al., 2002), resulting in different unfolded states which, in some cases, have the tendency to aggregate. Amyloid propagation is normally highly sequence specific. In PrP, this specificity leads to the “species barrier” phenomenon, which obstructs or slows down transmission (Scott et al., 1993; Telling et al., 1995; Gabizon and Taraboulos 1997; Raymond et al., 1997; Baskakov, 2004). Accordingly, a specific pattern of molecular interactions, rather than nonspecific interactions, play a role in the formation of amyloid structures. It has been demonstrated by Van Eldik et al., (1989) that side-chain hydrogen bonds and aromatic stacking play an important role in the stabilization of amyloid structures. They suggested that stacking interactions and/or stabilized hydrogen bonds between newly formed β-sheet strands may contribute energetically as well as directionally in the assembly of well-packed amyloid structures. More research is required to establish the presence and the role of these structures in TSE related neurodegenerative diseases.

Alternate allelic forms of PrP affect amyloid formation and disease susceptibility

The human PrP gene (PRNP) has two common
allelic forms that encode either methionine or valine at codon 129. This polymorphism modulates disease susceptibility and the clinico-pathological phenotypes in sporadic and acquired prion diseases (Collinge et al., 1991; Palmer et al., 1991; Brown et al., 2001; Mead et al., 2001) and may affect age of onset (Baker et al., 1991; Cervenakova et al., 1998; Parchi et al., 1999). All tested cases of variant CJD, to date, have been homozygous for codon 129-MetMet which is in contrast to an overall prevalence of 37% for this genotype in Caucasian populations, indicating a strong genetic risk factor. It is, however, possible that cases with an alternative codon 129 genotype may occur in the future as variations at this locus may influence incubation period (Parchi et al., 1999).

The molecular mechanisms by which the effect of this polymorphism is mediated remain unclear (Baskakov et al., 2005). Residue 129 is within the first β-strand (residues 128–131) in human PrP and might be a ‘nucleation site’ for a conformational transition from PrPC to PrPSc involving the loops connecting the β-sheet to the first helix (Fasman 1998; Santini et al., 2003). The folding, dynamics and stability of the physiological, α-helix-rich form of recombinant PrP are not affected by codon 129 polymorphism. In contrast, the misfolding pathway leading to the formation of β-sheet-rich, oligomer is favoured by the presence of methionine, compared with valine, at position 129. An alternative folding pathway, of amyloid formation, is favoured by the valine 129 allelomorph, forming amyloid fibres with a considerably shorter lag phase than the methionine 129 allelomorph whether under spontaneous conditions or when seeded with pre-formed amyloid fibres. Taken together, these studies demonstrate that the effect of the codon 129 polymorphism depends on the specific misfolding pathway and on the initial conformation of the protein.

The inverse propensities of the two allelomorphs to misfold as shown in vitro through the alternative oligomeric and amyloidogenic pathways could explain some aspects of prion diseases linked to this polymorphism such as age at onset and disease incubation time (Fasman 1998; Santini et al., 2003; Baskakov et al., 2005; Tahiri-Alaoui and James, 2005).

The codon 129 polymorphism may also influence the clinical phenotype (including age of onset and illness duration) of genetic prion diseases – ie diseases arising from a mutation in the PrNP gene. Most notably, the D178N mutation gives rise to a clinical picture of genetic CJD when associated with 129-Val on the mutant allele and yet results in FFI when associated with 129-Met.

The converse propensities of the 129-Met and 129-Val allelomorphs to fold into β-oligomer and amyloid forms in vitro, respectively, is highly intriguing, given the effect of the polymorphism on disease susceptibility in vivo. The differences in the misfolding behaviour between 129-Met and 129-Val could form the basis of human TSEs, particularly sporadic CJDs, that are associated with codon 129 variations such as age at onset of the disease and incubation time (Laplanche et al., 1994; Parchi et al., 1999; Schulz-Schaeffer et al., 1996; Windl et al., 1996; Alperovitch et al., 1999; Hauw et al., 2000). Reports mainly suggest that the age at onset is lower and the disease duration longer in 129-Val cases. This effect appears to be replicated in animal models of human disease: two out of three transgenic mice expressing human PrP 129-Val showed a shorter incubation period (Hill et al., 1997) as compared with transgenic mice expressing human PrP Met129 (Asante et al., 2002) when challenged with a matching genotype inocula.

Of concern are the findings in transgenic studies that a large proportion of infected animals develop subclinical disease (Asante et al., 2002; Wadsworth et al., 2004). Taken together with results from a large prevalence study in humans (Hilton et al., 2004) which suggests the presence of subclinical disease, it is possible that these individuals could act as a reservoir of carriers in the population and potentially cause further cases of secondary transmission which could result in vCJD persisting in the population at low levels for many years.

Conclusion

Although the cellular prion protein is remarkably well conserved in mammals, discerning the function(s) of this small protein has proved challenging. Indeed, PrPC is a case study in protein structural anomolies. There may now be sufficient experimental ‘fragments’ to begin assembling the larger picture of PrPc function, and to understand the origins and pathogenesis of prion disorders. This, along with developments in immunotherapy, gene therapy and proteomics, may allow design of intervention strategies to halt the progression of prion diseases.

References


Prion protein misfolding and Creutzfeldt-Jakob disease (CJD)


Accepted April 2, 2007