January 30, 2009

Evolutionary descent of prion genes from a ZIP metal ion transport ancestor

Gerold Schmitt-Ulms[1,2,*], Sepehr Ehsani[1,2], Joel C. Watts[1,2,3],
David Westaway[5], and Holger Wille[3,4]

[1] Centre for Research in Neurodegenerative Diseases, [2] Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON M5S 3H2, Canada.
[3] Institute for Neurodegenerative Diseases, [4] Department of Neurology, University of California San Francisco, CA 94143, USA.
[5] Centre for Prions and Protein Folding Diseases, University of Alberta, Edmonton, AB T6G 2M8, Canada.

*To whom correspondence should be addressed:

Centre for Research in Neurodegenerative Diseases,
University of Toronto, Room 209, Tanz Neuroscience Building,
Toronto, Ontario, M5S 3H2, Canada.
Tel: 416-946-0066; Fax: 416-978-1878
Email: g.schmittulms@utoronto.ca

Running title: Evolutionary descent of PrP from ZIP ancestor

Abbreviations: CID, collision-induced dissociation; Dpl, Doppel; Dr, *Danio rerio*; GPI, glycosylphosphatidylinositol; HMM, hidden Markov model; HPLC, high performance liquid chromatography; IP, immunoprecipitation; iTRAQ, isobaric tagging for relative and absolute quantitation; LC, liquid chromatography; Mm, *Mus musculus*; MS/MS, tandem mass spectrometry; PL, prion-like; PrP^{C}, cellular prion protein; PVDF, polyvinylidene fluoride; Sho, Shadoo; TM, transmembrane; Tr, *Takifugu rubripes*; ZIP, Zrt-, Irt-like Protein.

ABSTRACT

In the more than 20 years since its discovery, both the phylogenetic origin and cellular function of the prion protein (PrP) have remained enigmatic. Insights into the function of PrP may be obtained through a characterization of its molecular neighborhood. Quantitative interactome data revealed the spatial proximity of a subset of metal ion transporters of the ZIP family to mammalian prion proteins. A subsequent bioinformatic analysis revealed the presence of a prion-like protein sequence within the N-terminal, extracellular domain of a phylogenetic branch of ZIPs. Additional structural threading and ortholog sequence alignment analyses consolidated the conclusion that the prion protein gene family is phylogenetically derived from a ZIP-like ancestor molecule. Our data explain structural and functional features found within mammalian prion proteins as elements of an ancient involvement in the transmembrane transport of divalent cations. The connection to ZIP proteins is expected to open new avenues to elucidate the biology of the prion protein in health and disease.

Prion diseases are fatal neurodegenerative diseases of humans and animals which, in addition to sporadic and familial modes of manifestation, can be acquired via an infectious route of propagation. Prion diseases include sporadic and variant forms of Creutzfeldt-Jakob disease in humans. In sheep, cattle, and cervids prion diseases manifest as scrapie, bovine spongiform encephalopathy, and chronic wasting disease, respectively. The normal, cellular prion protein (PrP^{C}) is expressed at high levels in the central nervous system, but can also be found in other cell types within the body. In prion disease, PrP^{C} undergoes a structural transition to its disease-causing form (PrP^{Sc}) which possesses profoundly different physicochemical properties[1].

The prion protein may constitute the most intensively studied mammalian protein. Yet surprisingly, the evolutionary origin and physiological function of this protein have remained largely elusive[2, 3]. The function of a protein can sometimes be inferred from genomic investigations which may reveal proteins with similar sequences or sequence modules of known function. Extensive investigations of this kind provided evidence for PrP-related sequences in most species of the vertebrate lineage[4-6] and revealed two mammalian paralogous sequences encoding for the proteins Doppel (Dpl) and Shadoo (Sho), which together with PrP^{C} constitute the mammalian prion protein family[7]. Despite these advances, no conclusive biological role has emerged for PrP^{C} from the characterization of these additional members of the mammalian prion protein family[8].

Alternatively, by comparing the structural features of particular proteins, otherwise cryptic functional similarities may be inferred. Investigations of PrP^{C} revealed an unusual dichotomy of its structure, which consists of an extended, largely unstructured N-terminus and a globular and relatively stable C-terminal domain formed by two β-

strands and three α-helices[9]. While proteins with mixed β-strands and α-helices are widespread in nature, aside from Sho and Dpl which share structural features with the N-terminal and C-terminal regions of PrP, respectively, no other protein has been identified to infer the physiological function of PrP^C based on structural similarities. Extensive evidence has, however, been accumulated which demonstrates that the prion protein can bind a subset of divalent metal cations through two types of histidine-containing motifs embedded in its N-terminal domain.

Most proteins do not act in isolation but partner with other proteins to exert their biological roles[10]. Thus, the function of a protein can sometimes be deduced by characterizing its binding partners. Following this 'guilt-by-association' logic, we set out to identify the function of the cellular prion protein through a comprehensive interactome investigation. Surprisingly, this work not only revealed novel interactors but also shed light on the evolutionary origins and an ancient function of members of the prion gene family.

RESULTS

Quantitative interactome analyses

Various earlier attempts by us and others had already led to the identification of a few dozen proteins that co-purify with the cellular prion protein under a range of experimental conditions[11]. Often investigations of this kind result in long lists of candidate interactors. The challenge then remains to discriminate specific from unspecific binding partners. Here we incorporated quantitative mass spectrometry based on isotopic tagging of peptides into the workflow to overcome this limitation[12]. Furthermore,

investigations were extended to all three members of the mammalian prion protein family to further facilitate discrimination of potential interactors by differential interactome comparison. A murine neuroblastoma cell line (N2a), an established cell model for prion replication, served as the biological source material[13]. The cell line was stably transfected with a member of the mammalian prion protein family that had been FLAG-tagged in the vicinity of the N-terminus of the mature protein or with a negative control vector[14]. The conceptual choice of both the FLAG-tag and N-terminal attachment site was guided by data documenting that such insertions do not interfere with either PrP^C processing or conversion in transgenic animals[15], as well as studies of glycosylation and trafficking in Dpl- and Sho-transfected cells (J. Coomaraswamy *et al.*, in preparation). To stabilize physiologically relevant interactions, adherent cells were crosslinked by a short treatment with formaldehyde prior to the cell harvest step[16]. Following cell lysis in the presence of detergents, crosslinked protein complexes were affinity-purified based on the presence of the FLAG-affinity tag, trypsinized in solution, and subjected to labeling with isobaric tags for relative and absolute quantitation (iTRAQ)[17]. The samples for the separate prion proteins and the control were mixed and jointly analyzed by tandem mass spectrometry. It should be noted that subsequent to iTRAQ conjugation, the contribution of each sample to the identification of a peptide can readily be calculated by determining the relative intensity of signature iTRAQ mass signals in the relevant collision induced dissociation (CID) spectra. A comprehensive analysis of samples led to the identification of approximately one hundred proteins, more than thirty of which were observed with iTRAQ signature mass intensity ratios that suggested specific co-enrichment with members of the prion protein family (**Supplementary Fig. 1**; a detailed description of

this large dataset will be presented elsewhere). Strikingly, two proteins, ZIP10 (*slc39a10*) and ZIP6 (*slc39a6*), which were unequivocally identified with non-overlapping peptides in Dpl, Sho and PrP samples but not in the negative control, represent members of the ZIP family (Zrt-, Irt-like Protein) of transmembrane proteins **(Table 1)**. The *slc39a* gene family codes for fourteen distinct proteins (ZIP1-14) in humans. ZIP10 and ZIP6 populate together with their phylogenetically closest paralog ZIP5 a subbranch within this family[18].

Structural and sequence similarities of mammalian prion proteins and ZIPs

Intrigued by the presence of two phylogenetically related proteins, which we had never observed in similar datasets, the proteins were subjected to an extensive bioinformatic analysis. Surprisingly, the results revealed the presence of a domain within a subset of mouse ZIP proteins that displayed a substantial sequence similarity to both PrP and Dpl. The search was conducted with a SCOP (Structural Classification of Proteins, version 1.69)[19] linear hidden Markov model (HMM; Superfamily model 0037705) that had been trained on a multiple alignment of prion protein ortholog sequences (Superfamily 54098, designated as 'Prion-like')[20]. More specifically, a 107-amino acid fragment within ZIP10 (residues 285 to 391) not only showed good general alignment with the C-terminal globular domain of mouse PrP (41% similarity), but demonstrated perfect positional agreement of both conserved cysteine residues (which form a disulfide bridge in PrP^{C}) and the glycosylation site motif 'NxT' within that domain **(Fig. 1a)** (highly similar motifs were also found in ZIP6 and ZIP5). No other SCOP linear HMM model aligned to this region (from a total of 982 different models in this SCOP release). More importantly, aside from ZIPs and classical PrP family members

no other protein among the more than 120,000 mouse and human proteins represented in the LOCATE subcellular localization database (URL: http://locate.imb.uq.edu.au)[21] is recognized by the 'Prion-like' linear HMM in this manner. To further assess the statistical significance of the sequence similarity between PrP and ZIPs within this domain, we applied a method for homology inference using profile-profile alignments embedded within an algorithm that can be initiated from the program COMPASS[22]. A query of this algorithm with the abovementioned mouse ZIP10 segment returned the SCOP40 database entry 'Prion protein domain' (designator d.6.1.1.) as the only hit (E-value = $4.51e^{-4}$) which passed the statistical significance threshold of $5e^{-3}$, thereby classifying the respective domains within ZIPs and PrP as homologous.

Proteins of common evolutionary origin are frequently characterized by similar folds. Well-studied examples include proteins of the immunoglobulin and fibronectin type III superfamilies. Such proteins are often the result of divergent evolution during which they have accumulated differences in their primary structure but continue to sustain similar folds. The comparison of protein structures therefore constitutes an orthogonal approach in studying the relationship of proteins. Whereas a multitude of high-resolution structures of prion proteins from many species have been solved, similar information is currently not available for metal ion transporters of the ZIP family. We therefore employed structural threading algorithms to assess the predicted similarity of the folds of PrP and Dpl and the respective 'prion-like' (PL) domain in ZIPs. This comparison was based on an analysis utilizing the FFAS03 fold and function assignment server[23]. The profile-based sequence alignment revealed prominent sequence similarities

between PrP, Dpl and ZIP10 **(Fig. 1a),** which allowed threading the sequence of the PL domain of ZIP10 onto the structures of both PrP and Dpl **(Fig. 1b)**.

Initially, the FFAS03 structural alignment scores with ZIP10 were just below the threshold for statistical significance with values around -9.0 and a threshold of -9.5 (with the threshold indicating less than 3% false positives). However, relatively small changes to the sequence of ZIP10 or by replacing it with its paralog ZIP5 did bring the FFAS03 structural alignment scores into the statistically significant range (with values of -9.75 or below). Consistently, the models with the strongest scores were based on alignments of PL domains of ZIP proteins with mouse PrP or Dpl structures and all attempts to align the ZIP PL domain to structural templates outside the prion protein family produced insignificant fits. The PL domains of ZIP10 (and ZIP5, not shown) are predicted to contain a structural arrangement very similar to that of PrP or Dpl with three α-helices and possibly a short β-sheet composed of two β-strands **(Fig. 1b)**. The root mean square deviation (RMSD) between backbone carbon atoms of the nuclear magnetic resonance (NMR) structures for PrP and Dpl is 3.7 Å (as determined by the DaliLite server[24]). Surprisingly, the RMSD between the predicted structure for the PL domain of ZIP10 and the PrP and Dpl structures returned even lower values of 2.6 Å and 2.9 Å, respectively, indicating that the primary structure of the ZIP10 PL domain is highly compatible with the basic prion protein fold. Consequently, this observation argues that the secondary structure elements of PrP and Dpl originated from similar structural features in the ZIP PL domain.

Convergent evolution versus common evolutionary origin

It is well-established that similar secondary and tertiary protein structures can evolve independently, as is the case with structural similarity of prokaryotic subtilisin and eukaryotic chymotrypsin[25]. However, instances of convergent evolution at the protein sequence level are rare. Whenever reported, observations appear restricted to convergent adjustments of individual amino acids rather than independent evolutionary inventions of extensive blocks of similar primary structure. A frequently cited case represents the independent adaptation of the protein lysozyme to the acidic milieu found in the digestive tracts of unrelated species[26]. To investigate whether the similarities between PrP^C and ZIP10 constitute a case of convergent evolution, we analyzed ortholog sequences of both PrP^C and ZIP10 across a wide range of species within the chordate lineage. This was based on the conception that a divergent trend in sequence similarity would be indicative of a common phylogenetic ancestry, whereas the opposite scenario, in which the highest sequence similarity is seen in phylogenetically recent species, would constitute evidence for convergent evolution between PrP^C and ZIP10 sequences. PrP^C-related sequences can be found in all mammals and most species of the vertebrate lineage. ZIP proteins date back much further and are highly conserved throughout evolution. Related sequences can be found in all kingdoms of life, including bacteria and plants. The comparison of ZIP10 ortholog sequences revealed the presence of a highly conserved 'PALLY' motif, also noted in ref. 18, within the domain resembling PrP sequences. This motif is not found in mouse or human PrP^C sequences **(Fig. 2a)**. Subsequently, the sequence alignment of PrP-1 and ZIP10 from pufferfish (*Takifugu rubripes*) (i) not only revealed multiple additional amino acid residues conserved between pufferfish ZIP10 and PrP-1 — primarily in

highly conserved sequence positions (28% identity, 43% similarity); but also (ii) demonstrated the existence of the 'PAL' half motif within pufferfish PrP-1 in the expected location; and (iii) uncovered the presence of an extra cysteine in PrP-1 missing in mammalian prion proteins but present at the extracellular boundary of the first transmembrane domain in ZIP10 sequences from various species. A similar relationship was also observed between stickleback (*Gasterosteus aculeatus*) PrP-1 and ZIP10 sequences. PrP sequences from chicken (*Gallus gallus*) and frog (*Xenopus tropicalis*), organisms that lie between fish and mammals on the evolutionary scale, gave intermediate percent identity values when aligned with the corresponding ZIP10 sequences. In light of these data, convergence could be excluded as an explanation for the similarity of ZIP and PrP^C sequences. Instead, this analysis firmly established an evolutionary common origin of ZIP transporters and PrP^C.

N-terminal duplication versus PL domain insertion

Two alternative scenarios seemed plausible for the origin of modern day prion proteins: a gene rearrangement may have caused the insertion of the PL domain of a ZIP ancestor gene into a Sho-like precursor molecule thereby giving rise to a PrP founder gene. Alternatively, the complete PrP sequence including promoter elements and N-terminal sequences could have originated from a ZIP ancestor molecule. If the latter had occurred, similarities between ZIP proteins and PrP should not be restricted to the PL domain but extend to other features. We therefore compared the known biological features of prion proteins and ZIP proteins of the phylogenetic branch comprising ZIPs 5, 6 and 10. Both PrP^C and ZIPs 6 and 10 share a widespread expression in biological tissues with high transcript levels in the brain[18]. The expression of ZIP5 is more restricted

and no ZIP5 transcripts can be found in N2a cells (data not shown). Reminiscent of N-terminal repeat motifs in PrP^C, ZIPs 5, 6 and 10 are equipped with histidine-rich repeat sequences N-terminal to their PL domains. Also reminiscent of PrP^C, the PL domains within ortholog sequences of these ZIPs are more highly conserved than their N-terminal repeat sequences. The diversification of the latter is likely aided by an increased chance of unequal crossing-over recombination events generally associated with tandem repeat regions and previously suggested for prion sequences from different species[6]. The PL domains within ZIPs 5, 6 and 10 are part of their extracellular N-terminal domains and, with regard to orientation and relative distance to their downstream membrane anchorage sites, reminiscent of PrP^C. Both PrP^C and the above ZIPs are known to bind divalent metal ions through histidine-containing motifs embedded in N-terminal repeat sequences[18, 27]. While it is unknown whether PrP^C's metal binding ability serves a purpose for sensing, scavenging or transport of divalent cations *in vivo* , a role of the prion protein in the cellular zinc homeostasis has been proposed before[28], and members of the ZIP protein family are well-characterized in their ability to transport zinc and other divalent metals across membranes. ZIPs 5, 6 and 10 reside, like PrP^C, in the plasma membrane and are expected to facilitate metal import into the cytoplasm. Taken together, this context analysis uncovered multiple additional commonalities between ZIP proteins and PrP^C consistent with the interpretation that these molecules are related beyond sequence similarities in their PL domains. Our observation that ZIP10 and ZIP6 co-purified not just with FLAG-PrP and FLAG-Dpl but also with FLAG-Sho, as evidenced by strong iTRAQ117 signals in CID spectra which led to the identification of ZIP10 and

ZIP6, lent further support to this conclusion, because FLAG-Sho does not harbor a PL domain.

To further investigate this question, we undertook a systematic comparison of N-terminal sequences within and across the two protein families. Alignments of ZIP6 and ZIP10 with their closest phylogenetic neighbor ZIP5 suggested that N-terminal expansions of histidine-rich sequences present in ZIP6 and ZIP10 emerged following their divergence from a common ancestor with ZIP5. Thus, whereas ZIP5 shares the most conserved N-terminal sequence segments, it has a much shorter N-terminus and is only equipped with a short histidine cluster. Interestingly, the subsequent alignment of a Sho2 sequence from pufferfish not only showed remarkable sequence conservation to ZIP5 from zebrafish (34% identity, 41% similarity) but close examination revealed that a group of basic residues, the position of a largely hydrophobic sequence segment and the abovementioned cluster of histidines are shared between these sequences **(Fig. 2b)**. The above sequence alignment falls on the empirical threshold which distinguishes instances of significant similarity from uncertain 'twilight zone' alignments[29]. Nonetheless, in the context of other similarities presented above, it corroborates the conclusion that the homology between ZIP and prion protein families extends beyond the globular domain present in PrP and Dpl to include sequences N-terminal to this domain, present in Sho.

Having established the relation between prion and ZIP sequences, the order in which these molecules emerged during evolution could be deduced. Given that ZIP proteins are widespread in all kingdoms of life whereas PrP^C and Sho appear restricted to species of the Chordata lineage, a parsimonious interpretation is that PrP^C, and most likely Sho, evolved from a ZIP-like ancestor molecule. To sustain a viewpoint in favor of

a reverse phylogenetic relationship between these gene families it would be necessary to invoke a scenario in which prion gene sequences disappeared independently in most kingdoms of life except in vertebrates.

Divergent development of prion and ZIP genes

An in-depth analysis of the relative timing of the evolutionary origin of prion gene family sequences is beyond the scope of this article. Our current data suggest that the initial ZIP ancestor gene duplication event took place around the time when ZIP10 sequences diverged from the lineage that gave rise to today's ZIP5[18]. At that time, sequence features of the ZIP10 PL domain which we found conserved in the globular domain of PrP^{C}, had subtly diverged from those found in ZIP5 and ZIP6. Additionally, the extracellular N-terminal domain of this ZIP ancestor probably had not expanded beyond what is seen in today's ZIP5. N-terminal tetra-repeat and octa-repeat sequences present in current Sho and PrP proteins, respectively, most likely emerged by expansion of a short sequence motif containing basic residues following their divergence from the ZIP transporter lineage. The emergence of extended histidine-rich repeat sequences within the N-terminal domain of ZIP10 and ZIP6 may have served a similar purpose in the functional specialization of these proteins **(Fig. 3a)**. Interestingly, a strikingly similar histidine-rich repeat sequence with preferential zinc binding properties can also be found in zebrafish PrP[30] **(Supplementary Fig. 2)**.

Whereas the genomic organization of genes encoding for proteins of the prion protein family has been subjected to close scrutiny previously[31], relatively little is known about regulatory elements and the exon/intron structures of the *slc39a* gene family members. Our preliminary analysis revealed that a subset of *slc39a10* ortholog genes

have in common with a subset of genes from the prion family the existence of a first exon containing non-coding regulatory sequence elements and a downstream exon in which a very short segment of non-coding sequence is followed by the respective ATG start codons. The protein coding region of all members of the prion protein family is confined to a single exon. ZIP proteins, in contrast, are encoded by multiple exons. In particular, exon/intron arrangements of ZIP gene sequences encoding for protein sequences upstream of the PL domain are most diverse and may be encoded by a single exon (e.g. mouse ZIP10) or multiple exons with non-conserved intron/exon boundaries across ortholog ZIP sequences. A more complete understanding of genomic rearrangements underlying the emergence of prion sequences from a ZIP ancestor gene will have to be based on a comparison of all aspects of the respective genomic segments in multiple organisms.

DISCUSSION

Taken together, the data presented here constitute conclusive evidence that the mammalian prion protein family evolved from a specific branch of a highly conserved family of ZIP metal ion transporters, possibly as a result of a phylogenetically ancient and partial gene duplication event. This origin places a ZIP5/ZIP10-like ancestor gene at the root of the PrP gene family. It explains the existence of prion genes in the vertebrate lineage, and explains structural and biological features within prion proteins as remnants of an ancient involvement in the sensing and/or import of metal ions present in the extracellular milieu. While it remains to be determined to what degree prion proteins have retained functional involvement in metal ion transport, the co-purification of ZIP proteins and prion protein family members documented here suggests an ability of these

proteins to inhabit the same subcellular space and to recognize each other has been conserved throughout evolution.

The availability of genomic sequences from the teleost lineage proved highly useful for this work. The presence of multiple copies of prion genes in this vertebrate branch may have helped to preserve original ZIP-like features in a subset of these genes. A phylogenetic linkage of teleost and mammalian prion genes has been established earlier[32] and therefore conclusions drawn from teleost sequence alignments are also relevant for establishing evolutionary relationships of mammalian prions. Because teleost prion related sequences diverged from the tetrapod lineage hundreds of millions of years ago — a time period during which they evolved independently — they should, however, not be considered precursors of prion sequences in the tetrapod lineage.

Whereas PrP^C is inserted into the membrane by the presence of a C-terminal glycosylphosphatidylinositol (GPI) anchor[33], ZIP5 and ZIP10 are type III transmembrane proteins characterized by a relatively large N-terminal extracellular domain, encompassing about half of the molecule's primary structure, and a C-terminal domain harboring up to eight transmembrane spanning segments **(Fig. 3b)**. Evolutionary precedents exist in other protein families in which some members are inserted into the membrane through transmembrane domains while others employ a GPI anchor for membrane attachment. This dichotomy of membrane attachment modes has, for example, been documented among members of the carcinoembryonic antigen (CEA) protein family[34] and cadherins, a diverse family of GPI-anchored (T-cadherin), single-spanning (classical cadherins and protocadherins) and multi-spanning (7TM-cadherins) transmembrane proteins[35]. In some instances surprisingly small changes in the protein

sequence within or in proximity to the first transmembrane domain have been shown to cause a shift to GPI anchor-based modes of membrane attachment[36-38].

The function of the extracellular PL domain within the ZIP family branch containing ZIPs 5, 6 and 10 has yet to be characterized. However, multiple more distantly related paralogs (ZIPs 4, 8, 12 and 14) contain a homologous domain and for ZIP4, a protein which has been genetically linked to a rare recessive zinc deficiency disorder (acrodermatitis enteropathica), a role of N-terminal sequences in the sensing of metal ions has recently been suggested[39]. Interestingly, upon depletion of extracellular zinc, ZIP4 has been shown to undergo endoproteolytic cleavage near its 'PALLY' motif, which in turn appears to lower the threshold for cellular zinc uptake[39].

This work holds promise for efforts to elucidate the physiological function of members of the prion protein family. It is hoped that a mechanistic understanding of the workings of ZIP transporters as well as high-resolution structures of their extracellular domains provide insights into the origins and constraints underlying conformational changes associated with prion diseases. Additional work will also be needed to reveal whether membrane-inserted ZIP transporters or N-terminal fragments thereof shed into the extracellular space play a direct role in the manifestation or propagation of prion disease.

METHODS

Molecular clones

FLAG affinity tags were inserted before residue 29 of mouse PrP, residue 27 of mouse Dpl, and residue 26 of mouse Sho (all in the pcDNA3 mammalian expression vector) using standard PCR-based mutagenesis techniques. The identity of all constructs was verified by DNA sequencing.

Cell culture

Mouse neuroblastoma cells (N2a) were cultured in Dulbecco's Modified Eagle's Medium (DMEM) medium containing 10% fetal bovine serum (FBS) and 0.2 × penicillin/streptomycin and maintained in a humidified environment in the presence of 5% CO_2. N2a cells were either transiently or stably transfected with the FLAG-prion constructs or a native pcDNA3 vector (negative control) in OptiMEM using Lipofectamine 2000 (Invitrogen Canada, Burlington, ON, Canada) according to the manufacturer's protocol. For bulk selection of stably transfected cells, cultures were expanded in the presence of 1 mg/mL G418 and maintained at a concentration of 0.2 mg/mL G418.

In vivo crosslinking

Mild formaldehyde crosslinking of N2a cells followed a protocol described before [16]. Briefly, cells grown to confluency were washed with phosphate buffered saline (PBS) and subjected to 15 min crosslinking with 2% w/v formaldehyde in PBS at room temperature. The crosslinking reaction was quenched by incubating cells for 15 min with 125 mM glycine in PBS.

Affinity purification of bait proteins

Approximately 10^8 *in vivo* formaldehyde crosslinked cells each of control and FLAG-prion expressing N2a cell lines were lysed in homogenization buffer (50 mM NH_4Cl, 80 mM Tris, pH 8.0) supplemented with 1× Complete Protease Inhibitor Cocktail (Roche, Palo Alto, CA, USA). To ensure near quantitative extraction of membrane proteins, an equal volume of extraction buffer (20 mM NaCl, 1% deoxycholate, 1% NP-40, 20 mM Tris, pH 8.0) was added, followed by a 30-min incubation and 5-min sonication in a water bath sonicator. Insoluble cellular debris was removed by high-speed centrifugation (100,000 × *g*, 1 h). Subsequently, the crosslinked bait protein complexes were immunoaffinity-captured on anti-FLAG-agarose (Sigma-Aldrich, Oakville, ON, Canada). During this step samples were gently agitated on a turning wheel for 12 h, then washed extensively with 0.5 M NaCl, 0.05% SDS, 1% NP-40, 20 mM HEPES, pH 7.3, and detergents removed by a pre-elution wash with 10 mM NH_4HCO_3, pH 8.0. Proteins were eluted by acidification with 0.2% trifluoroacetic acid, 20% acetonitrile, pH 2.0.

Protein reduction, alkylation and trypsinization

Protein containing fractions were denatured in the presence of 6 M urea, 20 mM NH_4HCO_3, pH 8.0, followed by reduction with 1 mM Tris-(2-carboxyethyl)-phosphine for 30 min at 60°C and alkylation with 2.5 mM 4-vinylpyridine for 1 h at room temperature in the dark. Samples were diluted four-fold to ensure that the concentration of urea did not exceed 1.5 M. Tryptic digestion was initiated by the addition of 1% (wt/wt) of side chain-modified, TPCK-treated porcine trypsin and allowed to proceed at 37°C for 6 h.

iTRAQ labeling

Following trypsinization, equal quantities of tryptic peptide mixtures were spiked with 1 pmol of synthetic (Glu1)-Fibrinopeptide B (GluFib) (Sigma-Aldrich) to serve in the downstream analysis as an internal control for the efficiency of individual labeling reactions. Equal labeling with all four reagents was confirmed by equal intensities of 114:115:116:117 signature peaks upon forced fragmentation of the GluFib $[M+2H]^{2+}$ parent ion at m/z 785.85 . Any strong deviation from this ratio would have indicated problems with the labeling reaction or recovery of individual samples prior to the sample mixing step. Individual iTRAQ labeling reagents (Applied Biosystems, Foster City, CA, USA) were reconstituted in ethanol, added to peptide mixtures derived from the tryptic digestion of IP eluates (Control: iTRAQ 114, Dpl: iTRAQ 115, PrP: iTRAQ 116 , Sho: iTRAQ 117) and incubated at room temperature in the dark for 3 h.

Two-dimensional liquid chromatography

Strong cation exchange (SCX) chromatography was used to achieve peptide fractionation of the complex digest mixture. Samples digested with trypsin were adjusted to 25% acetonitrile and acidified (pH 3.0) by 20-fold dilution in 25% acetonitrile, 20 mM KH_2PO_4, pH 3.0. HPLC was carried out using the Ultimate System (Dionex, Sunnyvale, CA, USA) equipped with a microflow calibration cartridge, a Valco injection port and a 180 nL volume UV cell. Separation was achieved on a self-packed 0.5 mm × 110 mm Luna SCX (Phenomenex, Torrance, CA, USA) column at a flow rate of 18 µl/min with a steep salt gradient from 0–400 mM NH_4Cl in 25% acetonitrile, 20 mM KH_2PO_4, pH 3.0. Fractions eluted from the SCX column were desalted with C18 Empore (3M, Minneapolis, MN, USA) stop and go extraction (STAGE) tips[40] and subsequently

subjected to nano-flow RP-HPLC using the Ultimate LC system (Dionex) equipped with a nanoflow calibration cartridge at a flow rate of 250 nL/min. Peptides were separated on a 75-µm ID self-packed column containing Proteo C12 reverse-phase matrix (Phenomenex) using a 100-min gradient from 2%–34% acetonitrile in water, with 0.1% (wt/vol) formic acid as the ion-pairing agent.

ESI-QqTOF mass spectrometry analysis

The column effluent was coupled directly via a fused silica capillary transfer line to a QSTAR XL hybrid quadrupole/time-of-flight tandem mass spectrometer (Applied Biosystems; MDS Sciex, Concord, ON, Canada) equipped with a MicroIonSpray source. The progress of each LC/MS run was monitored by recording the total ion current (TIC) as a function of time for ions in the m/z range 300 to 1800. At 5-s intervals through the gradient, a mass spectrum was acquired for 1 s, followed by one collision-induced dissociation (CID) acquisition of 4 s each on ions selected by preset parameters of the information-dependent acquisition (IDA) method, using nitrogen as the collision gas. Singly-charged ions were excluded from CID selection. The collision energy was adjusted automatically for each CID spectrum using an empirically optimized formula which considers the charge state and m/z value of the precursor ion.

Database searches

Peak lists for database searching were created using Mascot Distiller (Version 1; MatrixScience, London, UK). Searches were performed using designated MS/MS data interpretation algorithms within Protein Prospector (Version 4.21.3; University of California, San Francisco, CA, USA)[41] and Mascot (Version 2.2; MatrixScience).

Modifications considered were oxidation of methionine, phosphorylations of serine and threonine, N-terminal (pyro)Glu and alkylation with 4-vinylpyridine. Searches further considered up to one missed cleavage and charge states ranging from +2 to +4. For a protein to be listed in the data tables it had to be identified by both search algorithms. In the few instances where only two peptides supported the identification of a protein, we required the underlying CID spectra to generate a Mascot score indicating a <5% probability that the match could be considered a random event[42] and further confirmed matches by the peptide sequence tag approach[43] and manual interpretation of spectra. As searches were carried out without species restriction, the correct assignment of matches to mouse entries served as an additional internal control. The mass tolerance range between expected and observed masses used for database searches was ±150 ppm for MS peaks, and ±0.15 Da for MS/MS fragment ions. These relatively large thresholds were used to capture more of the low intense peaks that frequently display broader distribution and thus are assigned with lower mass accuracy. Threshold levels were optimized based on LC/MS/MS datasets of tryptic digests of standard proteins. All samples were searched against the National Center for Biotechnology Information nonredundant database (nrNCBI, release: June 2008) and a 'decoy' database in which all entries of the above NCBI database had been inverted. The analysis of iTRAQ data was assisted by the software program ProteinPilot (Applied Biosystems; MDS Sciex). A feature of this software package was used to correct raw iTRAQ ratios for impurity levels of individual reagent lots determined by the manufacturer.

Western blotting

Cells were lysed in 0.5% sodium deoxycholate, 1% NP-40, 150 mM NaCl, 50 mM Tris/HCl, pH 8.0 and extracts clarified by centrifugation at 20,000 × *g* for 10 min at 4°C. Protein concentrations were determined using the bicinchoninic acid assay (BCA) (Pierce Biotechnology, Rockford, IL, USA). Proteins were separated on 4-12% NuPAGE gels (Invitrogen) and subsequently transferred to polyvinylidene fluoride (PVDF) membranes. Membranes were blocked with 5% milk in Tris-buffered saline containing 0.05% Tween-20 (TBST). Membranes were incubated overnight at 4°C with primary antibody, washed three times with TBST, and incubated with horseradish peroxidase-conjugated secondary antibody (Bio-Rad Laboratories, Hercules, CA, USA) for 2 h at room temperature. Following three washes with TBST, membranes were developed using Western Lightning ECL (PerkinElmer, Woodbridge, ON, Canada).

Multiple sequence alignments

Multiple sequence alignments were obtained using a combination of MAFFT[44] and ClustalX2's implementation of ClustalW2[45]. Small manual adjustments were made to further improve alignments. Sequence identity and conservation were calculated using the AlignX feature of Vector NTI Advance 10.3.1 (Invitrogen, Carlsbad, CA, USA)[46]. To determine the statistical significance of alignments, the COMPASS program was used[22]. Analyses were based on the 'SCOP40 iter5' database, GAP opening and extension penalties of 10 and 1, respectively, and the BLOSUM62 substitution matrix.

Accession numbers

The following sequences were retrieved from GenBank: *M. musculus* PrP (NP_035300.1), *M. musculus* Sho (CAF18554.1), *M. musculus* Dpl (AAF02544.1), *M. musculus* ZIP5 (AAH28990.1), *G. aculeatus* PrP-1 (CAL64057.1), *T. rubripes* PrP (AAN38988.1), *T. rubripes* Sho1 (CAG34291.1), *T. rubripes* Sho2 (CAG34292.1), *D. rerio* PrP (NP_991149.1), *D. rerio* ZIP10 (NP_956965.1) and *D. rerio* ZIP5 (XP_690258.2). Sequences for *M. musculus* ZIP10 (ENSMUSP00000027131), *G. gallus* PrP (ENSGALP00000040285), *G. gallus* ZIP10 (ENSGALP00000012604), *X. tropicalis* PrP (ENSXETP00000036722), *X. tropicalis* ZIP10 (ENSXETP00000040857), *G. aculeatus* ZIP10 (ENSGACG00000002320) and *T. rubripes* ZIP10 (ENSTRUP00000007404) were retrieved from Ensembl.

Structural threading

The sequence of the PL domain of ZIP10 was submitted to the FFAS03 fold and function assignment server[23] for structure prediction and structural threading. In order to eliminate threading artifacts based on uncertainties in determining the boundaries of the PL domain, we submitted a number of alternatively truncated sequences to the FFAS03 server. This approach helped to better define the boundaries of the folded region of the PL domain of ZIP10. The FFAS03 server provided a ranked list of structural threading models based on various Protein Data Bank (PDB) entries; invariably the highest scoring models were based on PrP[47] and Dpl[48] structures. We analyzed all sensible models for the position of secondary structure elements and derived a consensus prediction for the secondary structure of the PL domain of ZIP10. The closest model to the consensus prediction was chosen for the corresponding figure.

REFERENCES

1. Prusiner, S. B. Prions. Proc Natl Acad Sci U S A 117, 421-434 (1998).
2. Caughey, B. & Baron, G. S. Prions and their partners in crime. Nature 443, 803-810 (2006).
3. Aguzzi, A., Baumann, F. & Bremer, J. The Prion's Elusive Reason for Being. Annu Rev Neurosci 31, 439-477 (2008).
4. Rivera-Milla, E., Stuermer, C. A. O. & Malaga-Trillo, E. An evolutionary basis for scrapie disease: identification of a fish prion mRNA. Trends Genetics 19, 72-75 (2003).
5. Cotto, E., Andre, M., Forgue, J., Fleury, H. J. & Babin, P. J. Molecular characterization, phylogenetic relationships, and developmental expression patterns of prion genes in zebrafish. FEBS Journal 272, 500-513 (2004).
6. Rivera-Milla, E. et al. Disparate evolution of prion protein domains and the distinct origin of Doppel- and prion-related loci revealed by fish-to-mammal comparisons. Faseb J 20, 317-319 (2005).
7. Watts, J. C. & Westaway, D. The prion protein family: diversity, rivalry, and dysfunction. Biochim Biophys Acta 1772, 654-672 (2007).
8. Watts, J. C. et al. The CNS glycoprotein Shadoo has PrP(C)-like protective properties and displays reduced levels in prion infections. EMBO J 26, 4038-4050 (2007).
9. Wuthrich, K. & Riek, R. Three-dimensional structures of prion proteins. Adv Protein Chem 57, 55-82 (2001).
10. Date, S. V. Estimating protein function using protein-protein relationships. Methods Mol Biol 408, 109-127 (2007).
11. Schmitt-Ulms, G. et al. Time-controlled transcardiac perfusion crosslinking for the study of protein interactions in complex tissue. Nat Biotechnol 22, 724-731 (2004).
12. Gingras, A. C., Gstaiger, M., Raught, B. & Aebersold, R. Analysis of protein complexes using mass spectrometry. Nat Rev Mol Cell Biol 8, 645-654 (2007).
13. Race, R. E., Fadness, L. H. & Chesebro, B. Characterization of scrapie infection in mouse neuroblastoma cells. J Gen Virol 68, 1391-1399 (1987).
14. Einhauer, A. & Jungbauer, A. The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins. J Biochem Biophys Methods 49, 455-465 (2001).
15. Telling, G. C. et al. N-terminally tagged prion protein supports prion propagation in transgenic mice. Protein Science 6, 825-833 (1997).
16. Schmitt-Ulms, G. et al. Binding of neural cell adhesion molecules (N-CAMs) to the cellular prion protein. J Mol Biol 314, 1209-1225 (2001).
17. Zieske, L. R. A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. J Exp Bot 57, 1501-1508 (2006).
18. Taylor, K. M. et al. The emerging role of the LIV-1 subfamily of zinc transporters in breast cancer. Mol Med 13, 396-406 (2007).
19. Lo Conte, L. et al. SCOP: a Structural Classification of Proteins database. Nucleic Acids Research 28, 257-259 (2000).
20. Gough, J., Karplus, K., Hughey, R. & Chothio, C. Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. J Mol Biol 313, 903-919 (2001).

21. Sprenger, J. et al. LOCATE: a mammalian protein subcellular localization database. Nucleic Acids Res 36, D230-233 (2008).
22. Sadreyev, R. I. & Grishin, N. V. Accurate statistical model of comparison between multiple sequence alignments. Nucleic Acids Res 36, 2240-2248 (2008).
23. Jaroszewski, L., Rychlewski, L., Li, Z., Li, W. & Godzik, A. FFAS03: a server for profile-profile sequence alignments. Nucleic Acids Res 33, W284-288 (2005).
24. Holm, L. & Park, J. DaliLite workbench for protein structure comparison. Bioinformatics 16, 566-567 (2000).
25. Kraut, J. et al. The aromatic substrate binding site in subtilisin BPN' and its resemblance to chymotrypsin. Cold Spring Harb Symp Quant Biol 36, 117-123 (1972).
26. Kornegay, J. R. Molecular genetics and evolution of stomach and nonstomach lysozymes in the hoatzin. J Mol Evol 42, 676-684 (1996).
27. Davies, P. & Brown, D. R. The chemistry of copper binding to PrP: is there sufficient evidence to elucidate a role for copper in protein function? Biochem J 410, 237-244 (2008).
28. Watt, N. T. & Hooper, N. M. The prion protein and neuronal zinc homeostasis. Trends Biochem Sci 28, 406-410 (2003).
29. Rost, B. Twilight zone of protein sequence alignments. Prot Engineering 12, 85-94 (1999).
30. Szyrwiel, L. et al. Zn(II) ions bind very efficiently to tandem repeat region of "prion related protein" (PrP-rel-2) of zebra-fish. MS and potentiometric evidence. Dalton Trans 44, 6117-6120 (2008).
31. Lee, I. Y. et al. Complete genomic sequence analysis of the prion protein gene region from three mammalian species. Genome Res 8, 1032-1037 (1998).
32. Premzl, M., Gready, J. E., Jermin, L. S., Simonic, T. & Marshall Graves, J. A. Evolution of vertebrate genes related to prion and shadoo proteins - clues from comparative genomic analysis. Mol Biol Evol 21, 2210-2231 (2004).
33. Baldwin, M. A. Analysis of glycosylphosphatidylinositol protein anchors: the prion protein. Methods Enzymol 405, 172-187 (2005).
34. Rojas, M., DeMarte, L., Screaton, R. A. & Stanners, C. P. Radical differences in functions of closely related members of the human carcinoembryonic antigen gene family. Cell Growth Differ 7, 655-662 (1996).
35. Angst, B. D., Marcozzi, C. & Magee, A. I. The cadherin superfamily: diversity in form and function. J Cell Sci 114, 629-641 (2001).
36. Bell, L. M., Solomon, K. R., Gold, J. P. & Tan, K.-N. Cytoplasmic tail deletion of T cell receptor (TCR) b-chain results in its surface expression as glycosylphosphatidylinositol-anchored polypeptide on mature T cells in the absence of TCR-a. J Biol Chem 269, 22758-22763 (1994).
37. Lanier, L., Cwirla, S., Yu, G., Testi, R. & Phillips, J. H. Membrane anchoring of a human IgG Fc receptor (CD16) determined by a single amino acid. Science 246, 1611-1613 (1989).
38. Naghibalhossaini, F. & Stanners, C. P. Minimal mutations are required to effect a radical change in function in CEA family members of the Ig superfamily. J Cell Sci 117, 761-769 (2004).

39. Kambe, T. & Andrews, G. K. Novel proteolytic processing of the ectodomain of the zinc transporter ZIP4 (Slc39a4) during zinc deficiency is inhibited by acrodermatitis enteropathica mutations. Mol Cell Biol PMID 18936158 (2008).
40. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem 75, 663-670 (2003).
41. Chalkley, R. J. et al. Comprehensive Analysis of a Multidimensional Liquid Chromatography Mass Spetrometry Dataset Acquired on a Quadrupole Selecting Quadrupole Collision Cell, Time-of-flight Mass Spectrometer. II. New Developments in Protein Prospector Allow for Reliable and Comprehensive Automatic Analysis of Large Datasets. Mol Cell Prot 4, 1194-1204 (2005).
42. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20, 3551-3567 (1999).
43. Mann, M. & Wilm, M. Error-tolerant identification of peptides in sequence datases by peptides sequence tags. Anal Chem 66, 4390-4399 (1994).
44. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30, 3059-3066 (2002).
45. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947-2948 (2007).
46. Lu, G. & Moriyama, E. N. Vector NTI, a balanced all-in-one sequence analysis suite. Brief Bioinform 5, 378-388 (2004).
47. Riek, R. et al. NMR structure of the mouse prion protein domain PrP(121-321). Nature 382, 180-182 (1996).
48. Mo, H. et al. Two different neurodegenerative diseases caused by proteins with similar structures. Proc Natl Acad Sci U S A 98, 2352-2357 (2001).

ACKNOWLEDGMENTS

We thank Boris Steipe for suggestions, Janaky Coomaraswamy for the FLAG-tagged Doppel cDNA, and Nathalie Daude and Agnes Lau for assistance with bioinformatic analyses. Work on this project was funded through support from the Canadian Institutes of Health Research (MOP-74734-GSU and MOP36377-DW), a fellowship to J.C.W. from the Natural Sciences and Engineering Research Council of Canada (PGSD2-319161-2005), the Alberta Heritage Foundation for Medical Research (Grant 200600752) and the National Institutes of Health (NIH AG02132 and AG10770). G.S.-U. received support from the W. Garfield Weston Foundation.

AUTHOR CONTRIBUTIONS

G.S.-U. designed the study, carried out the interactome analyses, identified the evolutionary link between ZIPs and prion genes and wrote the paper. S.E. refined the bioinformatic analyses and assembled the sequence alignments. J.C.W. generated stable cell clones expressing FLAG-tagged expression constructs. D.W. provided prion expression vectors and logistical support. H.W. generated the structural threading models. All authors contributed to the bioinformatic analyses, discussed the data and edited the manuscript.

FIGURE LEGENDS

Figure 1: Structural similarity between mouse ZIP10, PrP and Dpl. a, Structural threading of the PL domain within ZIP10 predicts striking resemblance to PrP and Dpl with regard to relative position and order of secondary structure motifs. The secondary structure for the PL domain of ZIP10 was ranked according to the frequency of the prediction in separate threadings (**H** 75-100%, H 50-75%, **h** 25-50%, h 0-25% for α-helices; the same scale applies to β-sheeted structures "**E**/E/**e**/e"). **b**, Comparison of high-resolution nuclear magnetic resonance structures of PrP and Dpl with a predicted structure for the ZIP10 PL domain. Dark green, grey and black highlights depict conserved, similar and identical residues, respectively. 'S' and 'G' labels indicate sites of disulfide linkages and glycosylation, respectively.

Figure 2: Sequence evidence for common origin and divergent sequence evolution of members of ZIP and prion protein families. a, Multiple sequence alignment of PrP globular domain with ZIP10 PL domain from pufferfish and mouse. **b**, Multiple sequence alignment of N-terminal Sho and ZIP5 sequences from pufferfish, zebrafish and mouse. The addition of murine sequences served to document the greater divergence of orthologous proteins in mouse. Highlight colors are as in Figure 1. 'N', 'S', 'G' and 'C' labels indicate N-termini, disulfide linkages, glycosylation sites and C-termini, respectively. For full-length multiple alignments of these and related sequences, please see **Supplementary Fig. 2**.

Figure 3: Models depicting evolutionary origin and topology of members of prion protein family, hypothetical ZIP ancestor and ZIP10 transporter. a, Cartoon depicting the emergence of members of the prion gene family from ZIP ancestor gene(s): (i) Hypothetical Sho ancestor gene, (ii) prion gene founder and (iii) evolutionary intermediate prion gene family ancestor. Alternative hypotheses for the origin of prion genes: (1) insertion of ZIP ancestor-derived PL domain into Sho ancestor molecule; or (2) duplication of N-terminal ZIP ancestor fragment followed by expansion of hydrophobic domain and differentiation of former transmembrane domain into signal peptide for attachment of GPI anchor. Based on the above models, Sho genes (1) either evolved independently or (2) are themselves derived from a ZIP ancestor. Please note that our analyses strongly favor model (2). **b**, Comparison of orientation and membrane topology of prion proteins, hypothetical ZIP ancestor protein and today's ZIP10.

Supplementary Figure 1: Evidence for specific co-enrichment of ZIP10 and ZIP6 with all three members of the mammalian prion protein family. Comparison of FLAG-affinity chromatography eluates by quantitative tandem mass spectrometry. Side-by-side affinity purified bait protein complexes were trypsinized and subjected to iTRAQ labeling of peptides as follows: iTRAQ114 label: empty vector; iTRAQ115 label: FLAG-Dpl; iTRAQ116 label: FLAG-PrP; and iTRAQ117 label: FLAG-Sho. **a**, ZIP10 co-purified specifically with the three bait proteins. Collision induced dissociation (CID) spectrum from ZIP10 derived peptide with amino acid sequence QSTEEIGR ($[M+2H]^{2+}$, m/z 626.35). Inset: Low mass iTRAQ reporter ion region documenting relative contribution to the identification of this peptide by samples labeled with iTRAQ115 (FLAG-Dpl), iTRAQ116 (FLAG-PrP) and iTRAQ117 (FLAG-Sho) reagents but not

negative control sample labeled with iTRAQ114 reagent. **b**, CID spectrum derived from ZIP6 peptide with amino acid sequence ESASSSEVTSAVYNAVSEGTR ($[M+3H]^{3+}$, m/z 759.03). iTRAQ reporter ions document that ZIP6 co-purified specifically with the three bait proteins. **c**, Actin co-purified unspecifically in all four samples including the negative control. CID spectrum derived from actin peptide with amino acid sequence TTGIVMDSGDGVTHTVPIQEGYALPHAILR ($[M+4H]^{4+}$, m/z 666.35).

Supplementary Figure 2: Multiple full-length sequence alignment of selected mammalian and teleost ZIP and prion gene sequences. Due to the presence of large insertions found in a subset of depicted sequences, this alignment required extensive manual curation. Amino acid-specific colours were employed to facilitate visual comparison of sequences.

Table 1: Quantitative analysis of mouse Dpl, PrP and Sho interactomes identifies metal ion transporters of the ZIP protein family in spatial proximity to all three members of the mammalian prion protein family

Identified proteins[a]	IPI accession number	Amino acids[c]	Peptides[d]	Unique[e]	% Cov[f]	IPs & iTRAQ[b] 114 Control	115 Dpl	116 PrP	117 Sho
Specific binders									
Dpl	IPI00131622.1	178	8	13	39.1	2.0	92.5	3.0	2.5
PrP	IPI00120793.1	254	5	5	26.8	4.0	20.5	65.1	10.5
Sho	IPI00226455.1	147	3	6	50.4	2.9	11.4	20.5	65.2
ZIP10	IPI00273801.3	833	2	2	4.3	2.4	36.3	16.7	44.6
ZIP6	IPI00469000.4	765	3	3	7.3	3.2	29.9	32.8	34.2
Unspecific binders									
Actin	IPI00110850.1	375	21	46	88.5	19.1	21.6	31.8	27.5

[a]*Proteins were sorted into specific versus unspecific binder categories based on their iTRAQ distribution, i.e. proteins were considered unspecific interactors if their derived CID spectra revealed iTRAQ114 to 117 signature mass peak signal intensities which exceeded 10% of combined intensities for all samples including the unspecific control.*

[b]*For the calculation of iTRAQ values the intensity of individual peptide associated iTRAQ signature peaks was normalized to combine to 100% per peptide and subsequently averaged. Standard deviations were determined and are listed in* ***Supplementary Table 1****.*

[c]*Length of precursor molecules prior to posttranslational processing.*

[d]*Only CID spectra underlying different peptides were considered i.e. if the same peptide was identified with different charge states or modifications it counted as one hit.*

[e]*Total number of unique CID spectra. Please note that the same peptide was only counted more than once if it was identified with different charge states or modifications.*

[f]*Percent sequence coverage based on the presence of peptides for which no higher ranked assignment to other proteins could be made.*

Figure 1

a

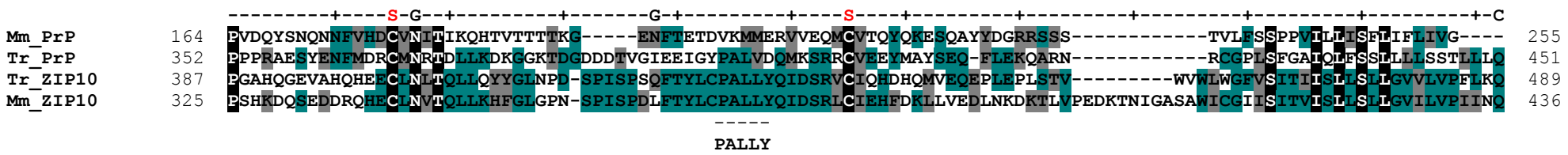
```
                 ---+---------+---------+---------+---------+---------+--S-G----+---------+----G----+---------+-S-------+---------
Mm_PrP       124 LGGYMLGSAMSRPMIH-FGNDWEDRYYRENMYRYPNQVYYRP-VDQYSNQNNFVHDCVNITIKQHTVTTTTKG----ENFTETDVKMMERVVEQMCVTQYQKESQAYYDGRRS 230
Mm_Dpl        56 RPGAFIKQG-RKLDID-FGAE-GNRYYAANYWQFPDGIYYEGCSEANVTKEMLVTSCVNATQAANQAEFSREK-----QDSKLHQRVLWRLIKEICSAKHCDFWLERGAA--- 157
Mm_ZIP10     285 HSGHELGHGHQELDPDNEGELRHTR-KREAPHVRKSAIYSTP-SHKDQSEDDRQHECLNVTQLLKHFGLGPNSPISPDLFTYLCPALLYQIDSRLCIEHFDKLLVEDLNKDKT 395
Mm_PrP 2D          EEEE            HHHHHHHHHHH     EEEE            HHHHHHHHHHHHHHH          HHHHHHHHHHHHHHHHHHH
Mm_Dpl 2D          EEE             HHHHHHHHHHH     EEE         HHHHHHHHHHHHHHHHHHHHHHHHH       HHHHHHHHHHHHHHHHH
Mm_ZIP10           eEEEee      hhhhHHHHHHHHHHHhhhh  eeEEe        HHHHHHHHHHHHHHHHhhhhhhhhhhhhhhh  HHHHHHHHHHHHHHHHHHHHHh
```

b

PrP Dpl ZIP10 PL domain

Figure 2

a



b

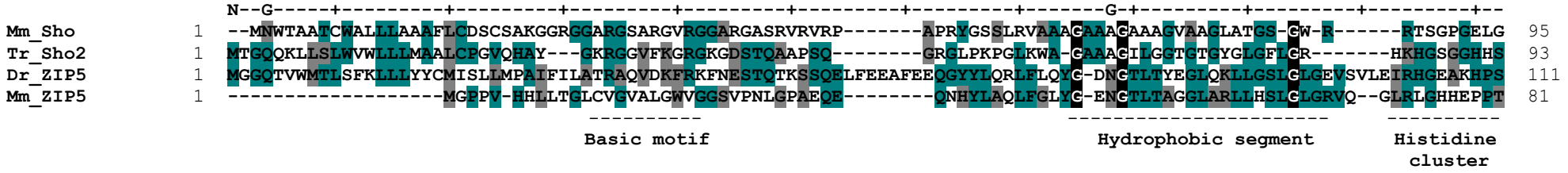
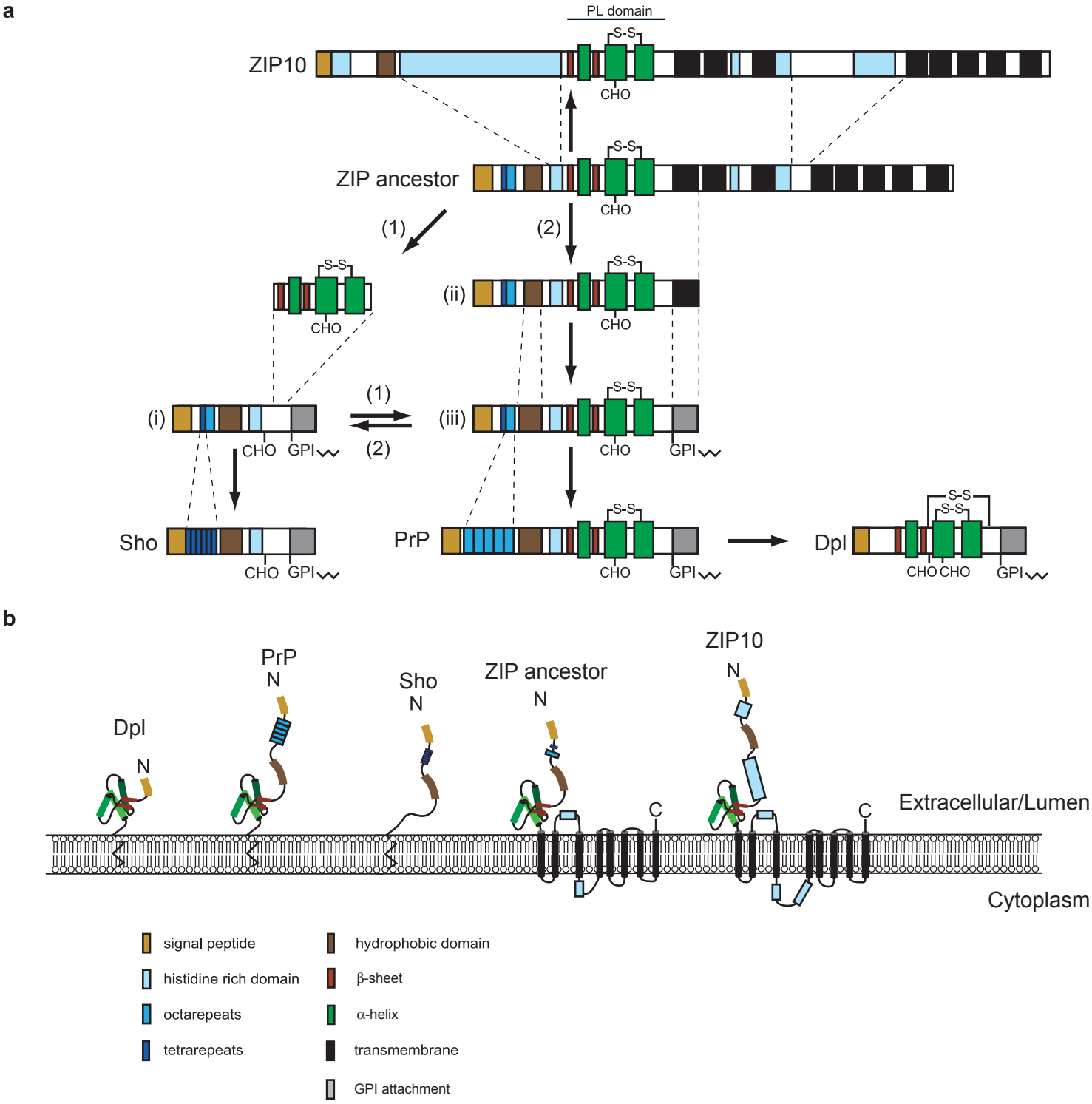
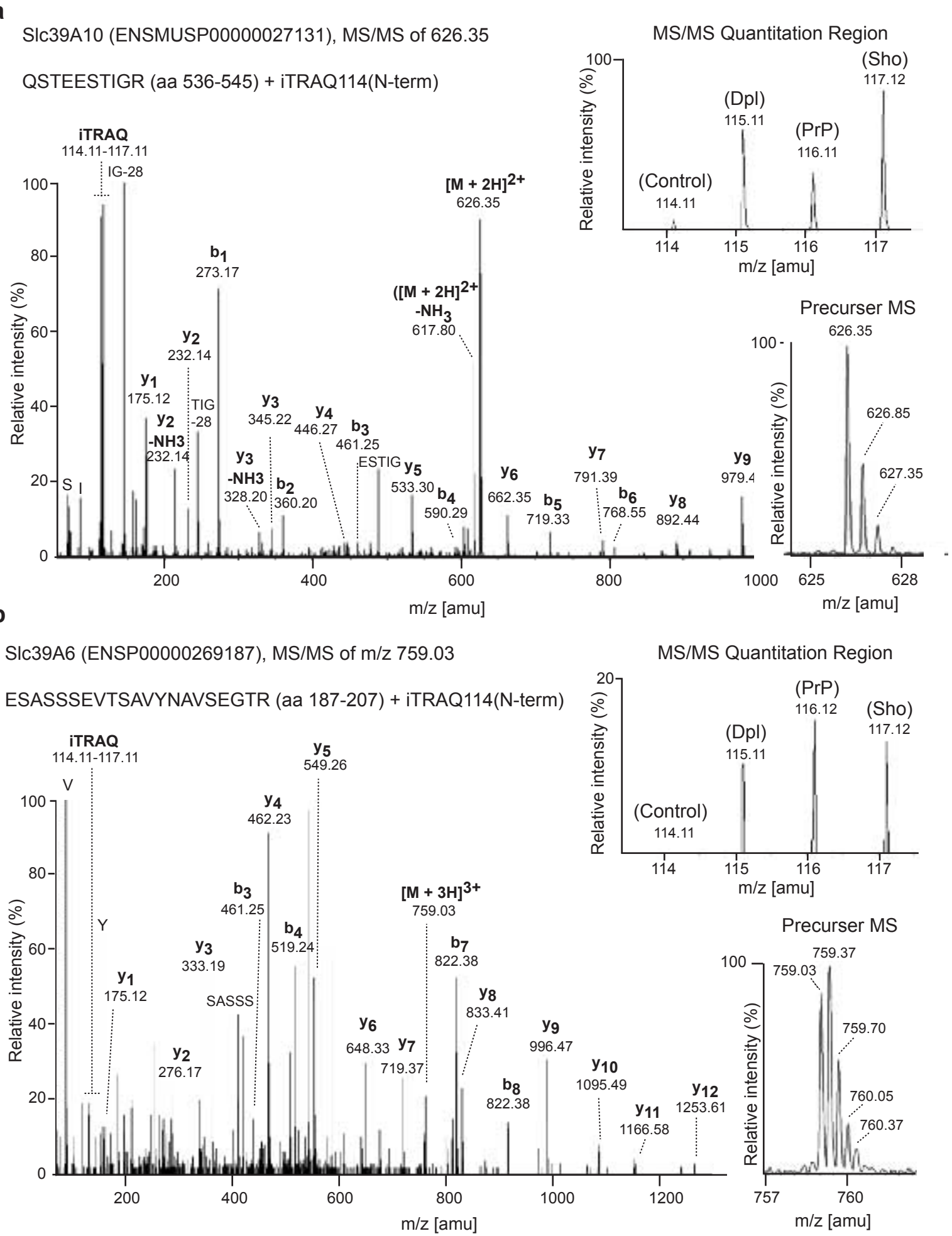


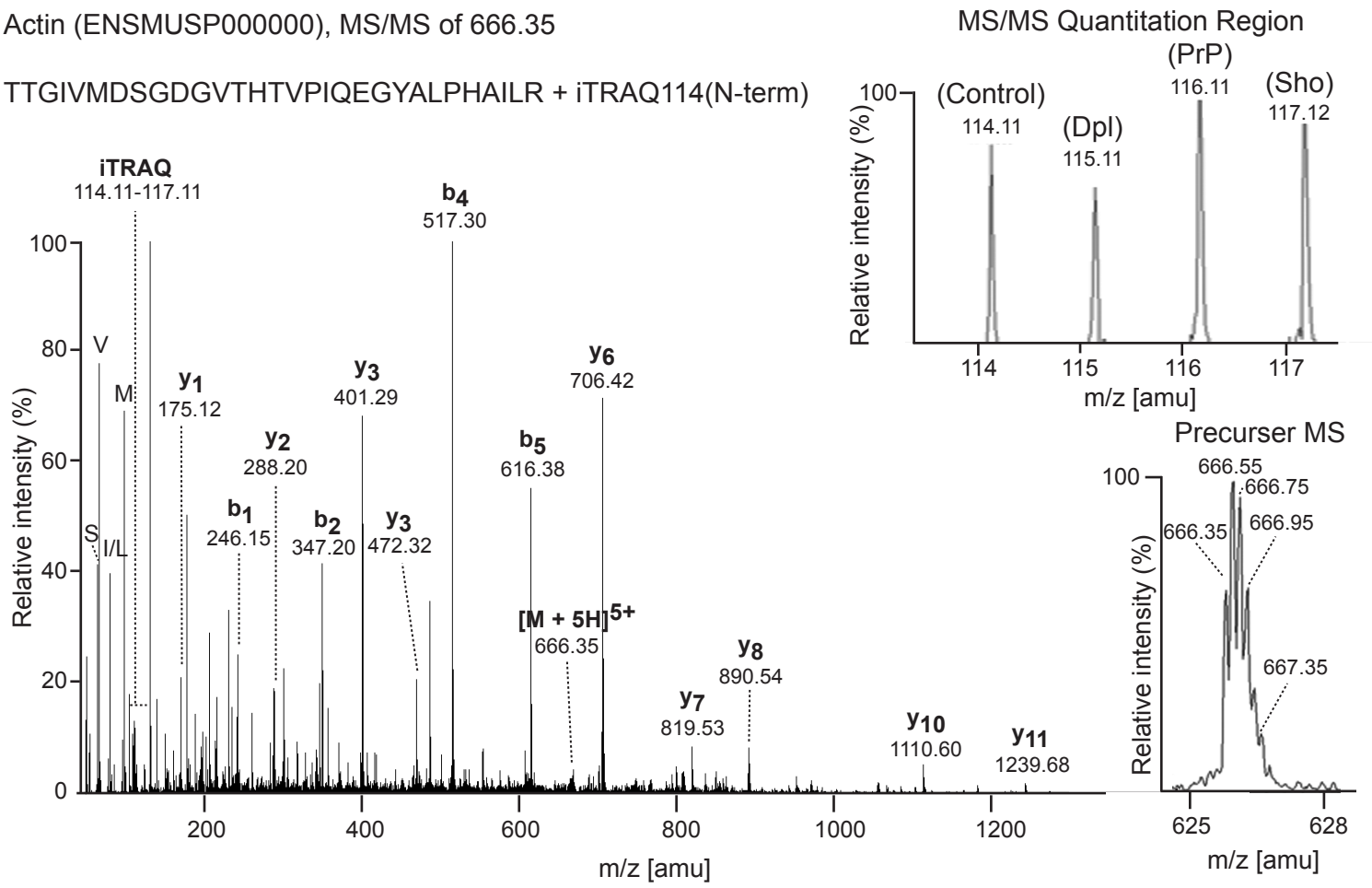
Figure 3



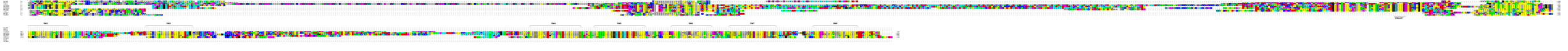
Supplementary Figure 1



c



Supplementary Figure 2



Supplementary Table 1

Accession number	Symbol	Protein	Conf	Sequence	Modifications	Cleavages	dMass	Prec MW	Prec m/z	Theor MW	Theor m/z	Theor z		Area 114	Area 115	Area 116	Area 117
IPI:IPI00131622.1	DPL	doppel	99	HCDFWLER	iTRAQ4plex@N-term; Pyridylethyl(C)@2		-0.03	1353.61	452.21	1353.64	452.22	3		6.0	89.0	3.4	1.6
			99	VAENRPGAF	iTRAQ4plex@N-term	cleaved F-I@C-term	0.00	1103.58	552.80	1103.58	552.80	2		0.7	94.7	1.7	2.9
			99	VLPSSGGQITEAR	iTRAQ4plex@N-term		-0.01	1457.79	729.90	1457.80	729.91	2		0.1	98.1	1.2	0.6
			99	LDIDFGAEGNR	iTRAQ4plex@N-term		-0.02	1349.65	675.83	1349.67	675.84	2		1.3	94.3	2.6	1.7
			99	VLPSSGGQITEAR	iTRAQ4plex@N-term; Ala->Val@12		0.01	1485.84	496.29	1485.83	496.28	3		1.6	89.9	2.4	6.0
			99	VLPSSGGQITEARVAENRPGAFIKQGR	iTRAQ4plex@N-term; Asn->His@17; Lys->Glu@24	missed R-V@13; missed K-Q@24	0.00	3005.59	752.41	3005.59	752.41	4		3.9	89.9	4.1	2.1
			99	VLPSSGGQITEAR	iTRAQ4plex@N-term; Deamidated(Q)@8; Dimethyl(R)@13		0.06	1486.87	496.63	1486.81	496.61	3		1.0	91.2	4.4	3.3
			99	KVLPSSGGQITEAR	iTRAQ4plex@N-term; iTRAQ4plex(K)@1	missed K-V@1	-0.02	1729.97	577.66	1729.99	577.67	3		1.7	91.0	4.1	3.2
			99	KLDIDFGAEGNR	iTRAQ4plex@N-term; iTRAQ4plex(K)@1	missed K-L@1	-0.04	1621.83	541.62	1621.87	541.63	3		2.4	87.6	6.3	3.7
			99	KVLPSSGGQITEAR	iTRAQ4plex@N-term; iTRAQ4plex(K)@1; HexNAc(T)@11	missed K-V@1	0.02	1933.09	645.37	1933.07	645.36	3		2.5	92.2	3.3	2.0
			99	VAENRPGAFIK	iTRAQ4plex@N-term; iTRAQ4plex(K)@11		0.96	1489.83	497.62	1488.87	497.30	3		0.0	99.4	0.0	0.6
			99	VAENRPGAFIK	iTRAQ4plex@N-term; Lys->Arg@11		-0.05	1372.72	687.37	1372.77	687.39	2		1.0	95.9	1.5	1.6
			92	KVLPSSGGQITEAR	iTRAQ4plex@N-term; Lys->Gln@1	missed K-V@1	0.05	1585.90	529.64	1585.85	529.63	3		3.5	89.0	3.8	3.7
													Average:	2.0	92.5	3.0	2.5
													StDev:	1.7	3.7	1.6	1.5
IPI:IPI00120793.1	PRNP	prion protein	99	HVAGAAAAGAVVGGLGGYMLGSAMSR	iTRAQ4plex@N-term	cleaved R-P@C-term	0.03	2474.29	825.77	2474.26	825.76	3		7.0	25.9	55.7	11.5
			96	QHTVTTTTK	Gln->pyro-Glu@N-term; iTRAQ4plex(K)@9		0.02	1142.62	572.32	1142.61	572.31	2		2.2	24.0	61.0	12.8
			99	PMIHFGNDWEDR	iTRAQ4plex@N-term	cleaved R-P@N-term	0.06	1659.82	554.28	1659.76	554.26	3		6.1	26.8	55.3	11.8
			97	ESQAYYDGR	iTRAQ4plex@N-term		0.00	1231.56	616.79	1231.56	616.79	2		3.4	11.5	76.4	8.6
			99	VVEQMCVTQYQK	iTRAQ4plex@N-term; Pyridylethyl(C)@6; iTRAQ4plex(K)@12		-0.09	1847.86	616.96	1847.95	616.99	3		1.1	14.2	76.9	7.7
													Average:	4.0	20.5	65.1	10.5
													StDev:	2.5	7.1	10.8	2.2
IPI:IPI00226455.1	SPRN	shadoo	99	ICLLLGGTLGALELLRP	iTRAQ4plex@N-term; Pyridylethyl(C)@2		-0.01	2000.19	667.74	2000.20	667.74	3		0.0	6.9	32.4	60.7
			99	VAAAGAAAGAAAGVAAGLATGSGWR	iTRAQ4plex@N-term		-0.04	2198.13	733.72	2198.17	733.73	3		0.0	21.6	18.3	60.1
			98	YGSSLR	iTRAQ4plex@N-term		0.00	825.45	413.73	825.45	413.73	2		0.8	2.3	23.2	73.8
			99	VAAAGAAAGAAAGVAAGLATGSGWR	iTRAQ4plex@N-term; Oxidation(W)@24		-0.01	2214.15	739.06	2214.16	739.06	3		8.0	13.4	15.0	63.5
			99	VAAAGAAAGAAAGVAAGLATGSGWR	iTRAQ4plex@N-term; Val->Asn@14		0.05	2213.19	738.74	2213.14	738.72	3		7.1	12.9	15.4	64.7
			99	VAAAGAAAGAAAGVAAGLATGSGWR	iTRAQ4plex@N-term; Val->Met@14		0.07	2230.21	744.41	2230.14	744.39	3		1.4	11.1	18.8	68.7
													Average:	2.9	11.4	20.5	65.2
													StDev:	3.7	6.5	6.5	5.2
IPI:IPI00469000.4	SLC39A6	ZIP6	99	ESASSSEVTSAVYNAVSEGTR	iTRAQ4plex@N-term		0.01	2274.08	759.03	2274.07	759.03	3		3.7	29.2	35.9	31.2
			95	AGMTVK	iTRAQ4plex@N-term; Dehydrated(T)@4; iTRAQ4plex(K)@6		-0.03	875.48	438.75	875.51	438.76	2		2.2	31.1	31.9	34.8
			99	YDSQLSSNEEK	iTRAQ4plex@N-term; Lys->Gln@11		0.06	1442.69	722.35	1442.63	722.32	2		3.7	29.2	30.5	36.6
													Average:	3.2	29.9	32.8	34.2
													StDev:	0.8	1.1	2.8	2.8
IPI:IPI00273801.3	SLC39A10	ZIP10	99	QSTEESTIGR	iTRAQ4plex@N-term		0.03	1250.65	626.33	1250.62	626.32	2		2.3	39.3	15.7	42.7
			98	NYLGVEEEK	iTRAQ4plex@N-term; Lys->Gln@9		0.07	1223.65	612.83	1223.58	612.80	2		2.5	33.2	17.7	46.5
													Average:	2.4	36.3	16.7	44.6
													StDev:	0.2	4.3	1.4	2.7
IPI:IPI00110850.1	ACTB	actin	99	AGFAGDDAPR	iTRAQ4plex@N-term		0.02	1119.56	560.79	1119.54	560.78	2		27.5	13.1	37.0	22.4
			99	AVFPSIVGR	iTRAQ4plex@N-term	cleaved R-P@C-term	-0.02	1088.63	545.32	1088.65	545.33	2		27.2	13.0	27.9	31.9
			99	AVFPSIVGRPR	iTRAQ4plex@N-term; Deamidated(R)@9		-0.03	1342.76	448.59	1342.78	448.60	3		25.5	14.2	33.8	26.5
			84	AVFPSIVGRPR	iTRAQ4plex@N-term; Dimethyl(R)@11		0.01	1369.85	457.62	1369.83	457.62	3		28.8	16.8	32.9	21.5
			99	DLTDYLMK	iTRAQ4plex@N-term; iTRAQ4plex(K)@8		0.00	1285.68	643.85	1285.68	643.85	2		18.2	23.6	37.5	20.7
			98	DLTDYLMK	iTRAQ4plex@N-term; Lys->Gln@8		0.02	1141.57	571.79	1141.54	571.78	2		10.6	8.5	30.5	50.4
			99	DLYANTVLSGGTTMYPGIADR	iTRAQ4plex@N-term		0.13	2358.29	787.10	2358.16	787.06	3		11.0	29.3	30.3	29.4
			99	DLYANTVLSGGTTMYPGIADR	iTRAQ4plex@N-term; Dimethyl(R)@21		0.03	2386.23	796.42	2386.20	796.41	3		30.1	18.6	0.0	51.3
			99	DLYANTVLSGGTTMYPGIADR	iTRAQ4plex@N-term; Leu->Pro@2		-0.02	2342.12	781.71	2342.13	781.72	3		5.7	28.2	29.4	36.7
			99	DSYVGDEAQSK	iTRAQ4plex@N-term; Dehydrated(S)@2; Lys->Gln@11		0.04	1323.61	662.81	1323.57	662.79	2		29.3	8.9	29.6	32.2
			99	DSYVGDEAQSK	iTRAQ4plex@N-term; iTRAQ4plex(K)@11		0.04	1485.76	496.26	1485.72	496.25	3		22.5	29.7	31.8	16.0
			99	DSYVGDEAQSK	iTRAQ4plex@N-term; Lys->Arg@11		0.04	1369.67	685.84	1369.62	685.82	2		21.7	10.8	31.2	36.4
			99	DSYVGDEAQSK	iTRAQ4plex@N-term; Lys->Gln@11		0.04	1341.62	671.82	1341.58	671.80	2		21.6	9.4	33.0	35.9
			87	DSYVGDEAQSK	iTRAQ4plex@N-term; Ser->Asp@10; iTRAQ4plex(K)@11		0.01	1513.72	505.58	1513.71	505.58	3		26.7	17.7	34.1	21.6
			92	DSYVGDEAQSK	iTRAQ4plex@N-term; Ser->Val@10; iTRAQ4plex(K)@11		0.01	1497.76	500.26	1497.76	500.26	3		27.3	25.1	28.8	18.8
			99	DSYVGDEAQSKR	iTRAQ4plex@N-term; iTRAQ4plex(K)@11	missed K-R@11	0.04	1641.86	548.29	1641.82	548.28	3		17.9	32.4	44.2	5.5
			96	DSYVGDEAQSKR	iTRAQ4plex@N-term; Lys->Gln@11	missed K-R@11	0.08	1497.76	500.26	1497.68	500.23	3		19.9	30.6	33.0	16.4
			80	EITALAPSTMK	Glu->pyro-Glu@N-term; iTRAQ4plex(K)@11		0.06	1286.76	644.39	1286.70	644.36	2		25.0	12.1	33.5	29.5
			99	EITALAPSTMK	iTRAQ4plex@N-term; Lys->Arg@11		0.01	1332.73	667.37	1332.72	667.37	2		19.8	11.4	33.8	35.0
			89	EITALAPSTMK	iTRAQ4plex@N-term; Lys->Gln@11		0.07	1304.75	653.38	1304.68	653.35	2		20.7	19.2	32.7	27.4
			99	EITALAPSTMKI	iTRAQ4plex@N-term; Lys->Ala@11; Ile->Ala@12	cleaved I-K@C-term; missed K-I@11	0.01	1318.70	660.36	1318.69	660.35	2		18.0	8.7	28.2	45.1
			99	GYSFTTTAER	iTRAQ4plex@N-term		0.00	1275.63	638.82	1275.62	638.82	2		13.6	14.7	36.8	34.9
			99	GYSFTTTAER	iTRAQ4plex@N-term; Dimethyl(R)@10		0.05	1303.71	652.86	1303.65	652.83	2		22.5	6.8	32.6	38.0
			99	HQGVMVGMGQK	iTRAQ4plex@N-term; iTRAQ4plex(K)@11		0.00	1458.77	487.26	1458.77	487.26	3		14.5	30.6	37.5	17.4
			90	HQGVMVGMGQK	iTRAQ4plex@N-term; Lys->Arg@11		0.03	1342.70	448.57	1342.67	448.56	3		11.4	13.7	34.2	40.7
			99	HQGVMVGMGQK	iTRAQ4plex@N-term; Lys->Gln@11		0.05	1314.68	439.23	1314.63	439.22	3		14.2	22.0	31.8	32.0
			99	KDLYANTVLSGGTTMYPGIADR	iTRAQ4plex@N-term; iTRAQ4plex(K)@1		0.04	2630.40	658.61	2630.36	658.60	4		12.5	33.2	28.5	25.8
			99	LDLAGRDLTDYLMK	iTRAQ4plex@N-term; iTRAQ4plex(K)@14	missed R-D@6	0.01	1911.05	638.02	1911.04	638.02	3		40.6	15.9	32.7	10.8
			98	MTQIMFETFNTPAMYVAIQAVLSLYASGR	iTRAQ4plex@N-term		0.07	3396.77	850.20	3396.70	850.18	4		8.9	52.0	28.3	10.8
			99	QEYDESGPSIVHR	iTRAQ4plex@N-term		0.04	1659.84	554.29	1659.80	554.27	3		6.6	42.9	29.0	21.5
			99	SYELPDGQVITIGNER	iTRAQ4plex@N-term		-0.02	1933.97	645.66	1933.99	645.67	3		11.2	41.0	30.4	17.5
			99	SYELPDGQVITIGNER	iTRAQ4plex@N-term; Asn->Arg@14		0.01	1976.05	659.69	1976.04	659.69	3		18.3	20.0	36.3	25.4
			99	SYELPDGQVITIGNER	iTRAQ4plex@N-term; Asp->Pro@6		-0.08	1915.93	639.65	1916.01	639.68	3		5.5	16.6	32.5	45.4
			99	SYELPDGQVITIGNER	iTRAQ4plex@N-term; Deamidated(N)@14; Glu->Arg@15		0.02	1962.05	655.02	1962.03	655.02	3		20.0	23.6	28.9	27.5
			99	SYELPDGQVITIGNER	iTRAQ4plex@N-term; Dehydrated(D)@6		-0.05	1915.93	639.65	1915.98	639.67	3		10.9	13.7	33.6	41.9
			81	SYELPDGQVITIGNER	iTRAQ4plex@N-term; Dehydrated(D)@6; Dimethyl(R)@16		0.01	1944.02	649.01	1944.01	649.01	3		13.0	21.2	41.0	24.8
			99	SYELPDGQVITIGNER	iTRAQ4plex@N-term; Dimethyl(R)@16		-0.01	1962.01	655.01	1962.02	655.01	3		18.6	14.0	37.5	29.9
			99	TTGIVMDSGDGVTHTVPIYEGYALPH	iTRAQ4plex@N-term	cleaved H-A@C-term	-0.04	2873.36	958.79	2873.40	958.81	3		7.7	19.7	30.5	42.0
			99	TTGIVMDSGDGVTHTVPIYEGYALPHAILR	iTRAQ4plex@N-term		0.04	3326.75	832.69	3326.71	832.68	4		24.8	20.5	29.3	25.4
			99	TTGIVMDSGDGVTHTVPIYEGYALPHAILR	iTRAQ4plex@N-term; Dimethyl(R)@30		-0.03	3354.72	839.69	3354.74	839.69	4		31.5	14.5	32.5	21.5
			99	TVLSGGTTMYPGIADR	iTRAQ4plex@N-term	cleaved N-T@N-term	0.03	1781.94	594.99	1781.91	594.98	3		21.1	16.9	34.7	27.2
			99	VAPEEHPVLLTEAPLNPK	Carbamyl@N-term; iTRAQ4plex(K)@18		-0.04	2140.13	714.38	2140.17	714.40	3		1.1	67.1	19.4	12.4
			93	VAPEEHPVLLTEAPLNPK	Formyl@N-term; iTRAQ4plex(K)@18		-0.02	2125.14	709.39	2125.15	709.39	3		43.3	19.8	23.3	13.7
			99	VAPEEHPVLLTEAPLNPK	iTRAQ4plex@N-term; Lys->Arg@18		-0.04	2125.12	709.38	2125.17	709.40	3		20.6	24.9	33.5	21.0
			99	VAPEEHPVLLTEAPLNPK	iTRAQ4plex@N-term; Lys->Gln@18		0.01	2097.14	700.05	2097.12	700.05	3		13.8	21.1	35.2	29.9
			99	YPIEHGIVTNWDDMEK	iTRAQ4plex@N-term; Methyl(E)@4; iTRAQ4plex(K)@16		-0.01	2248.10	750.37	2248.11	750.38	3		10.8	27.0	39.8	22.4
													Average:	19.1	21.6	31.8	27.5