Comparative analysis of the prion protein (PrP) gene in cetacean species

This is the author's manuscript

Original Citation:

Availability:
This version is available http://hdl.handle.net/2318/40432 since

Terms of use:
Open Access
Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)
Comparative analysis of the prion protein (PrP) gene in cetacean species

Pier Luigi Acutis a,⁎, Simone Peletto a, Elena Grego b, Silvia Colussi a, Maria Vittoria Riina a, Sergio Rosati b, Walter Mignone c, Maria Caramelli a

a CEA (National Reference Centre for TSEs), Istituto Zooprofilattico Sperimentale del Piemonte, Liguria e Valle d’Aosta, via Bologna 148, 10154 Turin, Italy
b Department of Animal Productions, Epidemiology and Ecology, Faculty of Veterinary Medicine, University of Turin, Grugliasco, Italy
c Department of Imperia, Istituto Zooprofilattico Sperimentale del Piemonte, Liguria e Valle D’Aosta, Imperia, Italy

Received 29 August 2006; received in revised form 4 December 2006; accepted 18 December 2006
Available online 12 January 2007
Received by M. Di Giulio

Abstract

The partial PrP gene sequence and the deduced protein of eight cetacean species, seven of which have never been reported so far, have been determined in order to extend knowledge of sequence variability of the PrP genes in different species and to aid in speculation on cetacean susceptibility to prions. Both the nucleotide and the deduced amino acid sequences have been analysed in comparison with some of the known mammalian PrPs. Cetacean PrPs present typical features of eutherian PrPs. The PrP gene from the species of the family Delphinidae gave identical nucleic acid sequences, while differences in the PrP gene were found in Balaenopteridae and Ziphidae. The phylogenetic tree resulting from analysis of the cetacean PrP gene sequences, together with reported sequences of some ungulates, carnivores and primates, showed that the PrP gene phylogenesis mirrors the species phylogenesis. The PrP gene of cetaceans is very close to species where natural forms of TSEs are known. From an analysis of the sequences and the phylogenesis of the PrP gene, susceptibility to or occurrence of prion diseases in cetaceans can not be excluded.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Transmissible spongiform encephalopathies; Maritime mammals; Polymorphism; Multiple sequence alignment; Phylogenetic tree

1. Introduction

The prion protein (PrPEC) is a cellular glycoprotein formed by a largely unstructured and flexible N-terminus and by a C-terminus forming a more rigid globular domain, with regions of secondary structure (three α-helices and a short, two-stranded, anti-parallel β-sheet). A segment of usually five or six repeats is incorporated in the N-terminal region (Hunter et al., 1994; Premzl et al., 2000). A disulfide bridge and two variably occupied N-linked glycosylation sites are involved in stabilization of the C-terminal domain (Maiti and Surewicz, 2001; Rudd et al., 2001). PrPEC is attached to the cell membrane surface by a glycosyl-inositol-phospholipid (GPI) anchor (Stahl et al., 1990). The function of PrPEC is not clear, but it seems to be implicated in copper binding (Brown et al., 1997) and copper transport and metabolism (Pauly and Harris, 1998). Imidazol nitrogens of the histidines in the repeat region and histidines at positions 96 and 111 are thought to bind Cu2+ ions, thus adding structure to the flexible repeat region (Viles et al., 1999). PrPEC is expressed in most tissues, but the highest levels are found in the central nervous system, notably associated with synaptic membranes. PrPEC is encoded by the PrP gene (PrP) which is present in a single copy and contains the entire open reading frame (ORF) within a single exon (Oesch et al., 1985; Martin et al., 1995; Prusiner, 1991).
avians, reptiles and fish. PrP\(^{C}\) is linked to the occurrence of transmissible spongiform encephalopathies (TSEs) or prion diseases, a group of fatal neurodegenerative diseases that affect humans and other mammals. Prion diseases are characterized by the post-translational conversion of the PrP\(^{C}\) to a conformationally abnormal isofrom (PrP\(^{Sc}\)) (Weissmann, 1999) that accumulates in the nervous system. PrP\(^{Sc}\) is characterized by an increased \(\beta\)-sheet content and by an increased resistance to proteases (Prusiner, 1982). It has been shown that only organisms expressing PrP\(^{C}\) are susceptible to TSEs (Bueler et al., 1993). A molecule other than PrP, designated “protein X”, is believed to be involved in the conversion of PrP\(^{C}\) to PrP\(^{Sc}\) (Telling et al., 1995). Natural prion diseases may occur as genetic, infectious or sporadic disorders in a variety of mammals, most notably in humans, sheep, cattle and deer. TSEs include Creutzfeldt–Jakob disease (CJD) in humans, scrapie in sheep and goats, chronic wasting disease (CWD) in deer and bovine spongiform encephalopathy (BSE) in cattle. The BSE agent crossed the species barrier, causing feline spongiform encephalopathy (FSE) in domestic and wild cats (Kirkwood and Cunningham, 1994) and variant CJD in human beings (Aguzzi et al., 2001). The species barrier is an important aspect in the risk of prion oral transmission. Studies using transgenic mice have shown that the species barrier might be related to the degree of sequence homology between the prion protein in different species (Scott et al., 1993), in both the amino acid sequence of the protein and its tridimensional conformation (Prusiner, 1998). This indicates that certain polypeptide segments of the prion protein seem to have a particularly large influence on the species barrier (Schätzl et al., 1995; Billetter et al., 1997; Vorberg et al., 2003) and that variability at these sites can create a “host barrier”, even between related species (Kocisko et al., 1995; Telling et al., 1994; Horiuchi et al., 2000).

In human beings, sheep and deer, the expression of prion diseases is influenced by polymorphisms of the host’s prion protein gene. In humans, inherited TSEs are caused by disease-associated coding mutations and insertions of different numbers of octarepeat units. In sporadic CJD, the polymorphism at codon 129, resulting in either methionine or valine, has a profound influence on susceptibility and phenotypic expression of this disease (Johnson and Gibbs, 1998). Methionine homozygosity at the same codon has also been found in all clinical cases of variant CJD (Peden and Ironside, 2004).

In sheep, amino acid changes at codons 136, 154 and 171 have been shown to be associated with susceptibility to scrapie: PrP allelic variants valine/arginine/glutamine (VRQ) and alanine/arginine/glutamine (ARQ) at codons 136, 154, 171, respectively, are generally associated with high susceptibility to scrapie, whereas the ARR allele has been linked to decreased susceptibility or even resistance (Belt et al., 1995; Bossers et al., 1996). In elk, the Met/Met132 genotype has been indicated as a predisposing factor to both natural and experimental CWD (O’Rourke et al., 1999; Hamir et al., 2006).

The PrP gene has been characterized in various species of mammals and birds (Gabriel et al., 1992; Martin et al., 1995; Wopfner et al., 1999; Zhang et al., 2002; van Rhee et al., 2003; Lysek et al., 2004; Wu et al., 2006), and corresponding cDNA has been identified in the turtle (Simonic et al., 2000) and Xenopus laevis (Strumbo et al., 2001). Analysis of amino acid sequences showed a high level of identity within mammals and within birds, whereas between birds and mammals the overall identity is low (Wopfner et al., 1999). Cloned and sequenced peacock and parakeet prion genes show conservation of the structural features of all known mammalian and avian PrPs (Yang et al., 2005). Different cDNAs coding for homologs of tetrapod PrP\(^{C}\) have been identified also in Fugu rubripes (Suzuki et al., 2002; Rivera-Milla et al., 2003), Atlantic salmon (Salmo salar) (Oidtmann et al., 2003) and zebrafish (Danio rerio) (Cotto et al., 2005).

Even though PrP sequences for many species are available from public databases, the PrP gene sequence of cetaceans has not yet been well characterized. Sequences for only two species are available on GenBank: the complete PrP open reading frame of Tursiops truncatus (bottle-nose dolphin) consisting of 774 nucleotides encoding 257 amino acids (GenBank acc. no. AY964056) and the partial PrP sequences of Physeter catodon (sperm whale) (GenBank acc. nos. AY133054 and AF117311). Moreover, cetaceans suffer from a lack of phylogenetic information that also mirrors the paucity of data on other aspects of their biology. This is illustrated by the fact that 56% of delphinids and 75% of ziphids are categorised as data deficient in the International Union for Conservation of Nature and Natural Resources (IUCN) Red List of threatened species (www.redlist.org).

To extend knowledge of PrP gene sequence variability in different species and to aid in speculation on cetacean susceptibility to prions, we have determined and analysed the prion genes of eight cetacean species. Both the nucleotide and the deduced amino acid sequences have been compared with some known mammalian PrPs.

2. Materials and methods

2.1. Molecular biology

Genomic DNA was isolated from 25 mg of frozen muscle tissue of Balaenoptera physalus (finback whale, 1 specimen), Ziphius cavirostris (Cuvier’s beaked whale, 1 specimen), T. truncatus (bottle-nose dolphin, 2 specimens), Steno bredanensis (rough-toothed dolphin, 1 specimen), Grampus griseus (Risso’s dolphin, 2 specimens), Globicephala melas (long-finned pilot whale, 1 specimen), Delphinus delphis (saddleback dolphin, 1 specimen) and Stenella coeruleoalba (striped dolphin, 4 specimens) using the DNeasy Tissue Kit (Qiagen). All the samples were collected from animals stranded on the Italian coasts. Liver and cutis/subcutis tissues were available for some species and total DNA was purified also from these matrices for data confirmation. DNA concentration was determined by spectrophotometric absorbance measurement.

The polymerase chain reaction (PCR) amplification of the PrP gene was performed in two independent laboratories. Following a multiple alignment of over 80 mammalian, avian, reptilian, amphibian and fish PrP sequences published on the GenBank database, two degenerate primers were designed on
highly conserved regions coding for the N- and C-terminal signal peptides in mammals. The sequences of these primers, which were obtained from Invitrogen (Carsbad, USA), were: WhaleFOR (5′-TGGATCTCTGGTTTGG-3′) and WhaleREV (5′-AATGAGGAARGADTASAGGG-3′) hybridising on the target PrP DNA at corresponding nucleotides 19–40 and 723–744 of human PrP ORF (GenBank acc. no. M13899), respectively. Standard conditions for the PCR reactions were: 150 ng of genomic DNA, 50 pmol of each primer, 100 μM of dNTPs (Fermentas), 1 unit of Taq polymerase (HotStarTaq, Qiagen) in a final volume of 50 μl of buffer consisting of Tris-Cl, KCl, (NH4)2SO4 and 1.5 mM MgCl2. Thermocycling parameters consisted of an initial denaturation step (95 °C, 15 min) followed by 40 cycles of denaturation (94 °C, 1 min), annealing (58 °C, 1 min) and extension (72 °C, 1 min) on temperature gradient cyclers (MyiQ, Bio-Rad Laboratories; GeneAmp 9700, Applied Biosystems).

PCR products with the appropriate length were directly purified by the silica column method (QIAquick PCR Purification Kit, Qiagen) or isolated by agarose gel electrophoresis and purified by silica-gel-based filtration (QIAquick Gel Extraction Kit, Qiagen).

2.2. Cloning and sequencing

The PrP gene sequence was determined by direct DNA sequencing on both strands of the PCR products by Big Dye terminator cycle-sequencing using the amplification primer pair, and analysed on an ABI Prism 3130 Genetic Analyser (Applied Biosystems) according to the manufacturer’s protocol. The purified PCR products of B. physalus and one S. coeruleoalba were cloned into pDRIVE cloning vector (Qiagen) and five recombinant clones were double strand sequenced. The PrP gene sequences of the maritime mammals were analysed using the Lasergene package (DNASTAR Inc., Madison, USA) and aligned with the consensus sequence of other known mammals available in the GenBank database. The newly determined cetacean PrP sequences have been deposited in the GenBank database under accession numbers DQ884467–DQ884475.

2.3. Phylogenetic analysis

Genetic distances were computed using MEGA (Kumar et al., 2004). Distance matrices were determined under the assumptions of Kimura’s two-parameter model and were used to infer dendrograms by the neighbor-joining method (Saitou and Nei, 1987). Confidence values for individual branches of the resulting tree were determined by bootstrap analysis with 1000 replicates (Felsenstein, 1981).

3. Results and discussion

3.1. PrP gene sequence and the encoded prion protein

In this work, the partial PrP gene sequence (688 bp, including the ORF of the complete mature protein) of eight cetacean species was determined. Except for those of the bottlenose dolphin, they are novel and expand the collection of known prion genes. Most of these newly analysed species belong to the suborder Odontoceti, which includes the family Delphinidae (5 species) and the family Ziphiidae (1 species). B. physalus is also included in this study and belongs to the suborder Mysticeti, family Balaenopteridae. The resulting PrP amino acid sequences are reported in Fig. 1, aligned with other mammalian PrPs (numbering system correlates to the human PrP). The PrP gene from the species of the family Delphinidae gave strikingly identical nucleic acid sequences, except for the G. griseus sequence that was heterozygous for a non-coding mutation at codon 69 (cac→cat). This inter-species homology was confirmed by multiple DNA extractions and PCR reactions in two different laboratories. The homology can be explained by the relatively recent evolutionary divergence of the family Delphinidae: dolphins evolved during the Miocene about 10 million years ago. This separation during the evolution of the families Delphinidae, Balaenopteridae and Ziphiidae can also explain the fact that differences in the PrP gene were found in B. physalus and Z. cavirostris.

3.1.1. N-terminal region

The N-terminal region (proteinase K-sensitive portion of PrP) includes amino acids 1–91 and comprises the signal peptide (amino acids 1–22) and the repeat region (amino acids 51–91).

In the signal peptide, B. physalus carried a cysteine at codon 16, encoded by the triplet tgc, while all the other maritime mammals presented a tryptophan (triplet tgg), like the other considered mammalian species. Codon 19 was characterized by the presence of a methionine, which appeared to be typical of cetaceans: in other mammals it has been reported only in the Asian elephant (van Rheede et al., 2003). At position 21 most mammals carry a leucine, while members of the family Delphinidae a phenylalanine. Codon 21 is located between the consensus residues for the cleavage enzyme and seems to be well conserved in mammals. Along with some other species, dolphins represent an exception: the same substitution has been detected in the PrP gene of the mink and the ferret (Wopfner et al., 1999), while valine replaces leucine in the manatee (van Rheede et al., 2003). An L21P polymorphism has been described in caprine PrP (Billinis et al., 2002). The residues flanking the signal peptide cleavage site, between cysteine 22 and lysine 23, were perfectly conserved in cetaceans; the presence of glycine 20 and cysteine 22 before the cleavage site agrees with the consensus residues for the cleavage enzyme (Udenfriend and Kodukula, 1995). Cetacean PrP conserved proline at position 44; this residue undergoes 4-hydroxylation and represents an important functional feature of the PrP N-terminal region (Gill et al., 2000). Nevertheless, several mammalian species lack this residue, raising questions as to the universal importance of this modification (van Rheede et al., 2003).

Analysis of the repeat region showed that the number of repeats was five in all examined maritime mammal species. One S. coeruleoalba specimen showed a polymorphism resulting in
Homo sapiens
Macaca mulatta
Oris aries
Bos taurus
Capra hircus
Cervus elaphus
Hippopotamus amphibius
Sus scrofa
Mus of vison
Felis catus
Canis familiaris
Physeter catodon
*Selaenoptera physalus
*Ziphius cavirostris
*Stenella coeruleoalba
*Grampus griseus
*Globicephala melas
*Delphinus delphis
*Steno bredanensis
*Tursiops truncatus
*Tursiops truncatus
*Tursiops truncatus
*Tursiops truncatus

<table>
<thead>
<tr>
<th>Homo sapiens</th>
<th>Macaca mulatta</th>
<th>Oris aries</th>
<th>Bos taurus</th>
<th>Capra hircus</th>
<th>Cervus elaphus</th>
<th>Hippopotamus amphibius</th>
<th>Sus scrofa</th>
<th>Mus of vison</th>
<th>Felis catus</th>
<th>Canis familiaris</th>
<th>Physeter catodon</th>
<th>Selaenoptera physalus</th>
<th>Ziphius cavirostris</th>
<th>Stenella coeruleoalba</th>
<th>Grampus griseus</th>
<th>Globicephala melas</th>
<th>Delphinus delphis</th>
<th>Steno bredanensis</th>
<th>Tursiops truncatus</th>
<th>Tursiops truncatus</th>
<th>Tursiops truncatus</th>
<th>Tursiops truncatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>L...N...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td>L...Y...V...Q...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Cetacean PrP sequences aligned with other mammalian PrPs. The numbering system correlates to the human PrP (upper line). Points indicate identical residue; deletions are indicated by dashes. Species are named at the left. Sequences marked by an asterisk (*) have been newly determined in this study. 6 repeat variant of *Stenella coeruleoalba* is marked by number (1). References to the Genbank sequences used for the alignment: *Homo sapiens* M13899, *Macaca mulatta* NM001047151, *Oris aries* M31313, *Bos taurus* DQ205538, *Capra hircus* X91999, *Cervus elaphus* AY748455, *Hippopotamus amphibius* AY133053, *Sus scrofa* NM001008687, *Mus of vison* S46825, *Felis catus* AY573555, *Canis familiaris* AF042843, *Physeter catodon* AY133054, *Tursiops truncatus* (2) DQ130070, (3) DQ130069, (4) AY964056, (5) AF117311.
heterozygosity for alleles with five and six repeats. The PCR product showed a double band in ethidium bromide-stained 2.8% agarose gel; cloning and sequencing revealed that a 24 bp extra repeat was inserted between the second and third repeat units of the wild-type sequence. The additional repeat had the same nucleotide sequence of the second repeat unit. Repeat length polymorphism is a frequent mutational process in the eutherian prion gene. Events leading to expansion and contraction of repeats are thought to involve unequal crossing-over and replication slippage (Collinge, 2001). Extension of the normal number of five repeats with one to nine copies has been observed in human prion disease kindreds, and addition of repeat units promotes the early onset of prion disease (Collinge, 2001). A reduction in the repeat number to four does not lead to prion disease, but an elderly patient suffering from a progressive dementia consistent with CJD was reported to carry heterozygosis for three repeats (Beck et al., 2001). A variant containing only three instead of the usual five copies of repeats has been described in goats (Goldmann et al., 1998). One goat heterozygous for the polymorphism and challenged experimentally with scrapie succumbed after an unusually long incubation period. In cattle, three PrP isoforms are known, ranging from five to seven repeats, unlike associated with BSE (Hunter et al., 1994; Neibergs et al., 1994; Humeny et al., 2002; Sander et al., 2004). Homozygosis for four or seven repeats occurs in several other mammalian species and alleles, with two repeats quite common in the lemur and the squirrel (van Rheede et al., 2003), even if it is not known whether homozygotes are viable. Cetacean repeats respected the eutherian consensus sequence P(Q/H)GGG(G/-)WGQ: the first repeat had Q at position 2 and a GGGG run and the following ones had H at position 2 and GGG. In the examined cetaceans, a GGGG track occurred in the last repeat; indeed, this is a common finding also in some terrestrial mammal PrPs and is indicative of repeat homogenization between maritime and terrestrial mammals. A truncated repeat is inserted in *P. catodon* (after position 91). In this species, the deletion may have been triggered by glycine runs on both sites of the WGQ triplet, as present in the last repeats in other placentalts (van Rheede et al., 2003). Such a feature was not present in the other related members of the suborder Odontoceti considered in this study, suggesting that the deletion occurred after the divergence of the family Physeteridae. The nucleotide sequence of the cetacean repeat region is shown in Fig. 2. The histidine residues of the internal octarepeats were generally encoded by cac, but cat was used in the second repeat of *Z. cavirostris* and in the third repeat of *P. catodon* and *G. griseus*. Triplets coding the first glycine showed a conserved behaviour respecting the same run in all the examined species: gga was used in the first repeat and ggt in the following ones. The ggg triplet encoded the second glycine in the first nona-repeat, while ggt encoded the same amino acid in the other units, but gcc was also used in *B. physalus*. The third glycine triplet in all the examined cetaceans changed with repeat length, being gga or gcc according to the presence of nine or eight residues, respectively. This variation can be associated with the fact that in mammals almost all glycine residues preceding the tryptophan are encoded by gcc (van Rheede et al., 2003). The triplet coding for the last glycine was quite inhomogeneous: ggt, gga and gcc were used, even if the third and fifth repeats showed conservation of ggt in all the considered species. Similarly, the terminal glutamine residue was usually encoded by cag, but it could also be encoded by caa except in *P. catodon*.

### 3.1.2. C-terminal region

In the C-terminal region (the proteinase-resistant part of PrP), important functional residues are present, like the

![Fig. 2. Nucleotide sequence alignment of the five to six repeats in cetacean species. Sequences marked by asterisk (*) have been newly determined in this study. The six repeat variant of *Stenella coeruleoalba* PrP is marked by number (1). The symbol “#” indicates the extra repeat of *Stenella coeruleoalba* six repeat variant. The additional repeat has the same nucleotide sequence of the second repeat.](image-url)
hydrophobic region (amino acids 115–126), the glycosylation sites, two cysteines forming a disulfide bridge and a serine at position 231, binding the GPI anchor. This region is structured by the presence of α-helices and β-strands.

Next to the repeat region, maritime mammal sequences revealed a deletion of codon 100, which led to a cetacean variant of the motif QWXKP (positions 98–102) that is well conserved in mammals. The only other species presenting this deletion is the hippopotamus: the molecular relationship of whales with the hippopotamus and ruminants in Cetartiodactyla is well-established now, although the existence of a whale–hippo clade is still debated among morphologists (Gatesy and O’Leary, 2001; Boisserie et al., 2005). Lysine 110 was present in all the analysed cetacean species. This amino acid seems to be very conserved in all mammalian species; this could be consistent with its supposed functional role as a post-translational cleavage site of PrPC (Harris, 1999). According to our findings, the PrP of Z. cavirostris was the only cetacean PrP with the same amino acid (methionine) at position 112 as in humans. The presence of methionine or valine might play a central role in species-barrier restrictions (Schätzl et al., 1995). Moreover, residue 112 is important in TSE diagnosis since it is involved in the formation of the epitope for antibody 3F4. This specific antibody has been shown to bind exclusively to PrPSc with a methionine at position 112 (Bolton et al., 1991; Kanyo et al., 1999).

The hydrophobic region was conserved in all cetacean sequences reported in this study. The hydrophobic stretch is perfectly conserved also in avian and reptiles PrP and is considered a PrP signature. It probably plays a key role in PrPC conversion to PrPSc (Salmona et al., 1999). Deletion of the N-terminus is able to prevent conformational transition to PrPSc only if it includes this hydrophobic region (Muramoto et al., 1996).

At position 143, situated immediately before α-helix 1 (H1) and believed to be implicated in human susceptibility cattle-derived prion (Krakauer et al., 1996), a serine residue was present in all species, relating cetaceans to most ruminants of the subfamily Bovinae and to species like the pig, the hippopotamus and the rabbit. A substitution unique to B. physalus was present at codon 159 (Asn-Ser), in close proximity to regions of PrP with a secondary structure (starting of β-sheet 2). Remarkably, this codon is usually well conserved in most species; an identical polymorphism has been identified in the squirrel monkey (van Rheede et al., 2003), and another coding mutation (Asn-Asp) was found in some canine species, including the domestic dog (Wopfner et al., 1999). Sequence position 159 of canine PrP is likely a candidate in association with positions 103 and 177 for contributing to the implicated species barrier for TSE transmission from cattle to canine species (Lysek et al., 2004). B. physalus PrP sequence carried a serine at position 171, the same substitution has been detected...
as a polymorphism in some human populations (Mead et al., 2003). At position 174, structurally the second amino acid of the \( \alpha \)-helical 2 (H2) region, all the analysed cetaceans showed a serine also found in the pig, banteng (Bos javanicus), lama, camel and rabbit. The cysteines at positions 179 and 214, forming the disulfide bridge, were well conserved in cetaceans, as in all mammals, birds and turtles. The disulfide bridge is essential for stabilizing the PrP structure. Substitutions of these residues result in insolubilization of the protein (Maiti and Surewicz, 2001). The two N-glycosylation sites at codon 181 and 197 (Rogers et al., 1990) were conserved in all cetaceans. N-linked sugars preserve the extracellular PrP surface from proteases and non-specific protein interactions (Rudd et al., 2001). The reported presence of lysine instead of asparagine at position 181 (Wopfner et al., 1999; Yanming et al., 2006) in T. truncatus was not present in the two bottle-nose dolphins analysed in the present study. These results suggest that position 181 is polymorphic in the bottle-nose dolphin. The frequency of the 181N and 181K alleles is not known, but we can speculate that neither is rare if they have been found in a small number of animals. To the authors’ knowledge, there are no other species with a PrP polymorphism causing the loss of the first glycosylation site, so it would be interesting to investigate its functional consequences. Transgenic mice carrying the PrP substitution threonine for asparagine 181, which eliminates the first glycosylation site, show that the lack of glycans does not influence PrP maturation and stability and that the presence of one sugar chain is sufficient for trafficking to the cell membrane (Cancellotti et al., 2005). In the other cetacean species examined in this study, the N-linked glycosylation sites were found to be conserved. Nonetheless, the overall identity of the PrP genes of the family Delphinidae does not rule out the possibility of the presence of this polymorphism also in other species: collecting further data on a larger number of samples could indicate whether it has a specific significance in marine mammals. Two silent mutations at codon 188 (aag→aca) and 201 (aag→aca) were found in heterozygosis in B. physalus: cloning showed that they were located on two different haplotypes. At positions 190–193 four threonine residues are present in the prion protein of all mammals so far characterized, while the cetacean sequences showed the insertion of a fifth threonine. This insertion localizes in the H2 region between the two glycosylation sites, a conserved region outside the structural domain which could be important for the conversion process of PrP isoforms (Wopfner et al., 1999). This variation could raise questions about the three-dimensional structure and flexibility of the prion protein in cetaceans. Isoleucine occurred at position 203 of all determined cetacean PrP sequences. The CJD-related mutation V203I seems to be an extremely rare polymorphism in humans (Peoc’h et al., 2000) while most eutherian PrP genes carry this variant (van Rheede et al., 2003). Position 212, implicated in Gerstmann–Straussler–Scheinker disease (GSS) in human (Q212P) (Piccardo et al., 1998), was not conservative in Z. cavirostris. Variant 212E occurred in this cetacean species; the same replacement has been observed exclusively in some microbats so far. The prion protein sequence of Z. cavirostris was the only identified cetacean gene carrying a threonine at codon 215. Members of orders Rodentia (i.e. Syrian hamster), Macroscelidea and Chiroptera are also known to carry this amino acid. A polymorphism involving codon 215 has been described in the goat (I215L) (Zhang et al., 2004). Position 215 forms with 168, 172, and 219 the epitope for the binding of protein X. An arginine residue was present at the highly variable position 220 in all the examined cetaceans. Position 223, immediately preceding the GPI anchor site, showed variability in cetacean species. Members of the family Delphinidae carried a glutamic acid residue, while a glutamine occurred in the species belonging to the families Balaenopteridae, Ziphiidae and Physeteridae. This region forms a flexible linkage to the GPI anchor (Riek et al., 1996; Liu et al., 1999). The serine residue at position 231, where the GPI anchor is attached, was strictly conserved in all cetaceans.

3.2. Phylogenetic analysis of cetacean PrP gene

Expanding the spectrum of mammalian prion genes may improve our ability to reconstruct ancient phylogenetic nodes and enhance our understanding of fixed mutational events during the evolution of the prion protein. Fig. 3 shows the phylogenetic tree resulting from an analysis of the cetacean PrP gene sequences, together with reported sequences of some ungulates, carnivores and primates. PrP gene phylogenesis mirrored species phylogenesis supporting the association of cetaceans with the hippopotamus and ruminants in the order Cetartiodactyla. Analysis of additional species will help to reconstruct the prion sequence at the time that cetaceans diverged from artiodactyls 60 millions years ago.

3.3. Conclusions

- This work adds so far unreported PrP sequences of cetaceans to the collection of known prion protein genes.
- Cetacean PrPs present typical features of eutherian PrPs.
- PrP gene evolution follows the evolution of the species.
- The PrP gene of cetaceans is very close to species where natural forms of TSEs are known.
- From an analysis of the sequences and the phylogenesis of the PrP gene, susceptibility to or occurrence of prion diseases in cetaceans cannot be excluded.

Acknowledgements

We would like to thank the Mediterranean Marine Mammals Tissue Bank of the Department of Experimental Veterinary Sciences (University of Padua) and especially Prof. Bruno Cozzi and Dr. Camilla Butti for sample contribution.

References


