JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Role of Prion Disease-Linked Mutations in the Intrinsically Disordered N-Terminal Domain of the Prion Protein

Xiaojing Cong,[†,‡,⊥,○] Nicola Casiraghi,[§,‡,⊥,○] Giulia Rossetti,*[,#,∥,⊥,‡,○] Sandipan Mohanty,[∥] Gabriele Giachin,[†] Giuseppe Legname,[†,▽] and Paolo Carloni*[,‡,⊥]

[†]Laboratory of Prion Biology, Department of Neuroscience, Scuola Internazionale Superiore di Studi Avanzati (SISSA), via Bonomea 265, 34136 Trieste, Italy

[‡]Laboratory for Computational Biophysics, German Research School for Simulation Sciences (GRS), Forschungszentrum Jülich−RWTH Aachen, 52425 Jülich, Germany

[§]Department of Biology, University of Bologna, via Selmi 3, 40126 Bologna, Italy

[∥]Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany

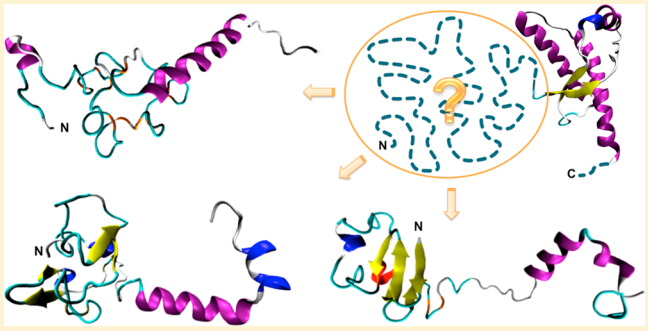
[⊥]Computational Biomedicine Section (IAS-5), Institute of Advanced Simulation (IAS), 52425 Jülich, Germany

[#]Institute for Research in Biomedicine and Barcelona Supercomputing Center Joint Research Program on Computational Biology, Barcelona Science Park, Baldiri I Reixac 10, 08028 Barcelona, Spain

[▽]ELETTRA Laboratory, Sincrotrone Trieste S.C.p.A., 34149 Basovizza, Trieste, Italy

Supporting Information

ABSTRACT: Prion diseases are fatal neurodegenerative disorders in mammals and other animal species. In humans, about 15% of these maladies are caused by pathogenic mutations (PMs) in the gene encoding for the prion protein (PrP^C). Seven PMs are located in the naturally unfolded PrP^C N-terminal domain, which constitutes about half of the protein. Intriguingly and in sharp contrast to other PMs clustered in the folded domain, N-terminal PMs barely affect the conversion to the pathogenic (scrapie, or PrP^{Sc}) isoform of PrP^C. Here, we hypothesize that the neurotoxicity of these PMs arises from changes in structural determinants of the N-terminal domain, affecting the protein binding with its cellular partners and/or the cotranslational translocation during the PrP^C biosynthesis. We test this idea by predicting the conformational ensemble of the wild-type (WT) and mutated mouse PrP^C N-terminal domain, whose sequence is almost identical to that of the human one and for which the largest number of *in vivo* data is available. The conformational properties of the WT are consistent with those inferred experimentally. Importantly, the PMs turn out to affect in a subtle manner the intramolecular contacts in the putative N-terminal domain binding sites for Cu^{2+} ions, sulphated glycosaminoglycans, and other known PrP^C cellular partners. The PMs also alter the local structural features of the transmembrane domain and adjacent stop transfer effector, which act together to regulate the protein topology. These results corroborate the hypothesis that N-terminal PMs affect the PrP^C binding to functional interactors and/or the translocation.



INTRODUCTION

Familial forms of human (Hu) transmissible spongiform encephalopathies (TSEs) or prion diseases are fatal and incurable neurodegenerative disorders. The key event in these maladies is the post-translational conversion of the ubiquitously expressed cellular form of the Hu prion protein (PrP^C) into the misfolded pathogenic isoform, PrP^{Sc}, without any detectable covalent modifications.[1]

Prion diseases are associated with 58 missense or insertional/deletional mutations identified so far in the gene coding for $HuPrP^C$ (*PRNP*). Missense mutations include 44 nonsynonymous codon substitutions, or point mutations (PMs), and five nonsense (or "stop") mutations.[2−4] These mutations lead to neurodegeneration and give rise to abnormal forms of HuPrP in the brain.[2] The PMs are located all over the protein, from the disordered N-terminal domain (N-term_$HuPrP^C$ hereafter, residues 23−124) to the folded C-terminal globular domain (GD, residues 125−230; Figure 1A).

Most PMs are located in the $HuPrP^C$ GD. They may affect GD secondary structure (SS) elements and structure flexibility[5−9] and accelerate the misfolding process *in vitro*.[2,10−12]

Received: June 24, 2013
Published: October 16, 2013

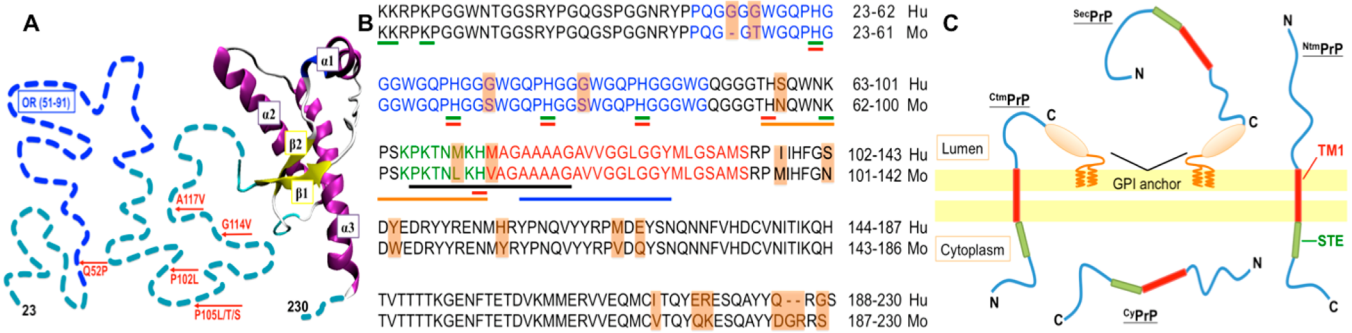


Figure 1. Structure, sequence, and topology forms of PrP^C. (A) Molecular structure of $HuPrP^C$. The GD NMR structure[42] is shown in cartoon format. The N-terminus (102 residues) and the C-terminus (three residues), not present in the NMR structure, are sketched in dashed lines. Red arrows indicate the PMs studied in this work. (B) Alignment of $HuPrP^C$ and $MoPrP^C$ sequences, obtained using BLAST.[86] The distinct residues in each of the two proteins are shadowed in orange. Those belonging to OR, STE, and TM1 are colored in blue, green, and red, respectively. Cu^{2+}, sulphated GAG, vitronectin, STI1, and Aβ multimers binding-site residues (Table 1) are underlined in red, green, black, blue, and orange, respectively. (C) Four topology forms of Hu/Mo PrP and PrP from SHa during biosynthesis:[19] ^{Sec}PrP, ^{Ntm}PrP, ^{Ctm}PrP, and ^{cy}PrP. The signal peptides (residues 1−22 and 231−253), which are cleaved off during the biosynthesis of the proteins, are not shown in this figure.

Table 1. Binding Partners of Hu/Mo/SHa PrP^C N-term so Far Identified by *in Vivo* and *in Vitro* Studies[a]

binding partner	proposed role(s) *in vivo*	PrP^C interacting residues (*in vitro*)	affinity (*in vitro*)	references
sulphated GAG	in the formation of Hu/Mo/SHa PrP^{Sc}[22,87]	in Hu/Mo/SHa PrP^C: K23, K24, K27, H60, H68, H76, H84, K101[23]	N/A	refs22,23, and87
Cu^{2+} ions	antioxidant activity regulation, copper homeostasis, calcium signaling, in the progression of the disease[24]	in Hu/Mo/SHa PrP^C OR: H60, H68, H76, H84	inter-repeat mode: nanomolar range; single-repeat mode: low-micromolar range	reviewed in refs24 and25
		in Hu/Mo/SHa PrP^C outside OR[b]: H95, H110	nanomolar range	
vitronectin	supporting axonal growth[26]	in $MoPrP^C$: residues 104−118[26]	12 nM[26]	ref26
STI1	triggering neuroprotection[27]	in $MoPrP^C$: residues 114−127[27]	140 nM[27]	ref27
Aβ multimers	suppression of long-term potentiation[28−30]	in $MoPrP^C$: residues 95−110[28−30]	N/A	refs28−30
LRP1	in a clathrin-dependent internalization mechanism[31]	N/A	nanomolar range[31]	ref31
NCAM	in the cell adhesion process[88]	N/A	N/A	refs32 and88

[a]The affinity and the binding site identified *in vitro* for these partners are also reported. [b]Few studies using N-terminally truncated murine PrP^C GD also detected a weak binding site in the GD with affinity ∼10 μM.[89,90]

Few PMs are located in N-term_$HuPrP^C$ (Figure 1A). Q52P has been recently found in this domain, and it is likely associated with familial Creutzfeldt−Jakob disease (fCJD);[3] G114V has been reported with both fCJD[13] and Gerstmann−Sträussler−Scheinker syndrome (GSS),[2] whereas P102L, P105L/S/T, and A117V are associated with GSS.[2] The corresponding mutation of P102L in mouse (Mo) PrP^C alters the incubation time in transgenic (Tg) mice upon infection by PrP^{Sc} from Hu, hamster, sheep, or murine sources in a strain-dependent manner.[14,15] These N-terminal PMs cause the disease without affecting the kinetics of the $HuPrP^C$ to $HuPrP^{Sc}$ conversion *in vitro*.[2] The observation that the N-terminal PMs cause Hu TSEs without affecting the thermodynamic stability of PrP^C suggests the presence of other cellular pathways leading to PrP^{Sc} formation.

On the one hand, the N-term_$HuPrP^C$, along with the Mo domain (N-term_$MoPrP^C$, featuring 93% sequence identity with the human domain, see Figure 1B and Supporting Information (SI), Table S1), contains a major part of the so-called transmembrane domain (termed TM1, comprising roughly residues 112−135) and the preceding "stop transfer effector" (STE, a hydrophilic region containing roughly residues 104−111)[16,17] (Figure 1B). STE and TM1 act in concert to control the cotranslational translocation at the endoplasmic reticulum (ER) during the biosynthesis of $HuPrP^C$ and of Mo and Syrian hamster (SHa) PrP^C.[18,19] Four topological forms of the protein can be generated during the translocation[19] (Figure 1C): some PrP (here, we use PrP to distinguish from the mature cellular form PrP^C) nascent chains pass completely through the translocon to generate secretory PrP (^{Sec}PrP). Others integrate into the lipid bilayer with either the N-terminus or the C-terminus in the ER lumen (^{Ntm}PrP and ^{Ctm}PrP, respectively); still others are retained in the cytoplasm (^{cy}PrP). Both the transmembrane forms ^{Ntm}PrP and ^{Ctm}PrP span the lipid bilayer with the TM1. Modified structural features of TM1 and STE can increase the proportion of the atypical forms ^{Ntm}PrP, ^{Ctm}PrP, and ^{cy}PrP, which can be detrimental *in vivo*.[16,19,20] It is worth noting that five out of the seven PMs are located in the STE and TM1 regions ("STE/TM1," hereafter).

On the other hand, N-term_$HuPrP^C$ is a broad-spectrum molecular sensor.[21] Indeed, along with the Mo species, it interacts with chemicals including (i) sulphated glycosaminoglycans (GAG),[22] which bind to the four histidines in the so-called octapeptide-repeat region (OR, see Figure 1A,B and SI, Table S1), as well as several lysines at the N-terminus[23] (Table 1), and (ii) Cu^{2+} ions,[24] which may bind at six histidine residues,[25] including those located in the OR (Table 1). In

addition, the Mo protein has been shown to interact with vitronectin,[26] the stress-inducible protein 1 (STI1)[27] and amyloid-β (Aβ) multimers,[28−30] which may bind N-term_MoPrP^C at residues 104−118, 114−127, and 95−110, respectively, as well as with the lipoprotein receptor-related protein 1 (LRP1)[31] and the neural cell adhesion molecule (NCAM)[32] (Table 1).

On the basis of these premises, here we make the plausible hypothesis that the PMs in the N-term of the protein cause prion diseases by (i) modifying the conformational properties of the STE/TM1 region and its interactions with trans-acting factors in the ER membrane and in the cytosol[33] and/or (ii) by affecting the interactions of N-term_HuPrP^C with chemicals and/or with protein partners *in vivo*. We test this hypothesis by biocomputing methods. We focus on MoPrP^C because of its high sequence identity (particularly in the N-term) to HuPrP^C and because experiments of protein binding have been so far reported mainly for this species. Our calculations predict the conformational ensemble of WT N-term_MoPrP^C and that of N-term_MoPrP^C containing the seven Hu PMs using replica-exchange (RE) Monte Carlo (MC) simulation[34] based on an implicit-solvent all-atom potential.[35] We used the latter previously to investigate the folding of structurally diverse sets of proteins, as well as to investigate the conformational properties of naturally unfolded proteins.[36−38] We did not consider (i) PMs present in MoPrP^C but not in HuPrP^C and artificial mutations,[39,40] their effect in humans being currently unknown; (ii) PMs in the signal peptides, which are cleaved off during the biosynthesis of the protein;[41] or (iii) adducts with metal ions other than Cu^{2+} (e.g., Zn^{2+}).

We find that N-term_MoPrP^C is composed of several regions characterized by different SS, in agreement with biophysical data.[42−46] The PMs do not dramatically change the conformational ensemble of N-term_MoPrP^C. Rather, they affect in a subtle manner the intramolecular contacts in regions responsible for interactions with cellular partners, as well as the local structural features of STE/TM1. These in turn may alter the protein function[21] and biosynthesis.[19]

MATERIALS AND METHODS

Bioinformatics. A full-length MoPrP^C was first constructed by homology modeling. Hereafter, we adopt the Mo numbering to match that of MoPrP^C. The HHpred web server[47] was used to select homologous templates (SI Table S2) from structured proteins in the RCSB Protein Data Bank (www.pdb.org). These templates were provided to MODELER 9.10[48] for multitemplate homology modeling. First, 100 preliminary models were constructed; we then used PROCHECK[49] to select the best model which had over 90% of the backbone dihedral angles within the core region of the Ramachandran plot. The full-length model was truncated at residue 127 to obtain the N-term_MoPrP^C. The Swiss-PdbViewer (DeepView 4.0)[50] was used to introduce the seven mutations (Q52P, P101L, P104L/T/S, G113V, and A116V) into the N-term_MoPrP^C for generating initial models of the seven PMs. These eight models were used as educated guesses for the subsequent molecular simulations and by no means represented a structural prediction.

PONDR-FIT[51] was used to predict the intrinsic disorder disposition of N-term_MoPrP^C. PSIPRED,[52] Jpred3,[53] Porter,[54] RaptorX,[55] and CABS-fold[56] were used to predict the SS in this domain.

Molecular Simulations. REMC simulations[34] using the PROFASI (PROtein Folding and Aggregation Simulator) code[35] were carried out for WT N-term_MoPrP^C and each of the seven PMs. PROFASI uses an all-atom model for the protein including hydrogen atoms and implicit water solvent. The model assumes fixed bond lengths, bond angles, and peptide torsion angles (180°), so that each amino acid has the Ramachandran dihedral angles φ and ψ as well as a number of side-chain torsion angles as its degrees of freedom. The interaction potential is composed of four terms: $E = E_{loc} + E_{ev} + E_{hb} + E_{sc}$. The term E_{loc} is a local potential that accounts for interactions between atoms separated by a few covalent bonds, such as the electrostatic interaction between adjacent peptide units along the chain. The other three terms are nonlocal in sequence. The excluded volume term E_{ev} is a $1/r^{12}$ repulsion between pairs of atoms. E_{hb} represents two kinds of hydrogen bonds: backbone−backbone bonds and bonds between charged side chains and the backbone. The last term E_{sc} represents simple pairwise additive approximations for hydrophobic attraction between nonpolar side chains and electrostatic interactions among charged side-chains such as those forming salt bridges. The interaction potential assumes the presence of a solvent environment for the proteins, although solvent molecules are not explicitly represented in the simulations. To explore the conformation space, we used a Markov Chain MC procedure. At any step, a conformation update involving a change in one or more backbone and side-chain torsion angles is proposed, and is either accepted or rejected using a Metropolis[57] criterion. A semilocal backbone update, called Biased Gaussian Steps,[58] producing smooth local deformations of the chains was used at the lower temperatures to improve the sampling of compact structures.

In each case studied here, the system was enclosed in a periodic cubic box of size $(300\ Å)^3$. We used a replica exchange (also known as parallel tempering) procedure[59] to enhance the sampling in our simulations. Simultaneous MC simulations were performed at eight temperatures: 293.0, 299.7, 306.5, 313.5, 320.6, 327.9, 335.4, and 343.0 K. Periodically, conformations of replicas at consecutive temperatures were swapped with a probability $P^{ij}_{swap} = \exp(1/(k_B T_i) - 1/(k_B T_j)(E_i - E_j))$, where T_k is the temperature and E_k is the energy of replica k and k_B is the Boltzmann constant. The exchange procedure improves sampling at the lower temperatures by stochastically seeding them with independent states obtained at high temperatures, and the above exchange probability ensures that the equilibrium is maintained at each temperature even for the replica exchange step.

The histogram of energy at different temperatures (SI Figure S1) demonstrated a massive overlap between each consecutive pair of temperatures. This showed that there were no sharp transitions in temperature in our system. In order to assess the reliability of the statistics obtained in these simulations, we examined the random walk of the replicas in temperature space (SI Figure S2). At the higher temperatures, the conformation of the system decorrelates fast by jumping over energy barriers; hence the statistics collected at the lower temperatures consist of uncorrelated samples, if the replica has visited a high temperature in between. The length of the simulations was chosen to allow our replicas to reach both ends of the temperature range repeatedly, so that our histograms received data from uncorrelated sample conformations from different parts of the energy landscape. We calculated the cumulative averages[60] of the radius of gyration (R_g) and of the SS contents

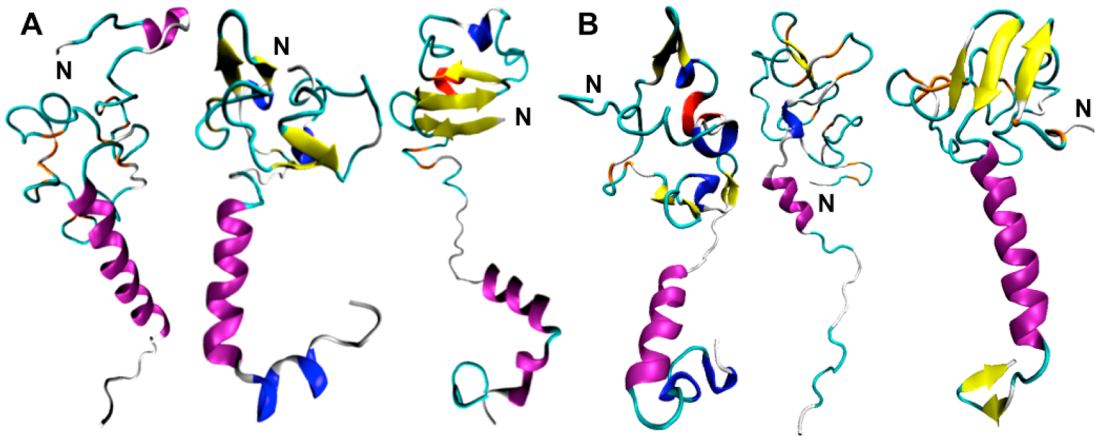


Figure 2. Selected conformations of (A) WT N-term_MoPrP^C and (B) one PM (N-term_MoPrP^C_Q52P) from our simulations. These contain transient α-helix (in violet), β-sheet (yellow), β-bridge (orange), β-turn (cyan), 3_{10}-helix (blue), and π-helix (red) elements.

Table 2. Content of Transient SS Elements (in %[b]) in Each Region of N-term_MoPrP^C Averaged over the Trajectory[a]

residues	α-helix	β-sheet	β-bridge	β-turn	bend	3_{10}-helix	π-helix	coil
23−127	19 ± 8	8 ± 5	7 ± 3	27 ± 5	12 ± 4	4 ± 3	1 ± 1	25 ± 5
23−30	3 ± 2	3 ± 3	6 ± 2	22 ± 5	10 ± 4	2 ± 2	0 ± 1	55 ± 9
31−50	7 ± 4	8 ± 5	12 ± 4	33 ± 6	13 ± 4	5 ± 3	0 ± 1	22 ± 5
51−90 (OR)	6 ± 2	12 ± 5	9 ± 4	36 ± 7	13 ± 5	6 ± 3	1 ± 1	18 ± 4
89−98	5 ± 2	12 ± 5	5 ± 3	23 ± 5	16 ± 4	4 ± 2	0 ± 2	35 ± 7
99−117	64 ± 9	1 ± 2	0 ± 1	7 ± 3	8 ± 3	2 ± 2	0 ± 1	16 ± 6
118−125	33 ± 8	3 ± 6	2 ± 3	33 ± 5	9 ± 3	5 ± 3	1 ± 1	15 ± 4

[a]According to the occurrence of different SS, we identify six sub-regions: (1) residues 23−30, mainly in coil/β-turn/bend conformation; (2) residues 31−50, mainly in β-turn/coil/bend/β-bridge conformations; (3) residues 51−90 (OR) mainly in β-turn/coil/bend/β-sheet conformations; (4) residues 89−98 mainly in coil/β-turn/bend/β-sheet conformations; (5) residues 99−117, which contain the highest percentage of α-helix among all six sub-regions; and (6) residues 118−125, which display a comparable percentage of α-helix and β-turn. [b]Standard deviations (SD) are indicated by " ± ".

over the sampled trajectory. The cumulative average was defined by $\sum_{i=1}^{N} X^i/N$, where N is the total number of frames sampled, i is the frame number, and X^i is the property of interest (the R_g or the SS contents) calculated for each frame. The cumulative averages appeared to converge rather than systematically drift (SI Figure S3). This, combined with the fact that the simulations continued to explore uncorrelated states by moving back and forth across temperatures, gave us reasonable confidence that our simulations have captured the statistical properties of these systems to the accuracy of the force field. In order to further reduce the effect of autocorrelation in our statistics, the conformations were sampled every 10^4 cycles (3.7×10^6 MC steps). We performed 1.8×10^8 cycles (or 6.7×10^{10} MC steps) of simulations for each system, and the first 10% of each run was discarded for thermalization. Production trajectories obtained at 293.0 K were extracted, which yielded 2000 frames per protein for analysis.

Data Analysis. The following properties were calculated with the production trajectory obtained for each N-term_MoPrP^C variant. (i) The average R_g was calculated using g_gyrate in Gromacs 4.5.[61] (ii) The SS in each frame, as well as the average SS content in the trajectory, was calculated with the DSSP software.[62] (iii) The intramolecular contact map for each frame, which was a distance matrix consisting of the smallest distance within 6 Å between residue pairs, was calculated using g_mdmat in Gromacs 4.5. The total number of contacts per residue was also calculated over the entire production trajectory. Every time two residues approached each other within 6-Å distance, this was counted as one contact for each residue. (iv) The backbone flexibility during the trajectory was calculated using the so-called T-PAD analysis.[63]

A clustering analysis was carried out on the contact map data in order to characterize the intramolecular contacts in each trajectory. The contact map of each frame was converted into an array of length $M \times M$ by concatenating matrix rows, where M was the number of residues in the protein. Distances between frames were computed using the Euclidean distance (eq 1) as a metric and stored in a proximity matrix. This distance indicated the similarity between two frames regarding their intramolecular contacts: the shorter the distance, the more similar the two frames. Finally, a hierarchical clustering[64] was performed by using an average linkage method.[65] In this method, the distance between two clusters is the average distance between pairs of observations, one in each cluster. Main clusters were obtained using a distance cutoff on the hierarchical tree so that the 10 most populated clusters cover about 70% of the entire data set. An average contact map in each cluster was computed to represent the intramolecular contacts in the cluster members.

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \ldots (p_n - q_n)^2)} \quad (1)$$

RESULTS

A total of 5.4×10^{11} MC steps of simulation are performed on bionformatics-based structural models of N-term_MoPrP^C and the seven PMs (6.7×10^{10} MC steps for each protein). The

dx.doi.org/10.1021/ct400534k | J. Chem. Theory Comput. 2013, 9, 5158−5167

Table 3. R_g and Content of Transient SS Elements (in %[a]) of the Entire Systems Studied (top) and of Residues 99−127 (bottom)

variant	R_g (Å)	α-helix	β-sheet	β-bridge	bend	β-turn	3_{10}-helix	π-helix	coil
					N-term				
N-term_MoPrP^C	18.5 ± 2.9	19 ± 8	8 ± 5	7 ± 3	12 ± 4	27 ± 5	4 ± 3	1 ± 1	24 ± 5
N-term_MoPrP^C _Q52P	17.1 ± 2.5	17 ± 8	14 ± 7	7 ± 3	12 ± 4	29 ± 5	3 ± 3	1 ± 2	22 ± 6
N-term_MoPrP^C _P101L	19.6 ± 3.0	16 ± 10	12 ± 10	7 ± 3	11 ± 4	29 ± 5	4 ± 3	0 ± 1	24 ± 6
N-term_MoPrP^C _P104L	19.3 ± 2.9	23 ± 8	9 ± 7	7 ± 3	11 ± 4	29 ± 5	4 ± 3	0 ± 1	22 ± 6
N-term_MoPrP^C _P104S	20.1 ± 2.9	21 ± 7	8 ± 5	7 ± 5	11 ± 3	27 ± 5	5 ± 4	0 ± 1	23 ± 5
N-term_MoPrP^C _P104T	19.7 ± 3.0	22 ± 8	8 ± 7	7 ± 3	11 ± 4	28 ± 5	4 ± 4	0 ± 1	24 ± 6
N-term_MoPrP^C _G113V	18.7 ± 2.9	20 ± 7	9 ± 5	7 ± 7	12 ± 3	28 ± 4	4 ± 3	0 ± 1	25 ± 5
N-term_MoPrP^C _A116V	18.8 ± 2.9	17 ± 8	10 ± 7	7 ± 3	11 ± 4	29 ± 5	4 ± 3	0 ± 1	25 ± 5
					residues 99−127				
N-term_MoPrP^C	12.3 ± 1.4	52 ± 10	2 ± 2	1 ± 1	8 ± 5	15 ± 4	3 ± 3	1 ± 1	20 ± 4
N-term_MoPrP^C _Q52P	12.2 ± 1.4	48 ± 10	4 ± 4	2 ± 2	9 ± 4	15 ± 5	3 ± 3	0 ± 1	19 ± 5
N-term_MoPrP^C _P101L	11.9 ± 1.2	41 ± 11	13 ± 4	1 ± 2	6 ± 4	17 ± 5	3 ± 3	0 ± 1	19 ± 4
N-term_MoPrP^C _P104L	12.2 ± 1.4	61 ± 12	3 ± 2	1 ± 1	4 ± 4	14 ± 5	2 ± 3	0 ± 1	14 ± 5
N-term_MoPrP^C _P104S	12.4 ± 1.2	63 ± 13	2 ± 2	1 ± 2	4 ± 4	14 ± 4	2 ± 2	0 ± 1	14 ± 4
N-term_MoPrP^C _P104T	12.3 ± 1.3	60 ± 11	3 ± 2	1 ± 1	14 ± 6	14 ± 4	2 ± 3	0 ± 1	15 ± 4
N-term_MoPrP^C _G113V	11.9 ± 1.2	53 ± 10	2 ± 3	1 ± 2	7 ± 4	14 ± 6	2 ± 2	1 ± 1	19 ± 5
N-term_MoPrP^C _A116V	11.8 ± 1.4	46 ± 10	4 ± 3	1 ± 1	8 ± 5	16 ± 4	3 ± 3	0 ± 1	21 ± 5

[a]SDs are indicated by " ± ".

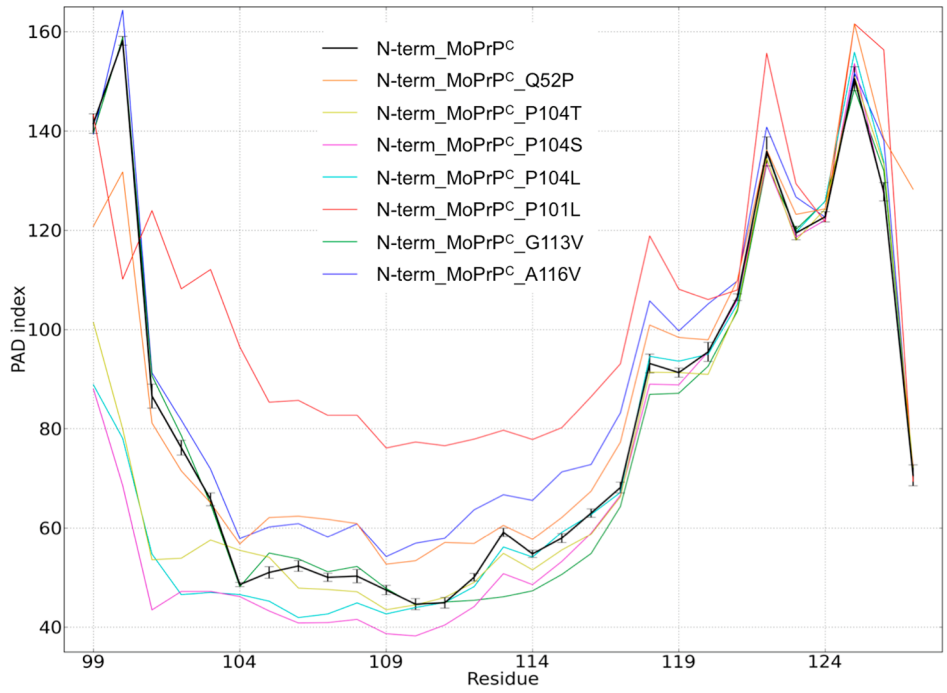


Figure 3. Flexibility of backbone units of residues 99−127. The higher the PAD index value,[63] the more flexible the backbone. The SDs in the 10 simulation runs of the WT N-term_MoPrP^C are indicated with vertical bars, whereas those of the PMs are shown in SI Figure S6.

number of MC steps performed is expected to be sufficient for the naturally unfolded proteins studied here. Only the simulations obtained at 293.0 K are discussed in this work because most of the *in vitro* experiments with which we compare our results have been carried out at room temperature.[42−46]

N-term_MoPrP^C WT. The equilibrated trajectory of N-term_MoPrP^C at 293.0 K provides an ensemble of conformations containing transient SS elements (Figure 2A) with R_g of 18.5 ± 2.9 Å. On average, it contains 19 ± 8% α-helix, 8 ± 5% β-sheet, 7 ± 3% β-bridge, 27 ± 5% β-turn, 12 ± 4% bend, 4 ± 3% 3_{10}-helix, and 1 ± 1% π-helix. We also calculated the average SS contents at 306.5 and 313.5 K since these values are the nearest to the physiological temperature. As it might be expected, the α-helix content decreases and the coil structure increases with respect to the results of 293.0 K (SI Table S3). However, the differences between the two higher temperatures and 293.0 K are in the range of the standard deviation along the simulations.

At 293.0 K, different regions in N-term_MoPrP^C appear to adopt different SS. According to the occurrence of diverse SS, N-term_MoPrP^C can be roughly divided into six subregions (Table 2). Notably, at residues 99−117 and 118−125, we observe a relatively high occurrence of α-helix (64 ± 9% and 33

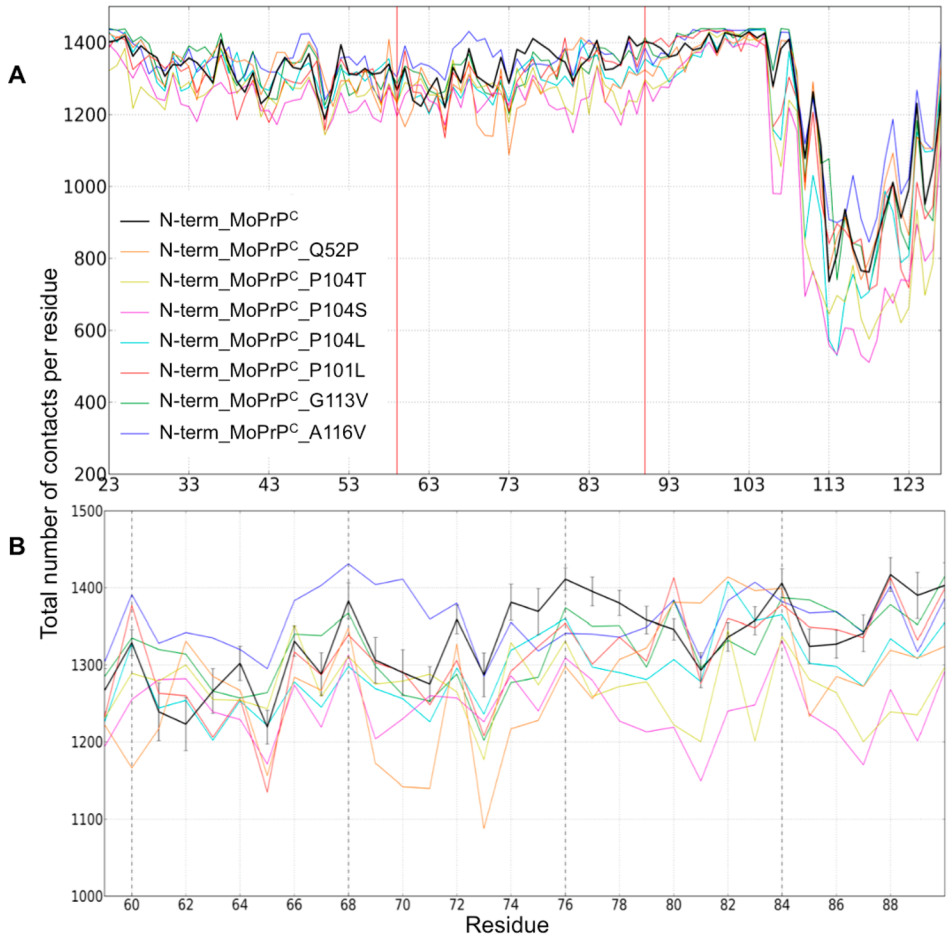


Figure 4. Total number of intramolecular contacts per residue in the WT and mutated N-term_MoPrP^C. (A) Total number of contacts between residues in the entire N-term over the trajectory. The SDs of the 10 simulation runs of each system are shown in SI Figure S7. (B) Total number of contacts in the OR over the trajectory where the four histidine residues are indicated with vertical dashed lines. The SDs in the 10 simulation runs of the WT N-term_MoPrP^C are indicated with vertical bars, whereas those of the PMs are in a similar range to that of the WT (SI Figure S7).

± 8%, respectively), which is in line with the prediction by PONDR-FIT[51] that this region has relatively low intrinsic disorder disposition (SI Figure S4). Consistently, commonly used bioinformatics tools for SS prediction (such as PSIPRED,[52] Jpred3,[53] Porter,[54] RaptorX,[55] and CABS-fold[56]) also indicate that residues ∼104−120 have the propensity to form an α-helix (SI Figure S4).

Our results are consistent with the finding by solution NMR that the OR peptide comprising residues 61−68 in N-term_HuPrP^C acquires a stable loop/β-turn conformation at pH ∼6.[46] In our simulations, indeed we find the loop/β-turn conformations in the OR region resembling the NMR structure.[46] These conformations show a backbone RMSD of less than 2.5 Å from the NMR structure,[46] which composes 8%, 11%, 3%, 9%, and 13% of the five sequential OR peptides in the OR region, respectively. We also find antiparallel β-sheets/β-bridges formed between adjacent OR peptides.

In residues 105−125 (known as the amyloidogenic region), we observe a transient helical structure spanning all over this region. However, the first part (residues 105−117) shows mainly α-helix motifs, while residues 118−125 display a comparable percentage of α-helix and β-turn. These results are consistent with the reported CD, NMR, and FTIR studies on HuPrP^C fragments.[43−45] Together, these studies indicate that the equivalent region in HuPrP^C may adopt α-helix, β-sheet, or coil structures under different experimental conditions.[43−45] We notice that this region roughly corresponds to the STE/TM1 in N-term_MoPrP^C, which comprises residues 103−127.

N-term_MoPrP^C PMs. The R_g of the seven PMs studied here is comparable to that of the WT (Table 3). This indicates that the mutations do not alter the compactness of the N-term_MoPrP^C. The average SS contents in the PMs are also mostly similar to those of the WT (Table 3, Figure 2B).

Differences between the WT and PMs are mainly in the β-sheet and α-helix contents in residues 99−127 (Table 3). Among these, N-term_MoPrP^C_P101L has 11% more β-sheet content and 11% less α-helix content in this region than those in the WT. This result is consistent with the *in vitro* data of MoPrP^C containing the P101L mutation.[66] N-term_MoPrP^C_P104L/T/S mutations lead to an increase in α-helix content by 9%, 11%, and 8%, respectively. The PMs feature different backbone conformational flexibility (calculated here with the T-PAD analysis,[63] SI Figure S5) from that of N-term_MoPrP^C WT. In particular, the flexibility of STE/TM1 increases in N-term_MoPrP^C_Q52P, N-term_MoPrP^C_P101L, and N-term_MoPrP^C_A116V. It decreases in N-term_MoPrP^C_G113V and N-term_MoPrP^C_P104L/T/S with respect to the WT (Figure 3).

dx.doi.org/10.1021/ct400534k | J. Chem. Theory Comput. 2013, 9, 5158−5167

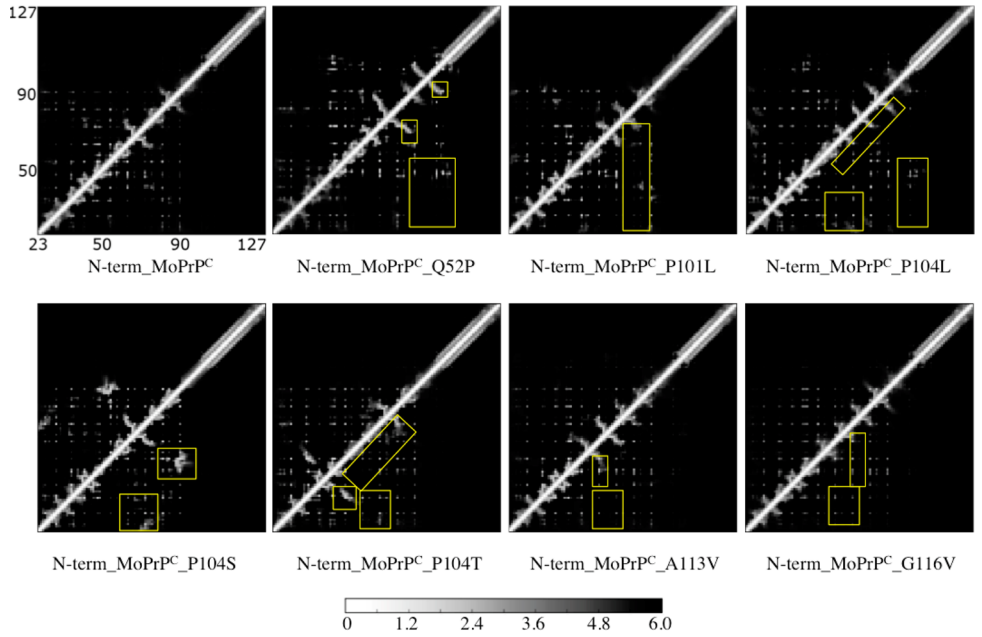


Figure 5. Contact maps illustrating the intramolecular interactions in WT and mutated N-term_MoPrP^C. Contacts within 6 Å are plotted in gray scale according to the distances as shown at the bottom. Residue numbers are labeled for the N-term_MoPrP^C and are the same for all the PMs. For each protein, only the most populated cluster is shown (see SI Figure S8 and Table S4 for the other clusters and the detailed data of each cluster, respectively). Major differences between the WT and the PMs are highlighted in the lower half of the (diagonally symmetric) maps with yellow boxes.

DISCUSSION

Our calculations suggest that N-term_MoPrP^C WT features a large ensemble of conformations containing transient SS, rather than completely disordered ones. These results are consistent with several sets of experimental data.[42−46]

While many PMs in the GD are known to modify significantly the folded structure and to increase its flexibility,[2,5,7,8] our calculations suggest that those in the N-term do not impact significantly the global structural properties of the N-term. This is consistent with *in vitro* experiments, which have shown that PMs in N-term_HuPrP^C do not affect the thermostability or the misfolding kinetics of the protein molecule.[2,10−12] However, the PMs at the N-term do modify local features from those of the WT at the binding sites for known cellular partners. This might affect the interactions with these partners and interfere with the related physiological functions.

Specifically, (i) the intramolecular contacts of residues binding Cu^{2+} and sulphated GAG (Table 1) decrease on passing from the WT to the PMs, as shown by the total number of contacts per residue during the simulations of each system (Figure 4). Ten main clusters of contact maps characterize the intramolecular contacts during the simulations (SI Figure S8). These are obtained by cluster analysis on the intramolecular contact maps of each system. In the most populated cluster (Figure 5), the N-term_MoPrP^C WT and the PMs differ mostly for the residues binding Cu^{2+} and sulphated GAG (i.e., the OR region and the H110 Cu^{2+}-binding site (Table 1)). In the less populated clusters of contact maps (SI Figure S8), minor differences between N-term_MoPrP^C WT and the PMs are observed also in other regions. These are not statistically significant; thus they are not discussed here. Alteration of the Cu^{2+}-PrP^C binding process has been suggested to play a role for Hu/Mo PrP^C aggregation and for the progression of the disease.[24] Interestingly, a recent study in our lab has shown that the GSS-linked Q211P PM[5] (Q212P in Hu numbering) in the HuPrP^C GD can influence the Cu^{2+} binding coordination at H95 and H110,[67] implicating a role of abnormal Cu^{2+} binding in the pathology of PMs in HuPrP^C. Sulphated GAG binding to PrP^C is likely involved in the formation of PrP^{Sc} *in vivo*, although its role remains to be established.[22] In addition, (ii) the PMs affect the SS (Table 3) and the flexibility (Figure 3) and increase the hydrophobicity of STE/TM1 (SI Table S5). The latter contains the putative binding sites for *in vivo* binding partner proteins such as vitronectin[26] and STI1.[27] This might affect the biological function of these interactions, which involves the signaling for axonal growth[26] and that for neuroprotection,[27] respectively.

The PM Q52P in the OR region affects the flexibility of STE/TM1 (Figure 3), while the other six PMs in STE/TM1 also alter the intramolecular contacts in the OR region (Figures 4 and 5). This suggests that the PMs may affect the transient interactions between the OR region and STE/TM1. However, these observed alterations on passing from the WT to the PMs in (i) the intramolecular contacts and (ii) the backbone flexibility turn out not to be quantitatively correlated. The number of contacts takes into account the backbone and side-chain atoms and is related to the 3D fold of the protein conformations sampled in the simulations. The PAD index, instead, measures the fluctuations of the backbone dihedral angles of the residues during the simulations.[63] Therefore, the local backbone flexibility might not be directly related to the 3D fold of the protein.

Growing evidence point to the role of PrP^C in mediating the memory impairment by Aβ multimers *in vivo*.[28,29] On the one hand, Aβ multimers have been found to bind PrP^C N-term, which is required for the suppression of long-term potentiation by Aβ multimers *in vivo*.[28,30] On the other hand, *in vitro* experiments suggest a Cu^{2+}-dependent mechanism that links the neurotoxicity of Aβ multimers to PrP^C.[68] The fact that

MoPrP^C residues 95−110 in the N-term are essential for the binding of both Aβ multimers[28] and Cu^{2+} suggests the significance of this region in the yet undefined role of PrP^C in Alzheimer's disease. Our findings that PMs affect the structural features of this region are in line with the hypothesis that a Cu^{2+}-dependent mechanism may unify the pathogenesis of neurodegenerative diseases such as TSEs, Alzheimer's, and Parkinson's disease.[69]

The altered local features in STE/TM1 might also impact the interactions of the protein with trans-acting factors in the cytosol and in the ER membrane.[33] This result is consistent with the *in vitro* data that PMs P101L, P104L, and A116V increase the interactions between MoPrP^C STE/TM1 and a membrane mimetic at pH 7.[70] These changes may affect the relative proportions of the ^{Sec}PrP, ^{Ntm}PrP, ^{Ctm}PrP, and ^{Cy}PrP topological variants in humans and mice.[19] This in turn could have repercussions on the disease, as an elevated ^{Ctm}PrP level has been related to both familial (in GSS patients) and infectious (in Tg mice) prion diseases.[16,20]

Implications for PrP^C Molecular Recognition. Experimental structural information is available for several fragments of PrP^C N-term in complex with antibodies. These include residues 119−122 from HuPrP^C (equivalent to MoPrP^C 118−121) in β-strand,[71] residues 104−113 from SHaPrP^C (equivalent to MoPrP^C 103−112) in bend conformation,[72] residues 107−115 from bovine (Bo) PrP^C (equivalent to MoPrP^C 95−103) in bend conformation,[73] and a MoPrP^C 2-OR peptide (residues 67−82) in an extended conformation.[74] SHaPrP^C and BoPrP^C share 86% and 95% of sequence identity with MoPrP^C, respectively. In addition, heparin binds to N-term from SHaPrP^C and stabilizes a repeated loop/β-turn conformation encompassing 4 OR.[23] Interestingly, the former three are present in our calculated ensemble. This indicates a possible "conformational selection" mechanism[75] upon binding in these cases. The extended conformation of the MoPrP^C 2-OR peptide is embedded in a groove of its cognate antibody (named POM2 Fab)[74] which prevents the formation of the loop/β-turn seen by NMR[46] and by our simulation. Similarly, the 4-OR conformation in N-term from SHaPrP^C bound with heparin is absent in our simulation, although, as discussed in the Results section, the conformation of each single OR is similar to that of the NMR structure,[46] and it is present in our simulation. Hence, we suggest that the binding of the POM2 Fab antibody and heparin to the OR involves an "induced fit" mechanism,[75] since the resulting OR conformations are not formed before binding.

In our test simulations on full-length WT MoPrP^C (discussed in the next paragraph), the above-mentioned β-strand structure in residues 118−121 (denoted β_0, hereinafter) extends the two-strand β-sheet in the GD into a three-strand antiparallel β-sheet. Such extension has been very recently reported in the crystal structure of HuPrP^C bound with an antibody[71] and in the MD simulations on truncated SHaPrP^C at acidic pH[76] or upon the G131V or M129V mutations in truncated SHaPrP^C.[77] Our simulations on full-length WT MoPrP^C and N-term_MoPrP^C show that β_0 can be formed with or without the GD (Table 2 and SI Table S6), suggesting the propensity of residues 118−125 to form a transient β-strand. The occurrence of β_0 along with other β-sheet structures within residues 99−127 is higher in the P101L mutant than that in the other mutants and the WT (Table 3). This appears to be accompanied by a decrease of α-helix in this mutant (Table 3), consistent with the *in vitro* data of full-length MoPrP^C_P101L.[66]

Limitations. As with any modeling study, this work has limitations. First, we do not include the GD in the calculations. Test calculations on full-length WT MoPrP^C (with the GD constrained) show that its N-term conformational ensemble exhibits features very similar to those of the N-term_MoPrP^C (see SI Table S6). On the one hand, this suggests that the absence of GD does not notably influence the conformational ensemble of the N-term. On the other hand, simulating the full-length MoPrP^C (containing 208 residues) for all of the systems investigated here might face issues of convergence and limitations of the accuracy of the force field used. The latter is calibrated for systems much smaller than the full-length MoPrP^C.[35,78,79] Constraining the GD introduces artifacts regarding the interactions between the GD and the N-term. NMR studies have reported interactions between the two domains in the presence or absence of metal ions (such as Cu^{2+} or Zn^{2+}).[80−83] In our simulations, however, we do not observe statistically significant interactions between the two domains. Since in our simulations these metal ions are absent and the GD is constrained, a proper comparison between our results and these NMR data concerning the interactions between the GD and the N-term cannot be made. Second, without explicit water molecules, some potentially important questions cannot be addressed, such as the effect of ionic strength. Here, however, we stress that the full agreement between available experimental data and our findings does validate the computational approach, which has been successfully applied to a variety of intrinsically disordered proteins.[37,38] Third, our REMC simulations cannot provide dynamics properties, which are the disadvantages compared with RE molecular dynamics (MD) techniques (such as that used in ref 84). However, our REMC technique has a remarkable advantage in sampling efficiency, which allows studying processes that, because of the long time scales involved, would be currently infeasible with REMD simulations.

CONCLUSION

The seven N-terminal PMs do not dramatically impact the conformational ensemble. However, they do alter the local conformational features in STE/TM1, which are in control of PrP topology during its biosynthesis.[16,19] This region also contains the estimated binding sites for cellular partners such as vitronectin and STI1.[26,27] Furthermore, the PMs specifically affect the intramolecular contacts in the OR at the putative binding sites for Cu^{2+} [25] and sulphated GAG,[23] as well as in the Cu^{2+}-binding site around residue H110[85] (Table 1). Hence, unlike most of the PMs in the GD, which affect HuPrP^C thermostability and misfolding kinetics,[2] the PMs at the N-term induce local structural modifications in key regions for HuPrP^C biosynthesis,[16,19] for metal binding,[25] and for interactions with other cellular partners.[21]

ASSOCIATED CONTENT

Supporting Information

Plots of energy histogram at different temperatures, a random walk of a replica (in the temperature space), cumulative averages, secondary structure prediction, backbone flexibility, and number of intramolecular contacts per residue during the simulations. Secondary structure contents in the simulations of N-term_MoPrP^C at 306.5 and 313.5 K. Details of contact maps and clustering based on contact maps. Four different

hydrophobicity scales for amino acid residues in proteins. Results of simulation on full-length MoPrP^{C}. Protein structure templates used for homology modeling full-length MoPrP^{C}. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: g.rossetti@grs-sim.de.
*E-mail: p.carloni@grs-sim.de.

Author Contributions

$^{\bigcirc}$These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge the Jülich-Aachen Research Alliance (JARA) and the Jülich Supercomputing Center (JSC) for computing resources. The authors wish to thank Erica Sarnataro for editing of the manuscript. This work was supported by a grant from the Italian *Ministero dell'Istruzione dell'Università e della Ricerca (MIUR)* through a FIRB-Accordi di Programma 2011 entitled *Genomica Funzionale delle Malattie Neurodegenerative* (Prot. RBAP11FRE9_001) to G.L.

ABBREVIATIONS

PM, pathogenic mutation; WT, wild-type; GAG, glycosaminoglycans; STI1, the stress-inducible protein 1; Aβ, amyloid-β; LRP1, the lipoprotein receptor-related protein 1; NCAM, the neural cell adhesion molecule; TM1, transmembrane domain; STE, stop transfer effector; Hu, human; Mo, mouse; SHa, Syrian hamster; Bo, bovine; fCJD, familial Creutzfeldt-Jakob disease; GSS, Gerstmann-Sträussler-Scheinker syndrome; PrP^{C}, the cellular form of prion protein; PrP^{Sc}, the scrapie form of prion protein; SecPrP, secretory PrP; ^{Ntm}PrP/^{Ctm}PrP, the transmembrane forms of PrP; ^{Cy}PrP, cytosolic PrP; N-term, N-terminal domain; GD, globular domain; SS, secondary structure; RE, replica-exchange; MC, Monte Carlo; R_g, radius of gyration; SD, standard deviation

REFERENCES

(1) Prusiner, S. B. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 13363.
(2) van der Kamp, M. W.; Daggett, V. *Protein Eng., Des. Sel.* **2009**, *22*, 461.
(3) Schelzke, G.; Stoeck, K.; Eigenbrod, S.; Grasbon-Frodl, E.; Raddatz, L. M.; Ponto, C.; Kretzschmar, H. A.; Zerr, I. *Dementia Geriatr. Cognit. Disord.* **2013**, *35*, 229.
(4) Giachin, G.; Biljan, I.; Ilc, G.; Plavec, J.; Legname, G. *Molecules* **2013**, *18*, 9451.
(5) Ilc, G.; Giachin, G.; Jaremko, M.; Jaremko, L.; Benetti, F.; Plavec, J.; Zhukov, I.; Legname, G. *PLoS One* **2010**, *5*, e11715.
(6) Biljan, I.; Ilc, G.; Giachin, G.; Plavec, J.; Legname, G. *Biochemistry* **2012**, *51*, 7465.
(7) Rossetti, G.; Cong, X. J.; Caliandro, R.; Legname, G.; Carloni, P. *J. Mol. Biol.* **2011**, *411*, 700.
(8) van der Kamp, M. W.; Daggett, V. *J. Mol. Biol.* **2010**, *404*, 732.
(9) Meli, M.; Gasset, M.; Colombo, G. *PLoS One* **2011**, *6*, No. e19093.
(10) Apetri, A. C.; Surewicz, K.; Surewicz, W. K. *J. Biol. Chem.* **2004**, *279*, 18008.
(11) Liemann, S.; Glockshuber, R. *Biochemistry* **1999**, *38*, 3258.
(12) Swietnicki, W.; Petersen, R. B.; Gambetti, P.; Surewicz, W. K. *J. Biol. Chem.* **1998**, *273*, 31048.
(13) Liu, Z.; Jia, L.; Piao, Y.; Lu, D.; Wang, F.; Lv, H.; Lu, Y.; Jia, J. *Acta Neurol. Scand.* **2010**, *121*, 377.
(14) Barron, R. M.; Thomson, V.; King, D.; Shaw, J.; Melton, D. W.; Manson, J. C. *J. Gen. Virol.* **2003**, *84*, 3165.
(15) Barron, R. M.; Thomson, V.; Jamieson, E.; Melton, D. W.; Ironside, J.; Will, R.; Manson, J. C. *EMBO J.* **2001**, *20*, 5070.
(16) Hegde, R. S.; Mastrianni, J. A.; Scott, M. R.; DeFea, K. A.; Tremblay, P.; Torchia, M.; DeArmond, S. J.; Prusiner, S. B.; Lingappa, V. R. *Science* **1998**, *279*, 827.
(17) Li, A. M.; Christensen, H. M.; Stewart, L. R.; Roth, K. A.; Chiesa, R.; Harris, D. A. *EMBO J.* **2007**, *26*, 548.
(18) Ott, C. M.; Akhavan, A.; Lingappa, V. R. *J. Biol. Chem.* **2007**, *282*, 11163.
(19) Chakrabarti, O.; Ashok, A.; Hegde, R. S. *Trends Biochem. Sci.* **2009**, *34*, 287.
(20) Hegde, R. S.; Tremblay, P.; Groth, D.; DeArmond, S. J.; Prusiner, S. B.; Lingappa, V. R. *Nature* **1999**, *402*, 822.
(21) Beland, M.; Roucou, X. *J. Neurochem.* **2012**, *120*, 853.
(22) Silva, J. L.; Vieira, T. C. R. G.; Gomes, M. P. B.; Rangel, L. P.; Scapin, S. M. N.; Cordeiro, Y. *Methods* **2011**, *53*, 306.
(23) Taubner, L. M.; Bienkiewicz, E. A.; Copie, V.; Caughey, B. *J. Mol. Biol.* **2010**, *395*, 475.
(24) Leal, S. S.; Botelho, H. M.; Gomes, C. M. *Coord. Chem. Rev.* **2012**, *256*, 2253.
(25) Migliorini, C.; Porciatti, E.; Luczkowski, M.; Valensin, D. *Coord. Chem. Rev.* **2012**, *256*, 352.
(26) Hajj, G. N. M.; Lopes, M. H.; Mercadante, A. F.; Veiga, S. S.; da Silveira, R. B.; Santos, T. G.; Ribeiro, K. C. B.; Juliano, M. A.; Jacchieri, S. G.; Zanata, S. M.; Martins, V. R. *J. Cell Sci.* **2007**, *120*, 1915.
(27) Zanata, S. M.; Lopes, M. H.; Mercadante, A. F.; Hajj, G. N. M.; Chiarini, L. B.; Nomizo, R.; Freitas, A. R. O.; Cabral, A. L. B.; Lee, K. S.; Juliano, M. A.; de Oliveira, E.; Jachieri, S. G.; Burlingame, A.; Huang, L.; Linden, R.; Brentani, R. R.; Martins, V. R. *EMBO J.* **2002**, *21*, 3307.
(28) Lauren, J.; Gimbel, D. A.; Nygaard, H. B.; Gilbert, J. W.; Strittmatter, S. M. *Nature* **2009**, *457*, 1128.
(29) Nicoll, A. J.; Panico, S.; Freir, D. B.; Wright, D.; Terry, C.; Risse, E.; Herron, C. E.; O'Malley, T.; Wadsworth, J. D.; Farrow, M. A.; Walsh, D. M.; Saibil, H. R.; Collinge, J. *Nat. Commun.* **2013**, *4*, 2416.
(30) Chen, S. G.; Yadav, S. P.; Surewicz, W. K. *J. Biol. Chem.* **2010**, *285*, 26377.
(31) Parkyn, C. J.; Vermeulen, E. G.; Mootoosamy, R. C.; Sunyach, C.; Jacobsen, C.; Oxvig, C.; Moestrup, S.; Liu, Q.; Bu, G.; Jen, A.; Morris, R. J. *J. Cell Sci.* **2008**, *121*, 773.
(32) Schmitt-Ulms, G.; Legname, G.; Baldwin, M. A.; Ball, H. L.; Bradon, N.; Bosque, P. J.; Crossin, K. L.; Edelman, G. M.; DeArmond, S. J.; Cohen, F. E.; Prusiner, S. B. *J. Mol. Biol.* **2001**, *314*, 1209.
(33) Hegde, R. S.; Kang, S. W. *J. Cell Biol.* **2008**, *182*, 225.
(34) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsovvelyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776.
(35) Irback, A.; Mohanty, S. *J. Comput. Chem.* **2006**, *27*, 1548.
(36) Irback, A.; Mitternacht, S.; Mohanty, S. *PMC Biophys.* **2009**, *2*, 2.
(37) Li, D. W.; Mohanty, S.; Irback, A.; Huo, S. *PLoS Comput. Biol.* **2008**, *4*, e1000238.
(38) Jonsson, S. A.; Mohanty, S.; Irback, A. *Proteins* **2012**, *80*, 2169.
(39) Hara, H.; Okemoto-Nakamura, Y.; Shinkai-Ouchi, F.; Hanada, K.; Yamakawa, Y.; Hagiwara, K. *J. Virol.* **2012**, *86*, 5626.
(40) Abalos, G. C.; Cruite, J. T.; Bellon, A.; Hemmers, S.; Akagi, J.; Mastrianni, J. A.; Williamson, R. A.; Solforosi, L. *J. Biol. Chem.* **2008**, *283*, 34021.
(41) Zhao, H. X.; Klingeborn, M.; Simonsson, M.; Linne, T. *Virus Res.* **2006**, *115*, 43.
(42) Calzolai, L.; Zahn, R. *J. Biol. Chem.* **2003**, *278*, 35592.
(43) Degioia, L.; Selvaggini, C.; Ghibaudi, E.; Diomede, L.; Bugiani, O.; Forloni, G.; Tagliavini, F.; Salmona, M. *J. Biol. Chem.* **1994**, *269*, 7859.
(44) Miura, T.; Yoda, M.; Takaku, N.; Hirose, T.; Takeuchi, H. *Biochemistry* **2007**, *46*, 11589.

(45) Satheeshkumar, K. S.; Jayakumar, R. *Biophys. J.* **2003**, *85*, 473.
(46) Zahn, R. *J. Mol. Biol.* **2003**, *334*, 477.
(47) Soding, J.; Biegert, A.; Lupas, A. N. *Nucleic Acids Res.* **2005**, *33*, W244.
(48) Eswar, N.; Webb, B.; Marti-Renom, M. A.; Madhusudhan, M. S.; Eramian, D.; Shen, M. Y.; Pieper, U.; Sali, A. *Current Protocols in Bioinformatics*; Wiley: New York, 2006; Chapter 5, Unit 5.6.
(49) Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283.
(50) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714.
(51) Xue, B.; Dunbrack, R. L.; Williams, R. W.; Dunker, A. K.; Uversky, V. N. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 996.
(52) McGuffin, L. J.; Bryson, K.; Jones, D. T. *Bioinformatics* **2000**, *16*, 404.
(53) Cole, C.; Barber, J. D.; Barton, G. J. *Nucleic Acids Res.* **2008**, *36*, W197.
(54) Pollastri, G.; McLysaght, A. *Bioinformatics* **2005**, *21*, 1719.
(55) Kallberg, M.; Wang, H. P.; Wang, S.; Peng, J.; Wang, Z. Y.; Lu, H.; Xu, J. B. *Nat. Protoc.* **2012**, *7*, 1511.
(56) Blaszczyk, M.; Jamroz, M.; Kmiecik, S.; Kolinski, A. *Nucleic Acids Res.* **2013**, *41*, W406.
(57) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.
(58) Favrin, G.; Irback, A.; Sjunnesson, F. *J. Chem. Phys.* **2001**, *114*, 8154.
(59) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604.
(60) Johansen, A. M.; Evers, L. *Lecture Notes* **2008**, 5.
(61) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435.
(62) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.
(63) Caliandro, R.; Rossetti, G.; Carloni, P. *J. Chem. Theory Comput.* **2012**, *8*, 4775.
(64) Jain, A. K.; Dubes, R. C. *Algorithms for Clustering Data*; Prentice-Hall, Inc.: New York, 1988.
(65) Sneath, P. H.; Sokal, R. R. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*; W H Freeman & Co: New York, 1973.
(66) Cappai, R.; Stewart, L.; Jobling, M. F.; Thyer, J. M.; White, A. R.; Beyreuther, K.; Collins, S. J.; Masters, C. L.; Barrow, C. J. *Biochemistry* **1999**, *38*, 3280.
(67) D'angelo, P.; Della Longa, S.; Arcovito, A.; Mancini, G.; Zitolo, A.; Chillemi, G.; Giachin, G.; Legname, G.; Benetti, F. *Biochemistry* **2012**, *51*, 6068.
(68) You, H. T.; Tsutsui, S.; Hameed, S.; Kannanayakal, T. J.; Chen, L. N.; Xia, P.; Engbers, J. D. T.; Lipton, S. A.; Stys, P. K.; Zamponi, G. W. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1737.
(69) Viles, J. H. *Coord. Chem. Rev.* **2012**, *256*, 2271.
(70) Hornemann, S.; von Schroetter, C.; Damberger, F. F.; Wuthrich, K. *J. Biol. Chem.* **2009**, *284*, 22713.
(71) Abskharon, R. N. N.; Giachin, G.; Wohlkonig, A.; Soror, S. H.; Pardon, E.; Legname, G.; Steyaert, J. *J. Am. Chem. Soc.* **2013**, in press.
(72) Kanyo, Z. F.; Pan, K. M.; Williamson, R. A.; Burton, D. R.; Prusiner, S. B.; Fletterick, R. J.; Cohen, F. E. *J. Mol. Biol.* **1999**, *293*, 855.
(73) Luginbuhl, B.; Kanyo, Z.; Jones, R. M.; Fletterick, R. J.; Prusiner, S. B.; Cohen, F. E.; Williamson, R. A.; Burton, D. R.; Pluckthun, A. *J. Mol. Biol.* **2006**, *363*, 75.
(74) Swayampakula, M.; Baral, P. K.; Aguzzi, A.; Kav, N. N.; James, M. N. *Protein Sci.* **2013**, *22*, 893.
(75) Csermely, P.; Palotai, R.; Nussinov, R. *Trends Biochem. Sci.* **2010**, *35*, 539.
(76) Alonso, D. O. V.; DeArmond, S. J.; Cohen, F. E.; Daggett, V. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 2985.
(77) Santini, S.; Claude, J. B.; Audic, S.; Derreumaux, P. *Proteins* **2003**, *51*, 258.
(78) Favrin, G.; Irback, A.; Mohanty, S. *Biophys. J.* **2004**, *87*, 3657.
(79) Irback, A.; Mohanty, S. *Biophys. J.* **2005**, *88*, 1560.
(80) Riek, R.; Hornemann, S.; Wider, G.; Glockshuber, R.; Wuthrich, K. *FEBS Lett.* **1997**, *413*, 282.
(81) Zahn, R.; Liu, A. Z.; Luhrs, T.; Riek, R.; von Schroetter, C.; Garcia, F. L.; Billeter, M.; Calzolai, L.; Wider, G.; Wuthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 145.
(82) Thakur, A. K.; Srivastava, A. K.; Srinivas, V.; Chary, K. V. R.; Rao, C. M. *J. Biol. Chem.* **2011**, *286*, 38533.
(83) Spevacek, A. R.; Evans, E. G. B.; Miller, J. L.; Meyer, H. C.; Pelton, J. G.; Millhauser, G. L. *Structure* **2013**, *21*, 236.
(84) Baillod, P.; Garrec, J.; Colombo, M. C.; Tavernelli, I.; Rothlisberger, U. *Biochemistry* **2012**, *51*, 9891.
(85) Klewpatinond, M.; Davies, P.; Bowen, S.; Brown, D. R.; Viles, J. H. *J. Biol. Chem.* **2008**, *283*, 1870.
(86) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403.
(87) Wong, C.; Xiong, L. W.; Horiuchi, M.; Raymond, L.; Wehrly, K.; Chesebro, B.; Caughey, B. *EMBO J.* **2001**, *20*, 377.
(88) Santuccione, A.; Sytnyk, V.; Leshchyns'ka, I.; Schachner, M. *J. Cell Biol.* **2005**, *169*, 341.
(89) Cereghetti, G. M.; Schweiger, A.; Glockshuber, R.; Van Doorslaer, S. *Biophys. J.* **2001**, *81*, 516.
(90) Kramer, M. L.; Kratzin, H. D.; Schmidt, B.; Romer, A.; Windl, O.; Liemann, S.; Hornemann, S.; Kretzschmar, H. *J. Biol. Chem.* **2001**, *276*, 16711.