Published in final edited form as:

J Mol Biol. 2015 January 30; 427(2): 328–340. doi:10.1016/j.jmb.2014.10.024.

Conformation dependent epitopes recognized by prion protein antibodies probed using mutational scanning and deep sequencing

Kyle M. Doolan and David W. Colby

Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware 19716, USA.

Abstract

Prion diseases are caused by a structural rearrangement of the cellular prion protein, PrP^C, into a disease-associated conformation, PrP^{Sc}, which may be distinguished from one another using conformation specific antibodies. We used mutational scanning by cell-surface display to screen 1,341 PrP single point mutants for attenuated interaction with four anti-PrP antibodies, including several with conformational specificity. Single molecule real time gene sequencing was used to quantify enrichment of mutants, returning on average 26,000 high quality full-length reads for each screened population. Relative enrichment of mutants correlated to the magnitude of the change in binding affinity. Mutations that diminished binding of the antibody ICSM18 represented the core of contact residues in the published crystal structure of its complex. A similarly located binding site was identified for D18, comprising discontinuous residues in helix 1 of PrP, brought into close proximity to one another only when the alpha helix is intact. The specificity of these antibodies for the normal form of PrP likely arises from loss of this conformational feature after conversion to the disease-associated form. Intriguingly, 6H4 binding was found to depend on interaction with the same residues, among others, suggesting that its ability to recognize both forms of PrP depends on a structural rearrangement of the antigen. The application of mutational scanning and deep sequencing provides residue-level resolution of positions in the protein-protein interaction interface that are critical for binding, as well as a quantitative measure of the impact of mutations on binding affinity.

Graphical abstract

Correspondence to David W. Colby: Department of Chemical and Biomolecular Engineering, University of Delaware, 150 Academy Street, Newark, Delaware 19716, USA. Tel: +1 302 831 3649; Fax: +1 302 831 1048; colby@udel.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

EP1802Y; yeast surface display; epitope mapping; SMRT; discontinuous epitope

INTRODUCTION

The specificity of protein-protein interactions mediates many biological processes including receptor-ligand binding, protein signaling cascades, cell adhesion, and antibody recognition. Gene based approaches to investigating protein-protein interactions have made use of surface display technologies[1–4], in which proteins are expressed on the surface of a cell, virus, or ribosome, directly linking protein function with its underlying genetic basis. Surface display has been used to identify antibody binding epitopes through the identification of mutations which diminish binding [5–8]. In particular, yeast surface display has enabled the identification of conformational epitopes by taking advantage of yeast's eukaryotic protein folding machinery[6]. Deep sequencing has been used in conjunction with surface display approaches to identify peptide ligand binding sites of the WW domain[9], examine the protein fitness landscape of engineered influenza binding proteins[10], probe CDR biases in the isolation of poliovirus receptor antibodies[11], and explore CH3 domain stability[12]. In these studies, reversible dye-terminator deep sequencing technology was used to obtain a large number of sequences; however, the technology does not generate individual reads longer than 250 base pairs, requiring the construction of complex libraries and use of tandem reads to investigate whole proteins [13, 14]. Alternatively, single molecule real time (SMRT) deep sequencing[15] permits investigation of open reading frames spanning more than 600 base pairs[16].

Prions cause fatal neurodegenerative diseases which may be transmissible, genetic, or sporadic in etiology[17]. The underlying molecular basis for the pathogenesis of these diseases is the structural rearrangement of the cellular prion protein, PrP^C, into a disease-associated conformation, PrP^{Sc}. Conversion of PrP^C into PrP^{Sc} occurs via interaction of the two isoforms, resulting in self-propagation of PrP^{Sc}. Structural characterization of recombinant PrP by NMR and X-ray crystallography suggests a structure consistent with that of PrP^C [18, 19]. However, atomic resolution characterization of the disease-associated prion conformation has been hampered by insolubility. Due to these difficulties, epitope accessibility has been used to gain insight into the structure of PrP^{Sc} [20–25]. While such approaches are critically dependent on accurate epitope definitions, studies of anti-PrP antibody epitopes have raised enigmatic questions [26–28]. For example, while D18, ICSM18, and 6H4 all bind PrP^C and have been shown to have epitopes in helix 1 by peptide technologies [20, 29, 30], only 6H4 is able to recognize PrP^{Sc}[22, 28, 31], although it does

so with a lower binding affinity than it has for PrP^C and thus may still be considered conformation dependent. Recombinant prion amyloid fibers are recognized by D18 only after partial denaturation [32]. Recent investigations into the conformational specificity of D18 and 6H4 have provided insight into the roles of PrP secondary structure and disulfide bond formation [33].

We conjectured that yeast surface display and SMRT sequencing would enable characterization of conformational epitopes, which may require contacts at disparate locations along the entire length of PrP. When fully processed, PrP is a 208 amino acid protein in humans and contains a single disulfide bond. We used yeast surface display and SMRT sequencing to identify epitope residues critical to ligand binding for four anti-PrP antibodies: D18[30], ICSM18[34], 6H4[20] and EP1802Y. Exogenous expression of PrP in yeast produces protein that is post-translationally processed similarly to PrP^C in mammalian cells[35]. A mutant PrP library was constructed, sorted by fluorescence-activated cell sorting (FACS) for mutations diminishing antibody binding, and the responsible mutations were identified by SMRT sequencing. Statistical analysis of the genetic sequences identified critical epitope residues for each antibody and demonstrated detection of secondary and tertiary structure dependent contacts.

RESULTS AND DISCUSSION

Library construction and characterization

To display PrP on the surface of yeast, the gene encoding mouse PrP was subcloned for expression as a fusion to the yeast mating protein, Aga2p, with a c-myc tag at the C-terminus to verify full-length expression. Based on this construct, a library of PrP mutants was constructed by error prone PCR amplification and transformed into yeast to produce a library containing 8.9×10^5 total members. We performed SMRT sequencing on the 633 base pair DNA fragments encoding the library in order to determine the frequency of mutations. Sequencing returned 31,001 high quality reads containing 27,721 total nucleotide mutations. The sequencing error rate of 5.90×10^{-6} errors/nucleotide was found to be independent of nucleotide position (Supplementary Figure 1), suggesting a small contribution of sequencing errors to the estimated depth of mutagenesis (<1%). Translation of the genetic sequences to their corresponding protein sequences gives rise to an effective protein mutation rate of 0.593 mutations/protein with a distribution that favors single mutation sequences over those with two or more mutations (Table 1). The mutations were found to be distributed about the entire length of the sequence (Supplementary Figure 2), and demonstrated nucleotide substitution biases similar to published reports [36] (Supplementary Table 1)..

Identification of PrP residues which modulate binding affinity

Genes encoding single chain variable fragments (scFvs, Supplementary Table 2) of anti-PrP antibodies D18, ICSM18, and 6H4 were synthesized for recombinant expression by yeast as a secreted protein. Secreted scFvs bound native PrP expressed by N2a cells (Supplementary Figure 3). Anti-PrP scFvs and EP1802Y, a commercially available anti-PrP MAb, were each incubated with the yeast surface display PrP library and binding was analyzed by flow

cytometry (Figure 1A). For each scFv, two major populations were observed: yeast that do not express PrP on their surface due either to plasmid loss or because the PrP contained mutations which made expression unfavorable, and yeast that express PrP with a sequence recognized by the anti-PrP scFvs. Clones appearing within the polygon labeled “Sort” express PrP containing mutations that reduce antibody affinity. This population was isolated by 2–4 successive rounds of FACS at concentrations greater than the wild type K_d for each antibody (Supplementary Table 3). The isolated subpopulations exhibited reduced recognition by the scFvs (Figure 1B). For ICSM18-scFv, D18-scFv, and Mab EP1802Y, less than 1% of cells with any binding at the sort concentrations were observed while for 6H4-scFv approximately 8% of cells had a binding profile comparable to wild-type PrP. Gel electrophoresis of digested fragments of the plasmids within the sorted populations revealed DNA fragments corresponding to the molecular mass of *Prnp*, which was excised for sequencing, as well as a smaller fragment consistent with the insert of the parental vector.

Deep sequencing was performed on the *Prnp* gene fragments of the sorted populations to identify mutations in PrP which reduce antibody binding affinity. Sequencing yielded 19,000–30,000 high quality reads for each sorted population (Table 2). The number of wild-type sequences remaining in the sorted population was used as an experimental measure of the sorting stringency and showed that the population sorted for diminished binding to ICSM18-scFv was the most stringently sorted (4.3% wild-type), with other populations showing less stringent sorting (6H4-scFv 7.6% wild-type, Mab EP1802Y 11.3% wild-type, and D18-scFv 35.1% wild-type). The number of sequences encoding a single mutation at each amino acid position was determined (Figure 1C), and their enrichment relative to wild type sequences was calculated (Figure 1D). The enriched amino acid positions for both ICSM18-scFv and D18-scFv indicated that mutations of two residues in helix 1 predominantly disrupted binding, with modest enrichment of additional mutations to residues in helix 1 and helix 3 for D18-scFv. 6H4-scFv binding was disrupted by mutating residues in helix 1, previously identified as its epitope[20], as well as several residues in helix 3. Binding of the EP1802Y antibody was disrupted by mutations in a largely linear sequence of residues in helix 3 located at positions 217–226.

Sequences containing two or more mutations generally contained at least one of the deleterious mutations identified in the data containing a single mutation (Supplementary Table 4). Only in the case of ICSM18-scFv, which was the most stringently sorted population, did the multiple mutations data identify additional substitutions that may impact binding, N151D (238 observations).

The error rate of sequencing for the sorted populations was calculated at 3.83×10^{-5} errors/nucleotide on average, corresponding to a Q-value of 47.9, and suggesting 12% of apparent mutations actually arose from sequencing errors. However, the number of unique sequences returned exceeded the estimated number of clones in the sorted populations, suggesting that the sequencing error rate may have been higher than calculated. Based on these observations, we estimated the sequencing error to be 2.58×10^{-4} errors/nucleotide, although we cannot exclude the possibility that these unique sequences arose from sources other than sequencing error. Even this higher estimate of the sequencing error rate is negligible compared to the error associated with chi-squared testing where multiple

observations of a specific mutation were necessary to obtain statistical significance. However, if this rate of sequencing error was obtained in the initial library as well, our estimate of the number of true single point mutations would be reduced to 17.85%, providing 940-fold coverage rather than 1,143-fold, which is more than sufficient for the current study.

Determination of contact residues from analysis of substitutions

The preceding analysis broadly identified epitope regions; however, the affinity could have been reduced upon mutation for reasons other than the loss of a critical interaction residue, including the introduction of unfavorable interactions and alternate folding of PrP secondary or tertiary structure. Analysis of the substitutions present at each position may help identify which is the case. We calculated the enrichment relative to wild type for all substitutions within the sorted populations based on the number of times a substitution was observed at a particular position in the sorted populations, the number of codons corresponding to the introduced amino acid arising from a single nucleotide change, and the observed mutation rate. All substitutions that were enriched ($p < 10^{-5}$) are identified in Supplementary Tables 5–8.

Residues in contact with the antibody probed are likely to have diminished binding, and thus be enriched in the sorted populations, when mutated to most if not all substitutions present in the initial library. When examining those residues with the highest positional enrichment, this characteristic was generally observed (Figure 2A). In some instances, most available substitutions were enriched at a particular residue, while one or more substitutions available by a single nucleotide change were not. We observed this for D201 and Y148 in the population sorted for diminished 6H4-scFv binding (Figure 2B), where substitution of glutamic acid for aspartic acid at position 201 or substitution of phenylalanine for tyrosine at position 148 were not observed to be deleterious to 6H4-scFv binding, presumably due to physicochemical similarity of the sidechains. Reduction in expression efficiency may also account for the reduction or absence of particular substitutions, as was the case for the D201A mutation (Supplementary Figure 4), which was only moderately enriched compared to the majority of the other substitutions at D201.

The analysis of substitutions present also identifies substitutions that are not permissive to antibody recognition due to the introduction of repulsive interactions or changes to secondary or tertiary structure. At position E151 in the ICSM18-scFv sorted population, only E151K diminished binding (Figure 2C), likely as a result of the net +2 charge change introducing repulsive interactions. All possible substitutions at E151 express sufficiently well to be observed for the 6H4-scFv sorted population, suggesting that the glutamic acid side chain at position 151 does not contribute substantially to affinity in the ICSM18-scFv binding under the conditions tested. Studies have demonstrated persistence of alpha helix 1 in PrP for a variety of amino acid substitutions in this region [37, 38], including specifically the E151K substitution [37]. Substitution to proline and glycine prove to be exceptions, as these substitutions are known to break alpha helical structures based on allowable dihedral angles [39]. We observed that mutations to proline and glycine at R150 disrupted 6H4-scFv binding (Figure 2D). Finally, we also observed that mutations of several core residues,

including M212, reduced 6H4-scFv binding, consistent with the observation that perturbation of core residues in PrP disrupt tertiary structure [40].

In order to identify only potential contact residues, we first identified residues demonstrating enrichment in more than half of the substitutions arising from a single nucleotide mutation. We then surveyed the literature on alpha helical conformation tolerance for the remaining residues to exclude those that introduce structural changes. The F197S mutation is associated with Gerstmann–Sträussler–Scheinker syndrome[41], significantly reduces PrP stability[42] and alters PrP folding[43]. For this reason it was excluded. Studies into the sensitivity of hydrophobic core residues to substitution[44] or oxidation[40] revealed that modifications to core residues produce a protein structure rich in alpha-helical character but with reduced stability, and an altered orientation of residues in helix 1. We excluded the previously identified hydrophobic core residues Y156, V160, Y156, V209, M204, M205, and M212 based on these studies. Finally, we excluded proline and glycine residues such as P157 and G227 based on dihedral angle considerations[39].

Antibody binding affinity for mutants correlates with enrichment

In order to experimentally validate our identification of substitutions that modulate binding and examine the relationship between the enrichment of specific substitutions and contribution to binding affinity, we selected clones from the sorted libraries for further investigation. We first titrated the concentrations of ICSM18-scFv (Figure 3A) and D18-scFv (Figure 3B) incubated with wild-type, W144S and N152D PrP expressed on the yeast surface in order to determine differences in binding affinity. We found that the W144S mutation resulted in ablation of the interaction with both D18-scFv and ICSM18-scFv with PrP over the concentration range tested; while the N152D again ablated D18-scFv binding over the same concentration range and reduced the affinity of ICSM18-scFv with dissociation constant increase from 153 ± 67 nM to 3.22 ± 1.28 μ M (95% confidence interval). This result is consistent with the observation of these substitutions in the isolated population, though N152D mutations were only observed in sequences with multiple mutations for ICSM18-scFv, highlighting the fact that increasing stringency reduces the number of residues identified by this type of analysis.

We then examined the affinity of 6H4-scFv for five PrP mutants, in addition to wild type, over a range of enrichment values in order to quantitatively examine the relationship between affinity and enrichment (Figure 3C,D). Highly enriched substitutions such as N152D, D201G, and M212V were found to have the greatest reduction in affinity. The modestly enriched F140S substitution was found to have a small but statistically significant reduction in affinity compared to wild-type ($p < 0.05$). The K203E substitution, which was not enriched in the sorted population, did not significantly impact 6H4-scFv binding. The data experimentally demonstrate a positive correlation between enrichment and affinity with a Spearman coefficient of 0.94 ($p = 8.3 \times 10^{-3}$).

Spatial orientation of residues modulating binding affinity

The structural organization of residues enriched in the sorted populations were visualized by mapping onto the structure of mouse PrP [45] (Figure 4A–C). The most highly enriched

residues for both ICSM18-scFv and D18-scFv form a continuous interface on the outward face of helix 1. Those for 6H4-scFv reside along the inner face of helix 1 and at the interface between helix 1 and helix 3 (D201) and also forms a continuous binding pocket. In contrast to the other antibodies tested, a linear series of residues that are solvent exposed were enriched for EP1802Y. Mapping of the epitope residues onto the mouse prion protein structure identifies epitopes that are dependent on secondary and tertiary structure and characterizes two different binding pockets (that for ICSM18-scFv/D18-scFv and 6H4-scFv) along alpha helix 1.

The potential that some identified residues modulate binding as the result of the introduction of negative interactions or conformational changes can be observed by weaker colored regions or highlighting of residues buried deep within the hydrophobic core. Identified contact residues are shown in Figure 4D.

For each antibody tested, the number of contact residues identified by this analysis were inferred to be dependent upon the stringency of sorting (Supplemental Table 3). More stringent sorting conditions favor the identification of only mutations which dramatically reduce binding affinity, as was the case for ICSM18-scFv, for which two core residues are identified here of the six contact residues found in the crystal structure. Less stringent sorting conditions allow for the detection of mutations with intermediate and low impact on binding affinity. Therefore, the positive identification of contact residues here comes with the caveat that residues not identified may be in contact with the antibody but that their mutation causes a smaller decrease in binding affinity than those isolated.

Comparison of contact residues identified with epitopes determined by established methods

The contact residues identified here by mutagenic library sorting and deep sequencing are generally in agreement with previously reported epitopes for these antibodies using alternative methods, and are more precisely defined compared to those identified using peptide arrays (Table 3). The epitopes of D18 and ICSM18 were originally identified by peptide recognition and were determined to include overlapping 25 and 11-residue stretches of PrP which include helix 1 [29, 30]; the residues identified as the epitope of both antibodies here are present on the overlapping section of these peptides. The reported co-crystal structure for ICSM18 in complex with human PrP [46] shows a similar set of discontinuous epitope residues comprised of externally facing alpha helix 1 residues. In the crystal structure, S142, D143, R150 and E151 were identified by proximity in the crystal structure as epitope residues, however, the mutation data suggest that while these residues may be in close contact with the antibody, their contribution to binding affinity are small for the ICSM18-scFv/MoPrP interaction. Differences in the antibody format and species variability of PrP in the region of the interface may also play a role providing a distinct epitope.

For 6H4, the mutagenic library sorting and deep sequencing approach identified an epitope composed of several residues in helix 1, previously identified as the region of the epitope by peptide array, as well as D201 located in helix 3 at its interface with helix 1, a contribution that was not detected by peptide array [20]. The availability of a minimal epitope composed

of helix 1 only permits detectable binding to linear peptides at high concentrations despite the reduction in affinity due to the absence of D201 by direct mutation to D201 or potentially by mutations to the buried methionine residues on helix 3. The reduction, but not ablation, of PrP binding by 6H4 following the disruption of the disulfide bond has been described [33], consistent with an interaction that depends on tertiary structural organization.

CONCLUSIONS

In this study we identify epitope profiles for the variable domains of three antibodies targeting helix 1 of PrP and one monoclonal antibody targeting the C-terminus of helix 3 with residue level resolution. Two antibodies, ICSM18 and D18, were found to be dependent on interactions with discontinuous residues along the outer face of helix 1 that form a continuous surface only when the alpha helix is formed (Figure 4), indicating the mechanism by which these antibodies recognize PrP^C but not PrP^{Sc}, as conversion of PrP^C to PrP^{Sc} is accompanied by a loss of alpha helical secondary structure elements and an increase in beta content[47]. Some of the residues responsible for 6H4 interaction with PrP which exist at the helix1/helix3 interface are inaccessible in a majority of the NMR structures in the PrP^C configuration where some of the residues (E145, N152, and D201) are <15% solvent accessible. This suggests that either these residue become accessible for direct contact as the result of a structural rearrangement or that these residues play a critical role in maintenance of the alpha helical PrP^C structure. Structural dynamics may play a role in allowing access to D201 where the sidechain is accessible in two of the 20 NMR models. A configuration allowing for surface accessibility of contact residues is either already present in PrP^{Sc}, or more likely, may be achieved at a comparable energetic cost, as 6H4 can bind both conformations[31]. Modulation of binding affinity by D201 in helix 3, which is in close proximity to helix 1 in PrP^C (Figure 4), may also explain differences in the strength of interaction of 6H4 with PrP^C and PrP^{Sc}. This observation is also consistent with the decreased affinity of 6H4 for PrP following denaturation[33].

A variety of established techniques are available for the characterization of antibody epitopes including peptide arrays, alanine scanning, and co-crystallization. Peptide arrays are experimentally simple and inexpensive to employ, but offer limited resolution and often cannot identify tertiary structural requirements. Mutational scanning and deep sequencing permits the contributions of each amino acid position to be evaluated in biologically relevant conditions, much like alanine scanning [48]. However, evaluation of many different substitutions at each position may provide information about reactivity across species or naturally occurring genetic variants. We found this to be the case where the N142S mutation was not enriched in the 6H4-scFv sorted population, an attribute that is likely important for the recognition of native human PrP by 6H4 where serine is the wild-type amino acid [20]. We identified the Q222E mutation as ablative in the EP1802Y sorted population suggesting that the antibody may have reduced or ablated affinity to mink or elk PrP where glutamic acid is the native amino acid. An important consideration for mutational scanning is the possibility that some substitutions may dramatically alter global conformations by the introduction of unfavorable interactions. This concern is minimized but not entirely eliminated in alanine scanning and can be partially identified in mutational scanning by examining the enrichment of all constructed substitutions at each position. Both alanine

scanning and mutational scanning approaches are inherently limited to residue level resolution and are unable to match the atomic resolution of co-crystallization structural determination. The residue level resolution of the method allows residues that do not contribute significantly to binding affinity to be identified, such as E151 for ICSM18-scFv, R150 for D18-scFv and Y224 for Mab EP1802Y, permitting identification of potential hotspots.

The mutagenic library sorting and SMRT sequencing approach taken here may be applied to the study of other antibody-protein or protein-protein interactions, including conformation dependent antibodies that recognize other targets [49, 50]. The primary method limitation is the requirement that the protein be expressed in the yeast surface display construct; however, owing to yeast eukaryotic protein folding machinery it is amenable to many types of proteins including high molecular weight and multi-domain proteins[51–54]. HiSeq sequencing generates a greater number of reads than SMRT sequencing, an attribute that can be leveraged for more detailed calculations of the protein fitness landscape. However, the short read lengths are most amenable to targeted sequencing approaches that are dependent on *a priori* identification of areas of interest. Additionally, SMRT sequencing has the ability to identify double mutants regardless of their genetic proximity to one another. The ability to analyze double mutants also enables the possibility of identifying epistasis of genetically distant residues and may be useful for libraries with increased coverage of the double mutant sequence space. In this work we obtained full-length sequence coverage using SMRT sequencing and identified conformational contributions that may not have been identified by targeting the established epitopes.

MATERIALS AND METHODS

Production of anti-PrP scFv

Gene and protein sequences for D18, ICSM18, and 6H4 were obtained from NCBI databases (Supplementary Table 2) and gene sequences were commercially synthesized (Genescript USA Inc.) with a c-terminal FLAG tag for detection. Synthesized genes were cloned into yeast secretion vector pITY, linearized by MfeI digestion and integrated into Ty transposons of the yeast genome using a G-418 selectable marker as previously described[55]. The YVH10 yeast strain, which is a BJ5464 derivative strain containing an additional copy of protein disulfide isomerase inserted in tandem with the exogenous copy, was used [56]. Clones plated onto 300 µg/mL G-418 agar plates were subsequently screened in galactose media to drive scFv secretion under the Gal1-10 promoter to identify highly productive clones[57]. Selected clones were grown in 1L of galactose media for 72–96 hours at 20°C. Cells were harvested by centrifugation at 3,000×g and supernatant was concentrated by ultrafiltration (Millipore UFC701008) using 10kDa filters. Concentrated scFv concentrations varied from 0.5–100 µM as determined by anti-FLAG chemiluminescence and comparison to a FLAG standard (Sigma P7457). Mab EP1802Y was purchased from a commercial source (Abcam ab52604).

Generation of mutant PrP library

The wild-type mouse PrP gene encoding the mature form of the protein (residues 23–231) was obtained by removal of the 3F4 epitope tag from a PrP expression vector by site-directed mutagenesis (Addgene plasmid 1321)[58]. The wild-type PrP gene was cloned into yeast surface display vector pCTCON2[59] to provide C-terminal fusion to yeast mating protein Aga2 and a tryptophan selection marker. Mutagenic library construction was performed as previously described [60], without the addition of nucleotide analogs and using only a single 30 cycle *Taq* polymerase amplification. 9.5µg of the resulting library was combined with 0.5µg of pCTCON2 plasmid backbone and transformed into 50 µL of electrocompetent EBY100 strain yeast [2] ($\sim 2 \times 10^8$ yeast). The transformation and homologous recombination led to a library of 8.9×10^5 members determined by serial dilution plating onto tryptophan deficient plates.

Yeast labeling and sorting for ablated antibody binding

The yeast mutant PrP mutant library was grown at 30°C overnight in tryptophan deficient glucose media, harvested by centrifugation, resuspended to OD ~ 1.0 in tryptophan deficient galactose media, and grown for 20–24 hrs at 20°C to induce protein expression. Exogenous expression of PrP in yeast produces protein that is post-translationally processed similarly to PrP^C in mammalian cells[35]. Yeast labeling was performed as previously described[59], using near saturating scFv/antibody and saturating anti-cmyc concentrations (Mouse anti-FLAG M2 Sigma F1804, chicken anti-cmyc Invitrogen A21281, goat anti-rabbit PE Invitrogen P2771MP, goat anti-mouse-Alexa Fluor 647 Invitrogen A21235, goat anti-chicken Alexa Fluor 488 Invitrogen A11039). Labeled yeast were analyzed on an Accuri C6 flow cytometer and sorted on a FACS Aria or FACS Aria II. Yeast were sorted for 2–4 successive rounds with outgrowth in tryptophan deficient glucose media and expression in tryptophan deficient galactose media between each round until the cmyc+/antibody+ population was eliminated.

For clonal analysis, yeast clones were selected from the sorted sub-libraries by dilution onto tryptophan deficient plates or were constructed by site-directed mutagenesis and gene sequences were determined by Sanger sequencing. Clones were grown, induced for surface-displayed protein expression, and labeled as described above. Titrations for each clone were performed spanning over four-orders of magnitude in antibody concentration. To determine the normalized fluorescence value, the minimum and maximum fluorescence values were obtained from the wild-type with the mean fluorescence of the expressing population was used. A one-parameter model was used to determine the dissociation constant K_d ,

$$\text{Normalized Fluorescence Value} = \frac{[\text{Antibody}]}{[\text{Antibody}] + K_d}$$

Mutations leading to less than 0.15 normalized fluorescence value throughout all concentrations tested were fit to a horizontal line representing the average normalized fluorescence value.

Isolation of mutant PrP genes

Sorted libraries were grown overnight in tryptophan deficient glucose media and 0.5mL culture was used for plasmid DNA extraction using the Zymo Prep Yeast Plasmid Miniprep II kit (Zymo Research D2004). DNA was further purified and concentrated by isopropanol/ethanol precipitation to remove inhibitors. Purified DNA was used to transform MegaX DH10B T1R Electrocomp Cells (Invitrogen C640003) by electroporation. Transformed cells were passaged into 50 mL Luria-Bertani broth with 100µg/mL ampicillin for selection and serial dilution onto Luria-Bertani agar plates with 100µg/mL ampicillin was performed to determine library size. Typical transformations provided a total of 4,000–30,000 colony forming units per library. E. coli cultures were passaged once and 100mL were grown in Luria-Bertani broth with 100µg/mL ampicillin for DNA extraction using two QIAGEN Plasmid Plus Midi columns (QIAGEN #12943). Bacterially derived plasmid DNA was digested with NheI-HF and SalI-HF restriction enzymes (New England Biolabs R3131S & R3138S) to liberate the 633 base pair fragment containing the PrP genes with ablating mutations. DNA fragments were purified on a 2% agarose gel, extracted using 2–4 QIAquick Gel Extraction columns (QIAGEN 28704) and further purified by isopropanol/ethanol precipitation. Final DNA preparations contained 0.5–50 µg total DNA.

SMRT sequencing and read filtration

End repair was performed on restriction digested DNA and samples were prepared for SMRT sequencing using the DNA Template Prep Kit 2.0 (Pacific Biosciences). Samples were run in a single SMRT Cell with 2 × 45 min processing on PacBio RS II (Pacific Biosciences).

Data was initially processed by using the SMRT Portal RS_ReadsOfInsert module that utilizes the Quiver algorithm to obtain a maximum likelihood consensus sequence. The consensus sequences and accompanying quality values were exported in .fastq format and further processed using a linux script (Supplementary Information). Briefly, a pairwise sequence alignment was performed for each consensus sequence and the wild-type *Prnp* sequence using the needleall program from the EMBOSS suite[61]. Consensus sequences that produced poor alignments were removed from consideration. From the alignment both nucleotide and protein mutations were identified. Sequences containing a cysteine mutation or producing a premature stop codon were censored because of the possibility of alternative disulfide bond formation. Ultimately, each sequencing provided 19,000–30,000 high quality reads (Table 2) with an average Q-Value of 40.00 to 55.22 corresponding to error rates from 1.00×10^{-4} to 3.00×10^{-6} errors/nucleotide.

The algorithm used to calculate the error rate (Quiver) was trained on data providing a template for alignment; we chose to construct de novo assemblies for the sequencing subreads to provide consensus sequences and then used a traditional Needleman-Wunsch algorithm for alignment of the consensus sequences to prevent wild-type bias. We observed a higher appearance of sequences observed only one or a few times, considering the estimated coverage was approximately 10 fold suggesting that the actual error rate may be as high as 2.58×10^{-6} (Q-Value of 35.9). Filtering removed on average 13.6% of sequences

(CCS Reads) due to low quality, consistent with other similar approaches for SMRT sequencing[62] (See Supplementary Methods for detailed account of data filtering).

Statistical analysis of mutations by position and substitution

Single mutant reads were used to identify epitope residues involved in antibody binding to PrP. Enrichment values were calculated for each position based on the number of times a mutation was observed in the sorted populations, the number of codons corresponding to the introduced amino acid arising from a single nucleotide changes, the nucleotide substitution biases, and the observed sorting stringency. p-values were calculated using the χ^2 statistic for each pairwise comparison. The polymerase mutational biases in the library were characterized (Supplementary Table 1). For comparison to traditional epitope identification methods, epitope residues were identified as residues showing statistically significant positional enrichment (p-value < 10^{-5}) and statistically significant enrichment (p-value < 10^{-5}) for more than half of the substitutions arising from a single nucleotide mutations.

Mapping of Epitope Residues onto Mouse PrP NMR Structure

The NMR model coordinates of the structured domain of the mouse PrP (121–231) were obtained from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (PDB ID: [2L39](#), model 1)[45] and were loaded into the Pymol[63] for visualization. Surface accessibility was determined using the surface accessibility calculator of the DeepView/Swiss-PdbViewer[64] using a greater than or equal to 15% amino acid surface accessibility criteria.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Olga Shevchenko for assistance with the SMRT sequencing and Bruce Kingham, Karol Miaskiewicz, and Lynn Opdenaker for helpful discussions. Yeast expression plasmids pCTCON2 and pITY and yeast strains EBY100 and YVH10 were a kind gift from the laboratory of K. Dane Witttrup.

FUNDING

This work was supported by the NIH/NIGMS CBI Program [5 T32 GM008550-17] and NIH/NIGMS IDeA Program [8 P20 GM103446-13]. Additional support from NIH NS064173 and the University of Delaware Center for Bioinformatics and Computational Biology Core Facility was made possible through funding from the NIH (5 P20 RR016472-12), and NSF (EPS-081425).

Abbreviations Used

PrP^{Sc}	prion protein conformation associated with disease
PrP^C	native prion protein conformation
SMRT	single molecule real time
<i>Prnp</i>	mouse prion protein gene
PrP	prion protein

FACS	Fluorescence-activated cell sorting
scFv	single chain variable fragment
Fab	Fragment antigen-binding
Taq	<i>Thermus aquaticus</i>
Aga2	a-agglutinin adhesion subunit 2

References

1. Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*. 1985; 228:1315–7. [PubMed: 4001944]
2. Boder ET, Wittrup KD. Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol*. 1997; 15:553–7. [PubMed: 9181578]
3. Georgiou G, Stathopoulos C, Daugherty PS, Nayak AR, Iverson BL, Curtiss R 3rd. Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat Biotechnol*. 1997; 15:29–34. [PubMed: 9035102]
4. Hanes J, Pluckthun A. In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A*. 1997; 94:4937–42. [PubMed: 9144168]
5. de la Cruz VF, Lal AA, McCutchan TF. Immunogenicity and epitope mapping of foreign sequences via genetically engineered filamentous phage. *J Biol Chem*. 1988; 263:4318–22. [PubMed: 2450091]
6. Chao G, Cochran JR, Wittrup KD. Fine epitope mapping of anti-epidermal growth factor receptor antibodies through random mutagenesis and yeast surface display. *J Mol Biol*. 2004; 342:539–50. [PubMed: 15327953]
7. Christmann A, Wentzel A, Meyer C, Meyers G, Kolmar H. Epitope mapping and affinity purification of monospecific antibodies by Escherichia coli cell surface display of gene-derived random peptide libraries. *J Immunol Methods*. 2001; 257:163–73. [PubMed: 11687250]
8. Osada E, Shimizu Y, Akbar BK, Kanamori T, Ueda T. Epitope mapping using ribosome display in a reconstituted cell-free protein synthesis system. *J Biochem*. 2009; 145:693–700. [PubMed: 19228777]
9. Fowler DM, Araya CL, Fleishman SJ, Kellogg EH, Stephany JJ, Baker D, et al. High-resolution mapping of protein sequence-function relationships. *Nat Methods*. 2010; 7:741–6. [PubMed: 20711194]
10. Whitehead TA, Chevalier A, Song Y, Dreyfus C, Fleishman SJ, De Mattos C, et al. Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing. *Nat Biotechnol*. 2012; 30:543–8. [PubMed: 22634563]
11. Larman HB, Xu GJ, Pavlova NN, Elledge SJ. Construction of a rationally designed antibody platform for sequencing-assisted selection. *Proc Natl Acad Sci U S A*. 2012; 109:18523–8. [PubMed: 23064642]
12. Traxlmayr MW, Hasenhindl C, Hackl M, Stadlmayr G, Rybka JD, Borth N, et al. Construction of a stability landscape of the CH3 domain of human IgG1 by combining directed evolution with high throughput sequencing. *J Mol Biol*. 2012; 423:397–412. [PubMed: 22846908]
13. Melamed D, Young DL, Gamble CE, Miller CR, Fields S. Deep mutational scanning of an RRM domain of the *Saccharomyces cerevisiae* poly(A)-binding protein. *RNA*. 2013; 19:1537–51. [PubMed: 24064791]
14. DeKosky BJ, Ippolito GC, Deschner RP, Lavinder JJ, Wine Y, Rawlings BM, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol*. 2013; 31:166–9. [PubMed: 23334449]
15. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. *Science*. 2009; 323:133–8. [PubMed: 19023044]

16. Smith CC, Wang Q, Chin CS, Salerno S, Damon LE, Levis MJ, et al. Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature*. 2012; 485:260–3. [PubMed: 22504184]
17. Colby DW, Prusiner SB. Prions. *Cold Spring Harb Perspect Biol*. 2011; 3:a006833. [PubMed: 21421910]
18. Riek R, Hornemann S, Wider G, Billeter M, Glockshuber R, Wuthrich K. NMR structure of the mouse prion protein domain PrP(121–231). *Nature*. 1996; 382:180–2. [PubMed: 8700211]
19. Knaus KJ, Morillas M, Swietnicki W, Malone M, Surewicz WK, Yee VC. Crystal structure of the human prion protein reveals a mechanism for oligomerization. *Nat Struct Biol*. 2001; 8:770–4. [PubMed: 11524679]
20. Korth C, Stierli B, Streit P, Moser M, Schaller O, Fischer R, et al. Prion (PrP^{Sc})-specific epitope defined by a monoclonal antibody. *Nature*. 1997; 390:74–7. [PubMed: 9363892]
21. Peretz D, Williamson RA, Matsunaga Y, Serban H, Pinilla C, Bastidas RB, et al. A conformational transition at the N terminus of the prion protein features in formation of the scrapie isoform. *J Mol Biol*. 1997; 273:614–22. [PubMed: 9356250]
22. Stanker LH, Serban AV, Cleveland E, Hnasko R, Lemus A, Safar J, et al. Conformation-dependent high-affinity monoclonal antibodies to prion proteins. *J Immunol*. 2010; 185:729–37. [PubMed: 20530267]
23. Tayebi M, Jones DR, Taylor WA, Stileman BF, Chapman C, Zhao D, et al. PrP(Sc)-specific antibodies with the ability to immunodetect prion oligomers. *PloS one*. 2011; 6:e19998. [PubMed: 21625515]
24. Silva CJ. Using small molecule reagents to selectively modify epitopes based on their conformation. *Prion*. 2012; 6:163–73. [PubMed: 22436143]
25. Pan T, Chang B, Wong P, Li C, Li R, Kang SC, et al. An aggregation-specific enzyme-linked immunosorbent assay: detection of conformational differences between recombinant PrP protein dimers and PrP(Sc) aggregates. *J Virol*. 2005; 79:12355–64. [PubMed: 16160162]
26. Li R, Liu T, Wong BS, Pan T, Morillas M, Swietnicki W, et al. Identification of an epitope in the C terminus of normal prion protein whose expression is modulated by binding events in the N terminus. *J Mol Biol*. 2000; 301:567–73. [PubMed: 10966770]
27. Jones M, McLoughlin V, Connolly JG, Farquhar CF, MacGregor IR, Head MW. Production and characterization of a panel of monoclonal antibodies against native human cellular prion protein. *Hybridoma (Larchmt)*. 2009; 28:13–20. [PubMed: 19132894]
28. Khalili-Shirazi A, Kaiser M, Mallinson G, Jones S, Bhelt D, Fraser C, et al. Beta-PrP form of human prion protein stimulates production of monoclonal antibodies to epitope 91–110 that recognise native PrP^{Sc}. *Biochim Biophys Acta*. 2007; 1774:1438–50. [PubMed: 17936697]
29. Khalili-Shirazi A, Quarantino S, Londei M, Summers L, Tayebi M, Clarke AR, et al. Protein conformation significantly influences immune responses to prion protein. *J Immunol*. 2005; 174:3256–63. [PubMed: 15749856]
30. Williamson RA, Peretz D, Pinilla C, Ball H, Bastidas RB, Rozenshteyn R, et al. Mapping the prion protein using recombinant antibodies. *Journal of virology*. 1998; 72:9413–8. [PubMed: 9765500]
31. Fischer MB, Roeckl C, Parizek P, Schwarz HP, Aguzzi A. Binding of disease-associated prion protein to plasminogen. *Nature*. 2000; 408:479–83. [PubMed: 11100730]
32. Novitskaya V, Makarava N, Bellon A, Bocharova OV, Bronstein IB, Williamson RA, et al. Probing the conformation of the prion protein within a single amyloid fibril using a novel immunoconformational assay. *J Biol Chem*. 2006; 281:15536–45. [PubMed: 16569635]
33. Kang HE, Weng CC, Saijo E, Saylor V, Bian J, Kim S, et al. Characterization of conformation-dependent prion protein epitopes. *J Biol Chem*. 2012; 287:37219–32. [PubMed: 22948149]
34. White AR, Enever P, Tayebi M, Mushens R, Linehan J, Brandner S, et al. Monoclonal antibodies inhibit prion replication and delay the development of prion disease. *Nature*. 2003; 422:80–3. [PubMed: 12621436]
35. Li A, Dong J, Harris DA. Cell surface expression of the prion protein in yeast does not alter copper utilization phenotypes. *J Biol Chem*. 2004; 279:29469–77. [PubMed: 15090539]
36. Vanhercke T, Ampe C, Tirry L, Denolf P. Reducing mutational bias in random protein libraries. *Anal Biochem*. 2005; 339:9–14. [PubMed: 15766704]

37. Norstrom EM, Mastrianni JA. The charge structure of helix 1 in the prion protein regulates conversion to pathogenic PrP^{Sc}. *J Virol*. 2006; 80:8521–9. [PubMed: 16912302]
38. Speare JO, Rush TS 3rd, Bloom ME, Caughey B. The role of helix 1 aspartates and salt bridges in the stability and conversion of prion protein. *J Biol Chem*. 2003; 278:12522–9. [PubMed: 12551897]
39. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. *J Mol Biol*. 1963; 7:95–9. [PubMed: 13990617]
40. Younan ND, Nadal RC, Davies P, Brown DR, Viles JH. Methionine oxidation perturbs the structural core of the prion protein and suggests a generic misfolding pathway. *J Biol Chem*. 2012; 287:28263–75. [PubMed: 22654104]
41. Giaccone G, Verga L, Bugiani O, Frangione B, Serban D, Prusiner SB, et al. Prion protein preamyloid and amyloid deposits in Gerstmann-Straussler-Scheinker disease, Indiana kindred. *Proc Natl Acad Sci U S A*. 1992; 89:9349–53. [PubMed: 1357663]
42. Liemann S, Glockshuber R. Influence of amino acid substitutions related to inherited human prion diseases on the thermodynamic stability of the cellular prion protein. *Biochemistry*. 1999; 38:3258–67. [PubMed: 10079068]
43. Apetri AC, Surewicz K, Surewicz WK. The effect of disease-associated mutations on the folding pathway of human prion protein. *J Biol Chem*. 2004; 279:18008–14. [PubMed: 14761942]
44. Hart T, Hosszu LL, Trevitt CR, Jackson GS, Waltho JP, Collinge J, et al. Folding kinetics of the human prion protein probed by temperature jump. *Proc Natl Acad Sci U S A*. 2009; 106:5651–6. [PubMed: 19321423]
45. Damberger FF, Christen B, Perez DR, Hornemann S, Wuthrich K. Cellular prion protein conformation and function. *Proc Natl Acad Sci U S A*. 2011; 108:17308–13. [PubMed: 21987789]
46. Antonyuk SV, Trevitt CR, Strange RW, Jackson GS, Sangar D, Batchelor M, et al. Crystal structure of human prion protein bound to a therapeutic antibody. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:2554–8. [PubMed: 19204296]
47. Pan KM, Baldwin M, Nguyen J, Gasset M, Serban A, Groth D, et al. Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. *Proc Natl Acad Sci U S A*. 1993; 90:10962–6. [PubMed: 7902575]
48. Cunningham BC, Wells JA. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science*. 1989; 244:1081–5. [PubMed: 2471267]
49. Perchiaica JM, Ladiwala AR, Bhattacharya M, Tessier PM. Structure-based design of conformation- and sequence-specific antibodies against amyloid beta. *Proc Natl Acad Sci U S A*. 2012; 109:84–9. [PubMed: 22171009]
50. Moroncini G, Kanu N, Solforosi L, Abalos G, Telling GC, Head M, et al. Motif-grafted antibodies containing the replicative interface of cellular PrP are specific for PrP^{Sc}. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101:10404–9. [PubMed: 15240877]
51. Cochran JR, Kim YS, Olsen MJ, Bhandari R, Wittrup KD. Domain-level antibody epitope mapping through yeast surface display of epidermal growth factor receptor fragments. *J Immunol Methods*. 2004; 287:147–58. [PubMed: 15099763]
52. Rakestraw JA, Aird D, Aha PM, Baynes BM, Lipovsek D. Secretion-and-capture cell-surface display for selection of target-binding proteins. *Protein Eng Des Sel*. 2011; 24:525–30. [PubMed: 21402751]
53. van den Beucken T, Pieters H, Steukers M, van der Vaart M, Ladner RC, Hoogenboom HR, et al. Affinity maturation of Fab antibody fragments by fluorescent-activated cell sorting of yeast-displayed libraries. *FEBS Lett*. 2003; 546:288–94. [PubMed: 12832056]
54. Wang XX, Li Y, Yin Y, Mo M, Wang Q, Gao W, et al. Affinity maturation of human CD4 by yeast surface display and crystal structure of a CD4-HLA-DR1 complex. *Proc Natl Acad Sci U S A*. 2011; 108:15960–5. [PubMed: 21900604]
55. Parekh RN, Shaw MR, Wittrup KD. An integrating vector for tunable, high copy, stable integration into the dispersed Ty delta sites of *Saccharomyces cerevisiae*. *Biotechnol Prog*. 1996; 12:16–21. [PubMed: 8845105]

56. Robinson AS, Hines V, Wittrup KD. Protein disulfide isomerase overexpression increases secretion of foreign proteins in *Saccharomyces cerevisiae*. *Biotechnology (N Y)*. 1994; 12:381–4. [PubMed: 7764684]
57. Shusta EV, Raines RT, Pluckthun A, Wittrup KD. Increasing the secretory capacity of *Saccharomyces cerevisiae* for production of single-chain antibody fragments. *Nat Biotechnol*. 1998; 16:773–7. [PubMed: 9702778]
58. Ma J, Lindquist S. Wild-type PrP and a mutant associated with prion disease are subject to retrograde transport and proteasome degradation. *Proc Natl Acad Sci U S A*. 2001; 98:14955–60. [PubMed: 11742063]
59. Chao G, Lau WL, Hackel BJ, Sazinsky SL, Lippow SM, Wittrup KD. Isolating and engineering human antibodies using yeast surface display. *Nat Protoc*. 2006; 1:755–68. [PubMed: 17406305]
60. Colby DW, Kellogg BA, Graff CP, Yeung YA, Swers JS, Wittrup KD. Engineering antibody affinity by yeast surface display. *Methods in enzymology*. 2004; 388:348–58. [PubMed: 15289082]
61. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet*. 2000; 16:276–7. [PubMed: 10827456]
62. Jiao X, Zheng X, Ma L, Kutty G, Gogineni E, Sun Q, et al. A Benchmark Study on Error Assessment and Quality Control of CCS Reads Derived from the PacBio RS. *J Data Mining Genomics Proteomics*. 2013; 4
63. Schrodinger LLC. The PyMOL Molecular Graphics System. Version 1.3r1. 2010
64. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 1997; 18:2714–23. [PubMed: 9504803]

Highlights

- SMRT sequencing enables high throughput mutational scanning with long read lengths
- Conformation specific antibody D18 binds discontinuous residues in helix 1
- 6H4 depends on a structural rearrangement to bind residues near PrP surface
- Approach identifies distal residues that mediate protein-protein interactions

Figure 1. Deep sequencing subpopulations of PrP mutants sorted for loss of antibody binding affinity identifies critical epitope regions

(A) The initial mutant PrP library displayed on the surface of yeast was labeled with antibodies indicated on left as well as an antibody recognizing a c-myc tag genetically fused to PrP to monitor full-length expression level. The c-myc tag is C-terminal to PrP. The majority of PrP mutants in the library are recognizable by the anti-PrP antibodies, while a small number of mutants have diminished binding to the antibodies, indicated in the polygon labeled “Sort.” (B) Following several rounds of sorting by FACS the resulting subpopulations are primarily composed of mutant PrP clones with diminished binding to the antibodies. (C) Mutations observed by SMRT sequencing of PrP genes contained within populations with diminished antibody binding were mapped by position. (D) Regions of interest where residues containing mutations were enriched in the sorted population relative to wildtype PrP. Calculating the enrichment value accounts for unequal representation of mutations in the initial library and the stringency of sorting.

Figure 2. Substitution distributions of selected residues characterize residue interactions with antibody

The enrichment for each particular substitution is reported for all substitutions arising from a single nucleotide substitution. (A) Epitope residues show enrichment of all or near all substitutions arising from a single nucleotide substitution consistent with the sidechain of these residues directly contacting paratope residues in the antibody (B) Some epitope residues show tolerance of particular substitution, demonstrated by the absence of these substitutions in the sequences obtained from the sorted populations, as a result of physicochemical similarity of the sidechains (C) Positions where only one or a few substitutions show statistically significant enrichment can be the result of particular substitutions that introduce negative interactions (D) Mutations to proline or glycine at many residues were ablative likely by preventing alpha helical secondary structure to develop. Mutations to residues critical to protein folding also led to reduction in affinity.

Figure 3. Experimental validation of critical epitope residues indicates correlation between mutation enrichment and contribution to binding
 wild-type (●), N152D (▲), W144S (◻), K203E (◊), F140S (◆), M212V (▲), and D201G (★) PrP expressed on the yeast surface were incubated with (A) ICSM18-scFv, (B) D18-scFv, and (C) 6H4-scFv at indicated concentrations. (D) The affinities of each PrP mutant, determined from the 6H4-scFv titration, were plotted as a function of the enrichment observed in the 6H4 sorted sub-library, demonstrating a correlation (Spearman rank correlation coefficient = 0.94, $p = 8.3 \times 10^{-3}$) between enrichment and the change in binding affinity. Titration of yeast expressing N152D-PrP did not result in the observance of an increase in fluorescence associated with 6H4-scFv binding preventing calculation of a dissociation constant; N152D as shown represents that the reduction in affinity was greatest among substitutions tested.

Figure 4. Mutational landscape mapped onto native PrP structure reveal structural basis for discontinuous contact residues

(A) The entire mutational landscape arising from single mutant sequences in the sorted sub-libraries were mapped onto the ribbon structure of mouse PrP (PDB ID: [2L39](#)) (B) Magnified view of ribbon structure for epitope region. An arrow is drawn to D201 in the 6H4 condition to highlight its position relative to residues in helix 1. (C) Electron density map for epitope region identify the binding surfaces for each antibody. Despite not contacting adjacent residues, the ICSM18/D18 contact residues form a continuous binding patch along the outside/top of helix 1. The contact residues for 6H4 are composed of residues at the helix 1/ helix 3 interface; arrow is drawn to D201. (D) Electron density map

for epitope region where critical contact residues (characterized by statistically significant enrichment of more than 50% of substitutions arising from single nucleotide changes) are indicated in green while all non-epitope residues are indicated in white.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Mutational frequency observed for PrP library

	Library Size (members)	No. mutations per member				
		0	1	2	3	4
DNA	8.9×10^5	48.5%	27.7%	12.6%	4.7%	6.51%
Protein	8.9×10^5	60.6%	28.2%	8.0%	2.1%	0.98%

Table 2

Library sizes and sequence filtering

Class	Name	Library	D18	ICSM18	6H4	EP1802Y
Cell libraries	Yeast Library Size	8.9×10^5	45–2,500	260–1,500	2,300–8,000	250–2,900
		28,820	5,000	8,600	4,800	10,000
	E.coli Library Size	28,820	5,000	8,600	4,800	10,000
<hr/>						
Raw Sequences	SMRT Subreads	420,326	847,680	399,721	435,086	392,720
	SMRT Sequences	80,444	95,943	81,835	96,383	94,261
<hr/>						
Filtered Sequences	CCS Reads	35,880	45,152	31,628	40,354	31,837
<hr/>						
Alignment	<i>Global Misalignment</i>	440	367	332	534	406
	<i>Local Misalignment</i>	3,756	6,190	4,358	3,706	3,970
	Aligned Reads	31,684	38,595	26,938	36,114	27,461
<hr/>						
Translation	<i>Truncated</i>	683	367	450	9	176
	<i>Cysteine Mutants</i>	(1033)	9,488	7,148	6,073	4,612
	Accepted Sequences	31,001	28,740	19,340	30,032	22,673
<hr/>						
Depth of Mutation	Wild-type Sequences	18,175	10,100	838	2,272	2,554
	Single Mutant Sequences	8,448	12,122	9,489	16,785	12,530
	Multiple Mutation Sequences	4,378	6,518	9,013	10,975	7,589
<hr/>						
Sequence Quality	Average Basecall Q-Score	52.29	55.22	40.00	45.79	46.22

Table 3

Comparison of anti-PrP antibody epitope residues characterized in this and other studies.

Antibody	Form	PrP Species	Method	Epitope	Reference
ICSM18	Mab	Human	Peptide Array	142-GSDYEDRYYREN ^M -154	[Khalili-Shirazi et al. J. Immunol. 2005]
ICSM18	Fab	Human	Co-crystallization	142-GSDYEDRYYREN ^M -154 202-DV ^K M-205	[Antonyuk et al. PNAS 2009]
ICSM18	scFv	Mouse	Mutational Scanning	141-GNDWEDRYYREN ^M -153	This study
D18	Fab	Mouse	Peptide Phage Display	132-SAMSRPMTIHFGNDWEDRYYREN ^M YRYP-158	[Williamson et al. J. Virol. 1998]
D18	scFv	Mouse	Mutational Scanning	132-SAMSRPMTIHFGNDWEDRYYREN ^M YRYP-158	This study
6H4	Mab	Bovine	Peptide Array	153-GSDYEDRYYREN ^M -165	[Korth et al. Nature 1997]
6H4	scFv	Mouse	Mutational Scanning	141-GNDWEDRYYREN ^M YR-155 200-TDV-203	This study
EPI802Y	Mab	Mouse	Mutational Scanning	220-ESQAYYDGRSS-231	This study