

## FULL ATOMISTIC MODEL OF PRION STRUCTURE AND CONVERSION

Giovanni Spagnoli<sup># [1]</sup>, Marta Rigoli<sup>[1,2]</sup>, Simone Orioli<sup>[2,3]</sup>, Alejandro M. Sevillano<sup>[4]</sup>, Pietro Faccioli<sup>[2,3]</sup>,  
Holger Wille<sup>[5]</sup>, Emiliano Biasini<sup># [1]</sup> & Jesús R. Requena<sup># [6]</sup>

[1] Centre for Integrative Biology (CIBIO) – University of Trento, Via Sommarive, 9, 38123 Povo TN, ITALY

[2] Department of Physics, Via Sommarive, 14, 38123 Povo, Trento TN, ITALY

[3] INFN-TIFPA, Via Sommarive 14, Povo (Trento) I-38123, ITALY

[4] Department of Pathology – University of California, 94115 San Diego, USA.

[5] Department of Biochemistry and Centre for Prions and Protein Folding Diseases – University of Alberta,  
T6G 2R3 Edmonton (AB), CANADA.

[6] CIMUS Biomedical Research Institute & Department of Medical Sciences, University of Santiago de  
Compostela-IDIS, 15782 SPAIN.

<sup>#</sup>Contact: [giovanni.spagnoli@unitn.it](mailto:giovanni.spagnoli@unitn.it); [emiliano.biasini@unitn.it](mailto:emiliano.biasini@unitn.it); [jesus.requena@usc.es](mailto:jesus.requena@usc.es)

## Abstract

Prions are unusual protein assemblies that propagate their conformationally-encoded information in absence of nucleic acids. The first prion identified, the scrapie isoform (PrP<sup>Sc</sup>) of the cellular prion protein (PrP<sup>C</sup>), is the only one known to cause epidemic and epizootic episodes(1). Most aggregates of other misfolding-prone proteins are amyloids, often arranged in a Parallel-In-Register- $\beta$ -Sheet (PIRIBS)(2) or  $\beta$ -solenoid conformations(3). Similar folding models have also been proposed for PrP<sup>Sc</sup>, although none of these have been confirmed experimentally. Recent cryo-electron microscopy (cryo-EM) and X-ray fiber-diffraction studies provided evidence that PrP<sup>Sc</sup> is structured as a 4-rung  $\beta$ -solenoid (4R $\beta$ S)(4, 5). Here, we combined different experimental data and computational techniques to build the first physically-plausible, atomic resolution model of mouse PrP<sup>Sc</sup>, based on the 4R $\beta$ S architecture. The stability of this new PrP<sup>Sc</sup> model, as assessed by Molecular Dynamics (MD) simulations, was found to be comparable to that of the prion forming domain of Het-s, a naturally-occurring  $\beta$ -solenoid. Importantly, the 4R $\beta$ S arrangement allowed the first simulation of the sequence of events underlying PrP<sup>C</sup> conversion into PrP<sup>Sc</sup>. Our results provide the most updated, experimentally-driven and physically-coherent model of PrP<sup>Sc</sup>, together with an unprecedented reconstruction of the mechanism underlying the self-catalytic propagation of prions.

**Keywords:** Prion Protein | PrP<sup>C</sup> | PrP<sup>Sc</sup> | Amyloid | Protein Misfolding | Molecular Dynamics

## Significance

Since the original hypothesis by Stanley Prusiner, prions have represented enigmatic agents diverging from the classical concept of genetic inheritance. However, the structure of PrP<sup>Sc</sup>, the infectious isoform of the cellular prion protein (PrP<sup>C</sup>), has so far remained elusive, mostly due to technical challenges posed by its aggregation propensity. Here, we present a new high resolution model of PrP<sup>Sc</sup> derived from the integration of a wide array of recent experimental constraints. By coupling the information of such model with a newly developed computational method, we reconstructed for the first time the conformational transition of PrP<sup>C</sup> to PrP<sup>Sc</sup>. This study offers a unique workbench for designing therapeutics against prion diseases, and a physically-plausible mechanism explaining how protein conformation could self-propagate.

## Introduction

Prion diseases are infectious neurodegenerative disorders characterized by an invariably lethal outcome caused by a proteinaceous infectious agent named “prion”(1). The central event in these pathologies is the conversion of PrP<sup>C</sup>, a GPI-anchored protein of unknown function, into a misfolded isoform (PrP<sup>Sc</sup>) which accumulates in the central nervous system of affected individuals(6). While PrP<sup>C</sup> structure has been widely characterized, and consists of a N-terminal disordered tail and a C-terminal globular domain(7), no high-resolution information is available for PrP<sup>Sc</sup> due to technical challenges posed by its high insolubility and aggregation propensity(8). In order to fill this gap, different atomistic models based on low-resolution experimental data have been proposed, including a Left-handed- $\beta$ -Helix (L $\beta$ H) structure spanning residues 89 to 170 while retaining the two C-terminal  $\alpha$ -helices of PrP<sup>C</sup> (9), and a Parallel In-Register Beta-Sheet (PIRIBS) architecture, characterized by intermolecular stacking of aligned PrP monomers(10). The PIRIBS model is not consistent with recent cryo-EM data obtained using infectious, anchorless PrP<sup>Sc</sup> fibrils (Supp. Fig. 1)(4). Moreover, this model fails to accommodate glycosylated residues in PrP<sup>Sc</sup>, which would result in the introduction of excessive steric clashes(11). In turn, while consistent with the mentioned experimental constraints, the proposed L $\beta$ H model is incoherent with a recent re-evaluation of previous FTIR data suggesting that PrP<sup>Sc</sup> does not contain  $\alpha$ -helices(8).

## Results

To satisfy current experimental evidence and theoretical structural constraints, we built a new atomistic model of mouse PrP<sup>Sc</sup> based on the 4R $\beta$ S conformation, and tested its stability by means of all-atom MD simulations. The construction of the model took into account an array of experimental data, including: (i) cryo-EM(4) and X-ray fiber-diffraction studies(5), which showed that the fold of a mouse glycosylphosphatidylinositol (GPI)-anchorless, infectious PrP<sup>Sc</sup> is compatible with a 4R $\beta$ S architecture with L- or T- shaped cross-section(4); (ii) circular dichroism (CD) and FTIR spectroscopy, which ruled out the presence of  $\alpha$ -helices, and suggest that PrP<sup>Sc</sup> contains approximately 40-50%  $\beta$ -sheet and 50-60% coil/turns(8); (iii) Mass Spectrometry (MS) analyses indicating the presence of an intact disulphide bond between residues C178 and C213 (mouse sequence)(12), as well as mapping Proteinase K (PK)-sensitive residues, which reflect amino acids likely excluded from the resistant core of the protein(13)(14); and (iv) the possibility of accommodating complex glycans at positions N180 and N196(11). All these constraints were comprehensively included into a 2D threading scheme spanning mouse PrP (moPrP) residues 89-230 (Supp. Fig. 2), also considering the structural propensities of different residues: polyglycine tracts and prolines were positioned in loops due to their destabilizing effects on  $\beta$ -strands; charged sidechains were excluded from the inner core of the protein or counterbalanced by salt bridges. This scheme was then modelled onto the 3D arrangement of a naturally-occurring  $\beta$ -solenoid protein (*Dickeya dadantii* Pectate Lyase; PDB 1AIR). The resulting structure (depicted in Fig. 1) features an inner core containing mainly hydrophobic or mildly polar side-chains (T94, T106, L108, V111, Y127, M128, W144, Y149, V165, Y168, I181, I183, V188, F197, T198 and T200), few polar side-chains involved in hydrogen bonding (N142-HB-Y168, H168-HB-T198, Q216-HB-T200 and Q218-HB-S221), and a salt bridge (R147-SB-D166). Conversely, the majority of the highly-polar residues (N and Q) including the glycosylation sites (N180 and N196) and charged side-chains (E, D, K and R) are exposed to the solvent. The structure also encompasses identified PK cleavage sites localized in loops/turns, or at the edge of the  $\beta$ -strands, and the intact disulphide bond between C178 and C213. Importantly, the final model fitted with a previously described, low-resolution cryo-EM map of infectious PrP<sup>Sc</sup> (Supp. Fig. 3).

To test the physical consistency of the 4R $\beta$ S model, we challenged its stability by all-atom MD simulations in explicit solvent. First, three independent, 20 ns simulations were performed in the isothermal-isobaric ensemble (T = 300 K, P = 1 Bar) restraining hydrogen bonds distances between atoms involved in  $\beta$ -sheets. This process allowed relaxation of protein loops and side chains of the core (Supp. Fig. 4). Next, the imposed restraints were released, and three plain-MD trajectories of 100 ns each were simulated. The stability of the model was then assessed by both Root Mean Squared Deviation (RMSD) of atomic positions relative to the initial frame (t = 0 ns) and secondary structures content. Interestingly, we obtained values in the same range of fluctuation for our 4R $\beta$ S model and the prion-forming domain of the fungal protein Het-s, a naturally occurring  $\beta$ -solenoid whose structure has been solved by solid-state NMR (ssNMR; PDBs 2KJ3

and 2RNM; Fig. 2, and Supp. Fig. 5). Conversely, by applying an identical workflow to the previously proposed model of laterally-stacked L $\beta$ H trimer of PrP<sup>Sc</sup>, we observed a profound instability of the  $\beta$ -helical domain, which was already evident after few tens of ns of dynamics (Fig. 2 and Supp. Fig. 5 and 6).

To mimic the fibrillary behavior of PrP<sup>Sc</sup>, we built a tetrameric 4R $\beta$ S model by stacking monomers in a head-to-tail fashion. As expected, this assembly showed a comparable MD stability to the monomer (Fig. 3). Moreover, the 4R $\beta$ S tetramer fits with the two main signals obtained by Fourier transform single particle analysis of cryo-EM data (19.1 Å and the ~40 Å)(4), which reflect distances between the same residues on two contiguous or alternate monomers, respectively, as well as with the observed volume of the fibrils (Fig. 3). Finally, by introducing complex sugar precursors (GlcNAc2Man3Fuc) at positions N180 and N196 of each monomer, we observed the absence of steric clashes, confirming that the 4R $\beta$ S model accommodates the presence of glycans (Supp. Fig. 7). Collectively, these findings indicate that, in contrast to the L $\beta$ H model, the 4R $\beta$ S is a solid arrangement for PrP<sup>Sc</sup>, built on the most recent experimental data, and showing a conformational stability comparable to that of a natural prion.

The 4R $\beta$ S model allowed us to develop an original scheme to perform for the first time a simulation of the conformational transition from PrP<sup>C</sup> to PrP<sup>Sc</sup>. In order to bridge the gap between the computationally-accessible and the biologically-relevant time scales, we employed a specific kind of biased dynamics called ratchet-and-pawl MD (rMD)(15) in the framework of an all-atom realistic force field(16). rMD-based methods have been successfully applied to simulate protein folding and other conformational transitions(17, 18). However, this scheme provides a sampling of the transition path ensemble only if the biasing force is applied along a reliable reaction coordinate (19). Therefore, the first step towards developing our rMD simulation was to build a statistical model to identify the reaction coordinate of the process. Using the Markov State Mode formalism, we demonstrate that among all the possible misfolding patterns from PrP<sup>C</sup> to PrP<sup>Sc</sup> the prominent reaction mechanism is the sequential formation of rungs. A biasing coordinate was then built by coupling this dynamical information with the all atom 4R $\beta$ S structure. The associated rMD simulations yielded a transition pathway with full atomistic resolution in which the C-terminal rung of the solenoid acts as a primary conversion surface for PrP<sup>C</sup> unstructured N-terminus (residues 89-124). This first event initiates a cascade of conformational transitions in which each newly formed rung acts as a template for the formation of the following one, ultimately leading to the complete conversion of PrP<sup>C</sup> into PrP<sup>Sc</sup> (Fig. 4 and Supp. Movie). This analysis provides a rigorous description of the active role of fibril ends in the templated propagation of prions, and it is compatible with previous observations suggesting the presence of a structured intermediate conformer of PrP<sup>C</sup> in its transition to PrP<sup>Sc</sup> (20)

## Discussion

The search for an effective therapy against prion diseases has largely been unsuccessful, partly due to the lack of information regarding the structure of PrP<sup>Sc</sup>, which has hampered rational drug discovery efforts. Indeed, the elucidation of the structure of PrP<sup>Sc</sup> at atomic resolution has proven to be a phenomenal experimental challenge, mainly due to its high insolubility and aggregation propensity. Available computational models of PrP<sup>Sc</sup> have the virtue of providing a plausible 3D structure, but fail to comprehensively accommodate most recent experimental data. Here, we filled this gap by exploiting the information arising from cryo-EM and X-ray fiber-diffraction studies, which suggested a general architecture as a 4R $\beta$ S, and refined the structure by including experimental constraints obtained by mass spectrometry. This new model possesses several important features. It displays an intrinsic coherence with *state-of-the-art* knowledge about infectious PrP<sup>Sc</sup> fibrils, and it appears to be as stable as a naturally-occurring prion. It also represents a unique workbench for interpreting future structural data or available biological evidence, such as the effect of variations in the *PRNP* gene favoring or disfavoring prion propagation. Importantly, the 4R $\beta$ S model allowed us to perform the first reconstruction of how the information encoded into the conformation of a protein could be propagated in a directional fashion, a concept underlying the infectious nature of prions.

## Methods

### Template Selection & Model Building

In order to accommodate the moPrP 89-230 sequence, we selected a  $\beta$ -solenoid architecture capable of satisfying the following requirements: (i) Approximate number of residues in extended conformation higher than 12 per rung. Requirement needed to fit the secondary structure content of approximately 40% of  $\beta$ -sheets(8); (ii) Capacity of accommodating two extended loops per rung. Namely, the possibility of introducing arbitrarily long sequences between consecutive  $\beta$ -strands in order to allow the exclusion of PK cleavage sites, glycine tracts and prolines from the core of the solenoid; (iii) L or T cross-section solenoidal-shape. Architecture coherent with cryo-EM maps; (iv) Capacity of accommodating bulky aromatic residues in the hydrophobic core (e.g. moPrP 89-230 contains 11 Tyr, 3 Phe and 1 Trp). All these requirements are satisfied by the right-handed, L-shaped  $\beta$ -solenoid architecture. In contrast, left-handed-solenoid structures typically display smaller rungs(21), impairing the modelling of the desired number of residues in extended conformation as well as the accommodation of bulky residues in the core. The left-handed structure of the prion forming domain of Het-s (PDB 2KJ3 and 2RNM) is an exception in terms of  $\beta$ -sheet content, however its hydrophobic core is mainly composed of small side-chains and it allows the insertion of only one arbitrarily long loop connecting two consecutive rungs(3). The threading scheme for the PrP 89-230 sequence (Supp. Fig. 2) was obtained by following the L-shaped  $\beta$ -solenoid architecture. In the threading process, PK cleavage sites, prolines and glycine tracts were positioned in the loops (when this was not possible, at the edges of beta strands). Charged side-chains were excluded from the inner core of the solenoid or paired with residues forming salt bridges. The presence of an intact disulphide bond between C178 and C213 and solvent exposure of N180 and N196 sidechains (in order to accommodate glycans) were also considered.

The 3D structure of the monomeric model of PrP<sup>Sc</sup> was constructed following these steps: (i) 4 Rungs of *Dickeya dadantii* Pectate Lyase (PDB 1AIR; two repetitions of residues 168-235) were used to obtain the scaffold for the  $\beta$ -solenoid core; (ii) The original loops of the protein were removed; (iii) The residues in the hydrophobic core of the original PDB structure were replaced with PrP residues (as indicated in Supp. Fig. 2) using Chimera(22); (iv) New loops were generated using the MODELLER tool(23) in Chimera. Each loop was selected from a set of 20 proposed conformations. Structures resulting in extended atomic clashes were discarded, and finally the best performing model in term of DOPE score was selected; (v) Side-chain geometry was optimized using the Dunbrack's rotamer library. In particular, for highly polar and charged side-chains in the hydrophobic core, geometries capable of forming hydrogen bonds or salt bridges with nearby residues were selected; (vi) The energy of the system was minimized in vacuum with the steepest descent algorithm in Gromacs 4.6.5(24). System topology was generated using Amber99SB-ILDN force field(16). A restraining potential during energy minimization between the H and O atoms involved in hydrogen bond formation between backbone residues was added, as in equation 1:

$$(r_{ij}) \begin{cases} 0 & \text{for } r_{ij} < r_0 \\ \frac{1}{2}k_{dr}(r_{ij} - r_0)^2 & \text{for } r_0 \leq r_{ij} < r_1 \\ \frac{1}{2}k_{dr}(r_1 - r_0)(2r_{ij} - r_1 - r_0) & \text{for } r_{ij} > r_1 \end{cases} \quad (\text{Equation 1})$$

Where  $r_{ij}$  is the distance between the H and the O atoms involved in hydrogen bond formation;  $r_0$  is the original hydrogen bond distance after energy minimization ( $2 < r_0 < 2.5$ ) Å and  $r_1$  is an upper limit equal to 4 Å, while  $k_{dr}$  is set to  $2 \cdot 10^3$  kJ/(mol·nm<sup>2</sup>).

This strategy was applied to favor the movement of the side-chains and accommodation of loops, while impairing the backbone deviation of the residues involved in the  $\beta$ -solenoid core formation; (vi) Additional optimization of the backbone (in order to remove Ramachandran outliers) and side-chain geometry was performed using Coot(25); (vii) Absence of steric clashes was verified using Chimera, setting the VDW-overlap threshold as 0.6 Å, subtracting 0.4 Å to account H-Bonding pairs and ignoring contacts of pairs <4 bonds apart. The 3D structure of the tetrameric model of PrP<sup>Sc</sup> was assembled by stacking four monomers in a head-to-tail fashion using Chimera, maintaining the proposed threading. The strands 198-TETD and 215-TQYQKESQAYY were stacked on the top of 94-THNQ and 105-KTNLKHVAGAA of the forthcoming

monomer, respectively. The structure was then energy minimized following the same protocol used for the monomer. Glycans (DMan $\alpha$  1-6[DMan $\alpha$  1-3]DMan $\beta$  1-4DGlc $\beta$ NAc $\beta$  1-4[LFuc $\alpha$  1-6]DGlc $\beta$ NAc $\alpha$  1-Asn) attached to residues N180 and N196 were added using the doGlycans tool to the 4R $\beta$ S tetramer, and the structure was energy minimized in vacuum and then explicit solvent using Gromacs 5.1.2.

### Molecular Dynamics Simulations

Molecular Dynamics simulations were performed on the supercomputers Finis Terrae II (CESGA, Santiago de Compostela, Spain) and Marconi (CINECA, Bologna, Italy) using Gromacs 5.1.2. The following protocol was applied for the whole set of simulations performed in this work. Protein topologies were generated using Amber 99SB-ILDN force field. The structures were accommodated in a dodecahedral box with periodic boundary conditions. The minimum wall distance from the protein was set to 12 Å. The box was filled with TIP3P water molecules. The charge of the system was neutralized with the addition of Na<sup>+</sup> or Cl<sup>-</sup> ions. The system was energy minimized in explicit solvent using a steepest descent algorithm, retaining the restraining potential (equation 1). From the minimized structure, three independent equilibrations with position restraints on heavy atoms were launched: in the NVT ensemble (300 K) for 500 ps using the Nose-Hoover thermostat, and then in the NPT ensemble (300 K, 1 Bar) for 500 ps using the Nose-Hoover thermostat and the Parrinello-Rahman barostat. For each equilibrated system, a 20 ns MD simulation with restraining potential (Equation 1) was launched. Finally, restraints were released and each trajectory was extended with 100 ns of plain-MD. This protocol yielded three independent, unbiased 100 ns MD trajectories for each structure of interest. MD simulations were performed using a leap-frog integrator with step equal to 2 fs. LINCS algorithm was applied to handle bonds constraints. Cut-off for short-range Coulomb and Van der Waals interactions was set to 12 Å, while long range electrostatics was treated using Particle Mesh Ewald.

### Generation of the Propagation Model

A key step to perform reliable rMD simulations is the identification of a reasonable reaction coordinate. To this end, we first developed a statistical coarse-grained model which enables us to identify the dominant reaction pathway underlying the conversion of PrP<sup>C</sup> into a 4R $\beta$ S PrP<sup>Sc</sup>. In order to describe the reaction kinetics we developed a simple stochastic model in which the key rate-limiting processes are assumed to be the irreversible formation and docking of the four rungs of the 4R $\beta$ S. We indicated R<sub>0</sub> as the C-terminal rung of the pre-formed 4R $\beta$ S, and with R<sub>1</sub> R<sub>2</sub> R<sub>3</sub> R<sub>4</sub> the consecutive rungs of the converting monomer. The instantaneous state of the system can be represented as a 4-dimensional vector of binary entries  $S = [n_0, n_1, n_2, n_3]$ , where  $n_k = 1$  in the presence of docking between rung R<sub>k</sub> and rung R<sub>k+1</sub>, and  $n_k = 0$  otherwise. We emphasize that this model excludes the presence standalone rungs, which would correspond to an entropically-unfavorable single extended conformation, not stabilized by hydrogen bonds of nearby  $\beta$ -strands. On the other hand, misfolded rungs can be stabilized either upon docking to a pre-existing misfolded region (template mechanism) or through a process in which two rungs simultaneously form and dock. We modeled the transition of an initial state  $S_R = [0,0,0,0]$  (in which the PrP<sup>C</sup> monomer is in the native state and none of the rungs are formed) to the fully misfolded state  $S_P = [1,1,1,1]$  (where the monomer is completely misfolded and incorporated into PrP<sup>Sc</sup>). The resulting network is represented in Supp. Fig. 8a which contains all the possible combinations of docking events leading  $S_R$  to  $S_P$  through a sequence of irreversible transitions. The model can be simplified considering that rate  $k_2$  is expected to be negligible as compared to  $k_0$ . Indeed, while the event associated with  $k_0$  only requires the structuring of a disordered PrP region, the events associated with  $k_2$  require the loss of native content (breakage of hydrogen bonds in the helical regions) together with the simultaneous formation of two rungs (two-fold entropic cost). Thus, it is possible to disregard all the reaction pathways in which the first step of the reaction is a transition occurring at a rate  $k_2$  and consider only events starting from  $[1,0,0,0]$ . The network of the resulting simplified model is depicted in Supp. Fig. 8b. Furthermore, we can set  $k_3/k_1 \gg 1$  and  $k_3/k_2 \gg 1$ , since the docking of two pre-formed rungs occurs at a rate much faster than all processes involving misfolding. The reaction kinetics in this stochastic model was simulated through a Kinetic Monte Carlo algorithm (arbitrarily setting  $k_3/k_1 = 10^6$ ) and the resulting reaction mechanisms were studied as a function of the  $k_2/k_1$  ratio. Namely, we enumerated all the reaction pathways and computed the corresponding probability:



	Step 1	Step 2	Step 3	Step 4
Path 1	[1,0,0,0]	[1,1,0,0]	[1,1,1,0]	[1,1,1,1]
Path 2	[1,0,0,0]	[1,0,1,0]	[1,1,1,0]	[1,1,1,1]
Path 3	[1,0,0,0]	[1,0,0,1]	[1,0,1,1]	[1,1,1,1]
Path 4	[1,0,0,0]	[1,0,0,1]	[1,1,0,1]	[1,1,1,1]
Path 5	[1,0,0,0]	[1,1,0,0]	[1,1,0,1]	[1,1,1,1]
Path 6	[1,0,0,0]	[1,0,1,0]	[1,0,1,1]	[1,1,1,1]

We find that Path 1 (which consists in the consecutive formation of all rungs by templating on previously misfolded structures) dominates over all others as soon as  $k_1/k_2 \geq 4$ . Using Kramers' theory and assuming comparable pre-factors, we find that the templating mechanism is the most prominent reaction pathway when the activation energies for single and double rung formations obey the relationship:

$$\ln \frac{k_1}{k_2} = \Delta G_2^\ddagger - \Delta G_1^\ddagger \gg 1.5 k_B T \quad (\text{Equation 2})$$

We expect this condition to be always satisfied. Indeed, processes 1 and 2 are characterized by the formation of the same number of hydrogen bonds, leading approximately to the same enthalpic gain. However, the process described by rate  $k_2$  requires the breaking of the double amount of native contacts, together with a two-fold entropy loss (compared to the single rung formation) when forming two de novo rungs. We can therefore conclude that the propagation reaction proceeds through a subsequent formation of individual rungs, using as template pre-existing 4R $\beta$ S free end. We emphasize that our approach is likely to underestimate the dominance of the sequential misfolding mechanism. Indeed, it does not account for the direct cooperativity of hydrogen bonds and the long-range electrostatics favoring  $\beta$ -strand formation in presence of pre-formed  $\beta$ -sheets, as directly supported by previous computational and experimental evidence(26, 27).

The coarse-grained information about the reaction mechanism obtained by means of our theoretical model can be exploited to set up a fully atomistic rMD simulation of the PrP<sup>C</sup>  $\rightarrow$  PrP<sup>Sc</sup> transition, using the 4R $\beta$ S as a target structure (Fig. 4 and Supp. Movie). The 3D structure of moPrP<sup>C</sup> (residues 105-230) was obtained by linking moPrP 121-231 (PDB 1XYX) to the adapted N-terminal fragment (residues 105-120) of huPrP<sup>C</sup> (PDB 5yj5) which was mutated to the moPrP sequence. The initial state for the conversion simulation was generated by modifying a central dimer extracted from the tetrameric 4R $\beta$ S at the end of 20 ns of restrained molecular dynamics simulation. The initial contact point between PrP<sup>Sc</sup> and PrP<sup>C</sup> was generated by leaving residues 89-104 of the C-terminal monomer anchored to the  $\beta$ -solenoid, which was then replaced by moPrP<sup>C</sup> retaining the original disulfide bond. The rationale behind this modelling scheme derives from multiple previous reports indicating that the same region is a primary contact point between PrP<sup>C</sup> and PrP<sup>Sc</sup> (28), as well as by uncertainties regarding the state of the disulfide bond during PrP conformational transition. The complex was energy minimized in vacuum and then in explicit solvent in a dodecahedral box (with periodic boundary conditions and minimum wall distance of 20 Å) also containing six Cl<sup>-</sup> ions to neutralize the charge of the system, using a steepest descent algorithm (protein topology was generated by Amber99SB-ILDN force field). The system was then equilibrated in the NVT ensemble (350 K) for 200 ps using the Nose-Hoover thermostat, and then in the NPT ensemble (350 K, 1 Bar) for 200 ps using the Nose-Hoover thermostat and the Parrinello-Rahman barostat. The model was then subjected to a modified protocol of rMD simulations, adapting the method to a sequential biasing. This scheme resulted in a progressive rMD simulation in which the structure of PrP<sup>C</sup> was targeted to the growing solenoid in a rung-by-rung fashion. Target structures included an entire 4R $\beta$ S solenoid and additional rungs at the growing C-terminus: Step 1, residues 89-115; Step 2, residues 89-151; Step 3, residues 89-190; Step 4, residues 89-230. In this rMD scheme, the equation of motion is modified by an history-dependent biasing force  $\mathbf{F}_i^{\text{rMD}}(X,t)$ , defined as:

$$\mathbf{F}_i^{\text{rMD}}(X,t) \equiv -k_R \nabla_i z(X) (z[X(t)] - z_m(t)) \theta(z[X(t)] - z_m(t)) \quad (\text{Equation 3})$$

where  $k_R$  is the ratchet spring constant determining the strength of the biasing force applied on the system, in this case  $10^{-2}$  kJ/mol.  $\theta$  function is equal to 1 if its argument is positive, its value 0 otherwise.  $Z[X(t)]$  is a collective coordinate defined as:

$$z[X(t)] \equiv \sum_{|i-j|>35}^N \{C_{ij}[X(t)] - C_{ij}^0\}^2 \quad (\text{Equation 4})$$

where  $C_{ij}[X(t)]$  is the instantaneous contact map and  $C_{ij}^0$  is the contact map of the target state. Their entries are chosen to smoothly interpolate between 0 and 1, depending on the relative distance of the atoms  $i$  and  $j$ . The  $z_m$  function indicates the smallest value assumed by the reaction coordinate  $z[X(t)]$  up to time  $t$  and. The contact map  $C_{ij}(X)$  is defined as:

$$C_{ij}(X) = \begin{cases} \frac{1-(r_{ij}/r_0)^6}{1-(r_{ij}/r_0)^{10}} & \text{if } r_{ij} < r_c \\ 0 & \text{if } r_{ij} > r_c \\ \frac{6}{10} & \text{if } r_{ij} = r_c \end{cases} \quad (\text{Equation 5})$$

Where  $r_{ij}$  is the distance between the  $i$ -th and the  $j$ -th atom;  $r_0$  is a typical contact distance set to 7.5 Å;  $r_c$  is an upper limit to improve computational efficiency by excluding excessively distant atoms from the calculation (and it is set to 12.3 Å). In this way, no bias force acts on the system when the protein spontaneously proceeds towards the target state, while the external biasing force is only applied when the polypeptide tends to backtrack toward the initial state. We terminated each rMD simulation when the RMSD of the protein relative to the final state is stably lower than 0.8 Å. rMD simulations were performed in explicit solvent using Gromacs 4.6.5 with the plugin Plumed 2.0.2(29). Integration of motion was performed using a leap-frog algorithm with step equal to 2 fs. Temperature was maintained at 350 K (approximately PrP melting temperature) and pressure at 1 Bar using Nose-Hoover thermostat and the Parrinello-Rahman barostat. LINCS algorithm was applied to handle bonds constraints. Cut-off for short-range Coulomb and Van der Waals interactions was set to 10 Å, while long range electrostatics was treated using Particle Mesh Ewald.

### Data Analysis

Root Mean Squared Deviation of atomic positions (RMSD) was calculated “all-atoms” using Gromacs 5.1.2. The specific calculation of the  $\beta$ -solenoid/helix core RMSD involves all the following residues for the different systems:

4R $\beta$ S Model: T94, H95, N96, Q97, K105, T106, N107, L108, K109, H110, V111, A112, G113, A114, A115, M128, L129, G130, S131, G141, N142, D143, W144, E145, D146, R147, Y148, Y149, R150, E151, D166, Q167, Y168, S169, N180, I181, T182, I183, K184, Q185, H186, T187, V188, T189, T190, T198, E199, T200, D201, T215, Q216, Y217, Q218, K219, E220, S221, Q222, A223, Y224, Y225.

Het-s Dimer: R225, N226, S227, A228, K229, D230, I231, R232, T233, E234, E235, R236, A237, R238, V239, Q240, L241, G242, V244, T261, N262, S263, V264, E265, T266, V267, V268, G269, K270, G271, E272, S273, R274, V275, L276, I277, G278, N279, E280.

L $\beta$ H Model: G89, Q90, G91, G92, G93, T94, H95, N96, Q97, W98, N99, K100, N107, L108, K109, H110, V111, A112, G113, A114, A115, A116, A117, G118, A119, V120, V121, G122, G123, L124, G125, G126, T127, M128, L129, G130, S131, A132, M133, S134, R135, P136, M137, I138, H139, F140, G141, N142, D143, W144, E145, D146, D166, Q167, Y168, S169.

Secondary structure content and fraction of native/misfolded contacts were obtained using the Timeline and the Trajectory tools in VMD 1.9.2. Fraction of native and misfolded contacts are defined in the following equation



$$S(X) \equiv \frac{\sum_{|i-j|>4} \theta(r_0 - r_{ij}(X)) \theta(r_0 - r_{ij}(X_S))}{\sum_{|i-j|>4} \theta(r_0 - r_{ij}(X_S))} \quad (\text{Equation 6})$$

where  $r_{ij}$  is the distance between the  $i$ -th and  $j$ -th  $C_\alpha$  atoms,  $r_0$  is the typical distance defining a contact between two residues and it is set at 7.5 Å,  $X$  is the instantaneous configuration of the protein during the simulation,  $X_S$  is the atomic configuration of the protein in the reference state, that is the native PrP<sup>C</sup> structure for the native contacts calculation, and the PrP<sup>Sc</sup> structure for the misfolded contacts calculations. This quantity is evaluated only for the converting monomer.

Initial and final values of RMSD and secondary structure content were calculated by averaging the values over the last 5ns (of restrained or unrestrained MD) and then calculating the mean and standard deviation for the three trajectories. Inter-monomer distances in the tetramers are calculating using Chimera as an average of the distances between the mid-residue of each  $\beta$ -strand and the same residue in the first or the second next monomer. In particular, for  $\beta$ -strand-1: N96, G130, Y168, T200; for  $\beta$ -strand-2: N107, D143, T192, Y217; for  $\beta$ -strand-3: G113, Y149, V188, A223. Tetramer volumes were calculated using Chimera. Molecular graphics images were produced using the UCSF Chimera package from the Computer Graphics Laboratory, while graphs are plotted using Gnuplot.

## References.

1. Prusiner SB (1982) Novel proteinaceous infectious particles cause scrapie. *Science* 216(4542):136-144.
2. Gremer L, *et al.* (2017) Fibril structure of amyloid-beta(1-42) by cryo-electron microscopy. *Science* 358(6359):116-119.
3. Wasmer C, *et al.* (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 319(5869):1523-1526.
4. Vazquez-Fernandez E, *et al.* (2016) The Structural Architecture of an Infectious Mammalian Prion Using Electron Cryomicroscopy. *PLoS Pathog* 12(9):e1005835.
5. Wille H, *et al.* (2009) Natural and synthetic prion structure from X-ray fiber diffraction. *Proc Natl Acad Sci U S A* 106(40):16990-16995.
6. Prusiner SB (1998) The prion diseases. *Brain Pathol* 8(3):499-513.
7. Zahn R, *et al.* (2000) NMR solution structure of the human prion protein. *Proc Natl Acad Sci U S A* 97(1):145-150.
8. Requena JR & Wille H (2014) The structure of the infectious prion protein: experimental data and molecular models. *Prion* 8(1):60-66.
9. Govaerts C, Wille H, Prusiner SB, & Cohen FE (2004) Evidence for assembly of prions with left-handed beta-helices into trimers. *Proc Natl Acad Sci U S A* 101(22):8342-8347.
10. Groveman BR, *et al.* (2014) Parallel in-register intermolecular beta-sheet architectures for prion-seeded prion protein (PrP) amyloids. *J Biol Chem* 289(35):24129-24142.
11. Baskakov IV & Katorcha E (2016) Multifaceted Role of Sialylation in Prion Diseases. *Front Neurosci* 10:358.
12. Welker E, Raymond LD, Scheraga HA, & Caughey B (2002) Intramolecular versus intermolecular disulfide bonds in prion proteins. *J Biol Chem* 277(36):33477-33481.
13. Vazquez-Fernandez E, *et al.* (2012) Structural organization of mammalian prions as probed by limited proteolysis. *PLoS One* 7(11):e50111.
14. Sevillano AM, *et al.* (2018) Recombinant PrP<sup>Sc</sup> shares structural features with brain-derived PrP<sup>Sc</sup>: Insights from limited proteolysis. *PLoS Pathog* 14(1):e1006797.
15. Tiana G & Camilloni C (2012) Ratcheted molecular-dynamics simulations identify efficiently the transition state of protein folding. *J Chem Phys* 137(23):235101.
16. Lindorff-Larsen K, *et al.* (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78(8):1950-1958.
17. S AB, Fant L, & Faccioli P (2015) Variational scheme to compute protein reaction pathways using atomistic force fields with explicit solvent. *Phys Rev Lett* 114(9):098103.
18. Ianeselli A, *et al.* (2018) Atomic Detail of Protein Folding Revealed by an Ab Initio Reappraisal of Circular Dichroism. *J Am Chem Soc* 140(10):3674-3682.
19. Bartolucci G, Orioli S, & Faccioli P (2018) Transition path theory from biased simulations. *J Chem Phys* 149(7):072336.
20. Nicholson EM, Mo H, Prusiner SB, Cohen FE, & Marqusee S (2002) Differences between the prion protein and its homolog Doppel: a partially structured state with implications for scrapie formation. *J Mol Biol* 316(3):807-815.
21. Kajava AV & Steven AC (2006) Beta-rolls, beta-helices, and other beta-solenoid proteins. *Adv Protein Chem* 73:55-96.
22. Pettersen EF, *et al.* (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605-1612.
23. Sali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234(3):779-815.
24. Berendsen HJC, Vandespoel D, & Vandrunen R (1995) Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Comput Phys Commun* 91(1-3):43-56.
25. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D* 66:486-501.
26. Zhao YL & Wu YD (2002) A theoretical study of beta-sheet models: is the formation of hydrogen-bond networks cooperative? *J Am Chem Soc* 124(8):1570-1571.

27. Kung VM, Cornilescu G, & Gellman SH (2015) Impact of Strand Number on Parallel beta-Sheet Stability. *Angew Chem Int Ed Engl* 54(48):14336-14339.
28. Moroncini G, *et al.* (2004) Motif-grafted antibodies containing the replicative interface of cellular PrP are specific for PrPSc. *Proc Natl Acad Sci U S A* 101(28):10404-10409.
29. Bonomi M, *et al.* (2009) PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput Phys Commun* 180(10):1961-1972.
30. Friedman R & Caflisch A (2014) Wild type and mutants of the HET-s(218-289) prion show different flexibility at fibrillar ends: a simulation study. *Proteins* 82(3):399-404.
31. Mizuno N, Baxa U, & Steven AC (2011) Structural dependence of HET-s amyloid fibril infectivity assessed by cryoelectron microscopy. *Proc Natl Acad Sci U S A* 108(8):3252-3257.

## Figure legends

**Fig.1. View of the 4R $\beta$ S PrP<sup>Sc</sup> model.** The structure of PrP<sup>Sc</sup> modelled as a 4R $\beta$ S ( $\beta$ -strands represented as red arrows) is depicted in the center of the figure. Residues are displayed in each individual rung (1-4) with different colors. PK cleavage sites identified by mass spectrometry in two different previous reports are colored in red (13) or orange(14). Glycosylation sites are labelled in green. Prolines are colored in purple. Cystine is indicated in cyan.

**Fig.2. Comparison of stability by MD simulations of a Het-s dimer, 4R $\beta$ S and L $\beta$ H PrP<sup>Sc</sup>.** Upper graphs report the RMSD deviation from initial configuration for the entire structure (blue lines) or the  $\beta$ -solenoid core (orange lines) of the different proteins. Filled curves indicate standard error of the mean. Results show a comparable stability between the Het-s dimer (a) and the 4R $\beta$ S model (b), with an average RMSD of the hydrophobic core (calculated as the average of the three trajectories over the last 5 ns,  $\pm$  standard deviation) of  $2.4 \pm 0.2$  Å for the Het-s dimer (similar results for Het-s stability are reported(30)) and  $2.1 \pm 0.3$  Å for the 4R $\beta$ S. In contrast, the structural deviation of the L $\beta$ H (c) hydrophobic core is approximately two-fold higher, reaching a value of  $4.3 \pm 0.6$  Å. Lower graphs indicate the  $\alpha$ -helical (blue lines) or  $\beta$ -sheet (red lines) content of each protein. The initial and final  $\beta$ -sheet content was calculated as the average of the three trajectories over the last 5 ns of the restrained and unrestrained MD simulations, respectively. The Het-s dimer showed a variation from an initial  $41.8 \pm 2.3\%$  to a final  $33.2 \pm 2.8\%$ . Similarly, the 4R $\beta$ S model deviates from an initial  $41.3 \pm 2.3\%$  to a final  $36.6 \pm 5.4\%$ . Instead, the L $\beta$ H model deviates from a starting  $18.3 \pm 1.4\%$  to a  $7.5 \pm 3.1\%$ . These results are illustrated by the structures shown below the graphs, which represent the initial (top) and final (bottom) frames of the MD trajectories.

**Fig. 3. MD simulation and PK restriction map of a 4R $\beta$ S tetramer.** **a**, 3D representation of a 4R $\beta$ S tetramer at the beginning (left) and the end (right) of MD simulations. Blue bars indicate the distance between two residues in the same position on two consecutive monomers, which corresponds to  $19.2 \pm 0.4$  Å ( $t = 0$  ns) and  $19.8 \pm 2.2$  Å ( $t = 100$  ns). Purple bars indicate the distance between a residue in one monomer and the same residue on the second forthcoming monomer, which corresponds to  $38.4 \pm 0.5$  Å ( $t = 0$  ns) and  $39.6 \pm 1.8$  Å ( $t = 100$  ns). A similar pattern of signals reflecting monomeric and dimeric repeats has previously been observed by cryo-EM studies on Het-s(31). Both values are in almost perfect agreement with the two main signals obtained by Fourier transform single particle analysis in the cryo-EM experiment ( $19.1$  Å and the  $\sim 40$  Å signals). The average monomeric model volume calculated at beginning and at the end of the 100 ns tetramer simulations are equal to  $(18.5 \pm 0.7) \cdot 10^3$  Å<sup>3</sup> and  $(18.5 \pm 0.3) \cdot 10^3$  Å<sup>3</sup> respectively; while the estimated cryo-EM monomeric volume from the fibril is equal to  $18.9 \cdot 10^3$  Å<sup>3</sup> **b**, Upper graph shows the RMSD deviation of the tetramer from initial state for the entire structure (blue lines) or the  $\beta$ -solenoid core (orange lines). Structural deviation over the 100 ns of simulation corresponds to  $2.6 \pm 0.2$  Å. Lower graph report the secondary structures percentage, initial  $\beta$ -strand content is  $46.2 \pm 1.2$  %, while the final  $38.6 \pm 1.7$  %. Filled curves indicate standard error of the mean. These results indicate a comparable stability between monomeric and tetrameric 4R $\beta$ S structures. **c**, Map of the experimentally observed PK cleavage sites (identified by mass spectrometry in two previous reports are colored in red or orange) overlapped with the probability of each residue to be in an extended conformation, calculated over the last 5 ns of the MD trajectories.

**Fig.4. Graphs and frames extracted from the rMD simulation of PrP<sup>C</sup> to PrP<sup>Sc</sup> conversion.** Upper graph reports the fraction of native and misfolded contacts of the instantaneous configurations of the PrP<sup>C</sup>-PrP<sup>Sc</sup> complex, starting from the initial modelled configuration (depicted in 1) in which PrP<sup>C</sup> contacts the 4R $\beta$ S monomer, with respect to the final target configuration (4R $\beta$ S dimer). Lower graph shows the secondary structure content. Pictures of the evolving complex represent frames extracted from the entire conversion simulation at precise rMD steps, corresponding to the refolding of PrP<sup>C</sup> as follow: (2) residues 89-115; (3) residues 89-151; (4) residues 89-190; (5) residues 89-230. The process highlights the progressive unfolding and refolding of PrP<sup>C</sup> onto the 4R $\beta$ S template, which initially involves the unstructured region, followed by the loss of  $\alpha$ -helices and a progressive formation of  $\beta$ -sheets.

## Supp. figure legends

**Supp. Fig. 1. PIRIBS model of PrP<sup>Sc</sup> superimposed to the cryo-EM map.** The picture illustrates that the size of PIRIBS (80 Å x 50 Å) does not fit the fibril cross section (50 Å x 30 Å) obtained by cryo-EM (level of contouring equal to 3.05 in order to match the measured fibril diameter).

**Supp. Fig. 2. 2D threading scheme of moPrP residues 89-230.** 2D arrangement based on the general architecture of right-handed  $\beta$ -solenoid proteins with L-shaped cross section. Blue thick arrows indicate  $\beta$ -strands, thin cyan and thin orange arrows indicate the sidechain orientation (toward the solvent or toward the hydrophobic core, respectively). The scheme is used to thread the moPrP 89-230 sequence by considering different constraints. PK cleavage sites identified by mass spectrometry in two different previous reports are colored in red or orange. Glycosylation sites are labelled in green. Prolines are colored in purple. Cystine is indicated in cyan.

**Fig. 3. MD simulation and PK restriction map of a 4R $\beta$ S tetramer.** **a**, 3D representation of a 4R $\beta$ S tetramer at the beginning (left) and the end (right) of MD simulations. Blue bars indicate the distance between two residues in the same position on two consecutive monomers, which corresponds to  $19.2 \pm 0.4$  Å ( $t = 0$  ns) and  $19.8 \pm 2.2$  Å ( $t = 100$  ns). Purple bars indicate the distance between a residue in one monomer and the same residue on the second forthcoming monomer, which corresponds to  $38.4 \pm 0.5$  Å ( $t = 0$  ns) and  $39.6 \pm 1.8$  Å ( $t = 100$  ns). A similar pattern of signals reflecting monomeric and dimeric repeats has previously been observed by cryo-EM studies on Het-s(31). Both values are in almost perfect agreement with the two main signals obtained by Fourier transform single particle analysis in the cryo-EM experiment (19.1 Å and the  $\sim 40$  Å signals). The average monomeric model volume calculated at beginning and at the end of the 100 ns tetramer simulations are equal to  $(18.5 \pm 0.7) \cdot 10^3$  Å<sup>3</sup> and  $(18.5 \pm 0.3) \cdot 10^3$  Å<sup>3</sup> respectively; while the estimated cryo-EM monomeric volume from the fibril is equal to  $18.9 \cdot 10^3$  Å<sup>3</sup>. **b**, Upper graph shows the RMSD deviation of the tetramer from initial state for the entire structure (blue lines) or the  $\beta$ -solenoid core (orange lines). Structural deviation over the 100 ns of simulation corresponds to  $2.6 \pm 0.2$  Å. Lower graph report the secondary structures percentage, initial  $\beta$ -strand content is  $46.2 \pm 1.2$  %, while the final  $38.6 \pm 1.7$  %. Filled curves indicate standard error of the mean. These results indicate a comparable stability between monomeric and tetrameric 4R $\beta$ S structures. **c**, Map of the experimentally observed PK cleavage sites (identified by mass spectrometry in two previous reports are colored in red or orange) overlapped with the probability of each residue to be in an extended conformation, calculated over the last 5 ns of the MD trajectories.

**Supp. Fig. 3. 4R $\beta$ S model of PrP<sup>Sc</sup> superimposed to the cryo-EM map.** The pictures illustrate that the size of 4R $\beta$ S fits with the fibril cross section (50 Å x 30 Å) obtained by cryo-EM. **a**, Level of contouring was set to 3.05 to match the measured fibril diameter. **b**, Level of contouring was set to 3.60 to highlight the superimposition of the hydrophobic core of the 4R $\beta$ S with the highest electronic density region of the fibril.

**Supp. Fig. 4. Summary of restrained MD simulations of the Het-s dimer, 4R $\beta$ S and L $\beta$ H PrP<sup>Sc</sup>.** Upper graphs report the RMSD deviation from initial configurations for the entire structure (blue lines) or the  $\beta$ -solenoid core (orange lines) of the different proteins during the 20 ns of restrained MD simulations. Filled curves represents standard error of the mean. The graphs indicate a minor rearrangement of the  $\beta$ -solenoid core that is almost identical for the Het-s dimer and the 4R $\beta$ S model, characterized by a RMSD of  $2.3 \pm 0.2$  Å and  $2.4 \pm 0.2$  Å respectively (calculated as the average of the three trajectories over the last 5 ns,  $\pm$  standard deviation). In contrast, the hydrophobic core of the L $\beta$ H model displays a higher RMSD deviation ( $3.3 \pm 0.3$  Å). Lower graphs indicate the  $\alpha$ -helical (blue lines) or  $\beta$ -sheet (red lines) content of each protein. The latter is stable for all the three structures, likely due to the presence of the restraining potential for the whole length of the three simulations.

**Supp. Fig. 5. Probability distribution of  $\beta$ -sheet content for the Het-s dimer, 4R $\beta$ S and L $\beta$ H PrP<sup>Sc</sup>.** Graphs show the probability of each residue to be in an extended conformation at the beginning of the simulation

(green) or at the end of the simulation (purple) for the Het-s dimer, 4R $\beta$ S and L $\beta$ H PrP<sup>Sc</sup> models. Initial or final probabilities were calculated from the last 5 ns of the restrained or unrestrained simulations, respectively.

**Supp. Fig. 6. Initial and final conformations of L $\beta$ H PrP<sup>Sc</sup> after MD simulations.** Pictures show the energy minimized structure (1), the same structure at end of the restrained (2) or unrestrained (3) MD simulations of the L $\beta$ H model.

**Supp. Fig. 7. Illustration of the energy minimized 4R $\beta$ S tetramer carrying complex glycans.** A model of energy minimized tetrameric 4R $\beta$ S structure carrying glycan residues is depicted in two different orientations (a, b). Top view of a laterally stacked trimer of glycosylated 4R $\beta$ S (c), compatible with the 2D crystals diffraction data. A scheme of the complex glycan precursor added is depicted in (d), blue squares indicate N-acetylglucosamine, green circles indicate mannose and the red triangle indicate fucose.

**Supp. Fig. 8. Statistical model for prion propagation mechanism.** (a) Schematic representation of the network of transitions leading to the incorporation of PrP<sup>C</sup> into PrP<sup>Sc</sup>. The purple arrow describes the formation of a single rung by templating the unstructured region of PrP<sup>C</sup> (89-115) on PrP<sup>Sc</sup> (rate  $k_0$ ). The blue arrows describe the formation of a single rung by means of a templating process which involves the breaking the native contacts and the docking onto a pre-formed rung (rate  $k_1$ ). The orange arrows indicate the spontaneous formation and docking of two de-novo rungs (rate  $k_2$ ). Finally, the green arrows indicate the docking of preformed rungs belonging to two adjacent misfolded regions of the same chain (rate  $k_3$ ). (b) Simplified version of the network assuming as a priming reaction step the formation of the first rung. (c) Relative probability of the 6 reaction pathways as a function of the  $k_1/k_2$  rate ratio.

**Supp. Movie. Full-atomistic model of prion propagation.** Visualization of the rMD simulation reconstructing the entire refolding events of PrP<sup>C</sup> onto the 4R $\beta$ S PrP<sup>Sc</sup> model.



**Acknowledgements.** HW acknowledges support through grant 201600029 from the Alberta Prion Research Institute. JRR was funded by grants BFU2017-86692-P and BFU2013-48436-C2-1-P from the Spanish Ministries of Economy and Competitiveness, and Education and Science, respectively, both partially including FEDER funds from the European Union. Calculations were performed using supercomputing facilities Finis Terrae II (CESGA, Spain) and Marconi (CINECA, Italy). This work was supported by a grant from Fondazione Telethon (Italy, TCP14009). GS is a recipient of a fellowship from Fondazione Telethon. EB is an Assistant Telethon Scientist at the Dulbecco Telethon Institute.

**Author contributions.** JRR, PF, EB, HW and GS conceived and developed the project; JRR, AMS, GS and MR designed and built the 4R $\beta$ S PrP<sup>Sc</sup> model; GS and MR performed MD simulations and analyses; EB, PF, SO, GS and MR developed methods to reconstruct PrP misfolding; EB, GS, MR, PF and JRR wrote the original version of the manuscript, which was revised and accepted by all the authors.

**Competing interests.** All the authors declare no conflict of interests. EB, GS and PF are co-founders of Sibylla Biotech SRL, a startup company focused on developing new approaches for rational drug design.

**Materials & Correspondence.** Correspondence and/or material should be address to GS ([giovanni.spagnolli@unitn.it](mailto:giovanni.spagnolli@unitn.it)), EB ([emiliano.biasini@unitn.it](mailto:emiliano.biasini@unitn.it)) or JRR ([jesus.requena@usc.es](mailto:jesus.requena@usc.es)).

Figure 1

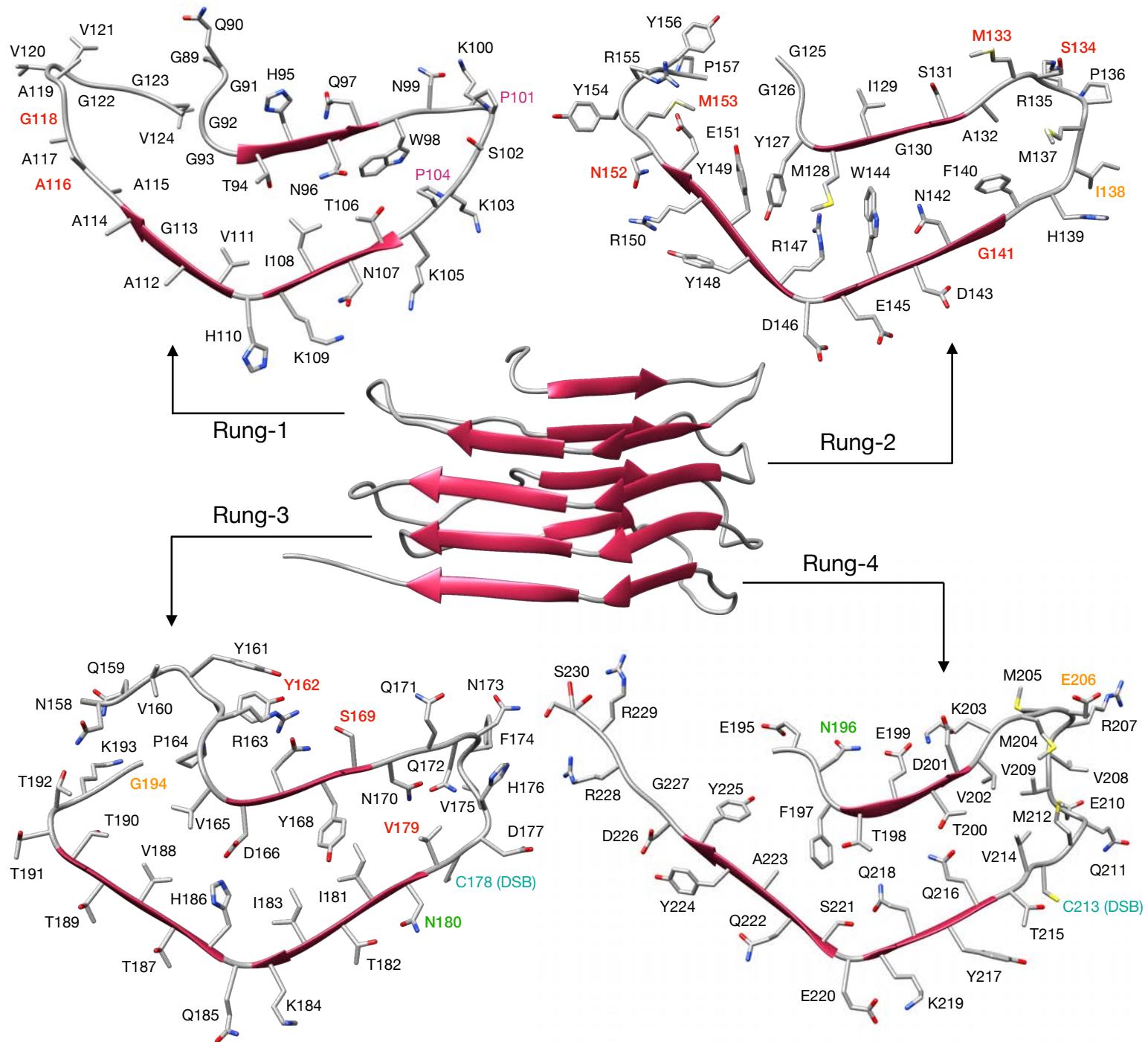


Figure 2

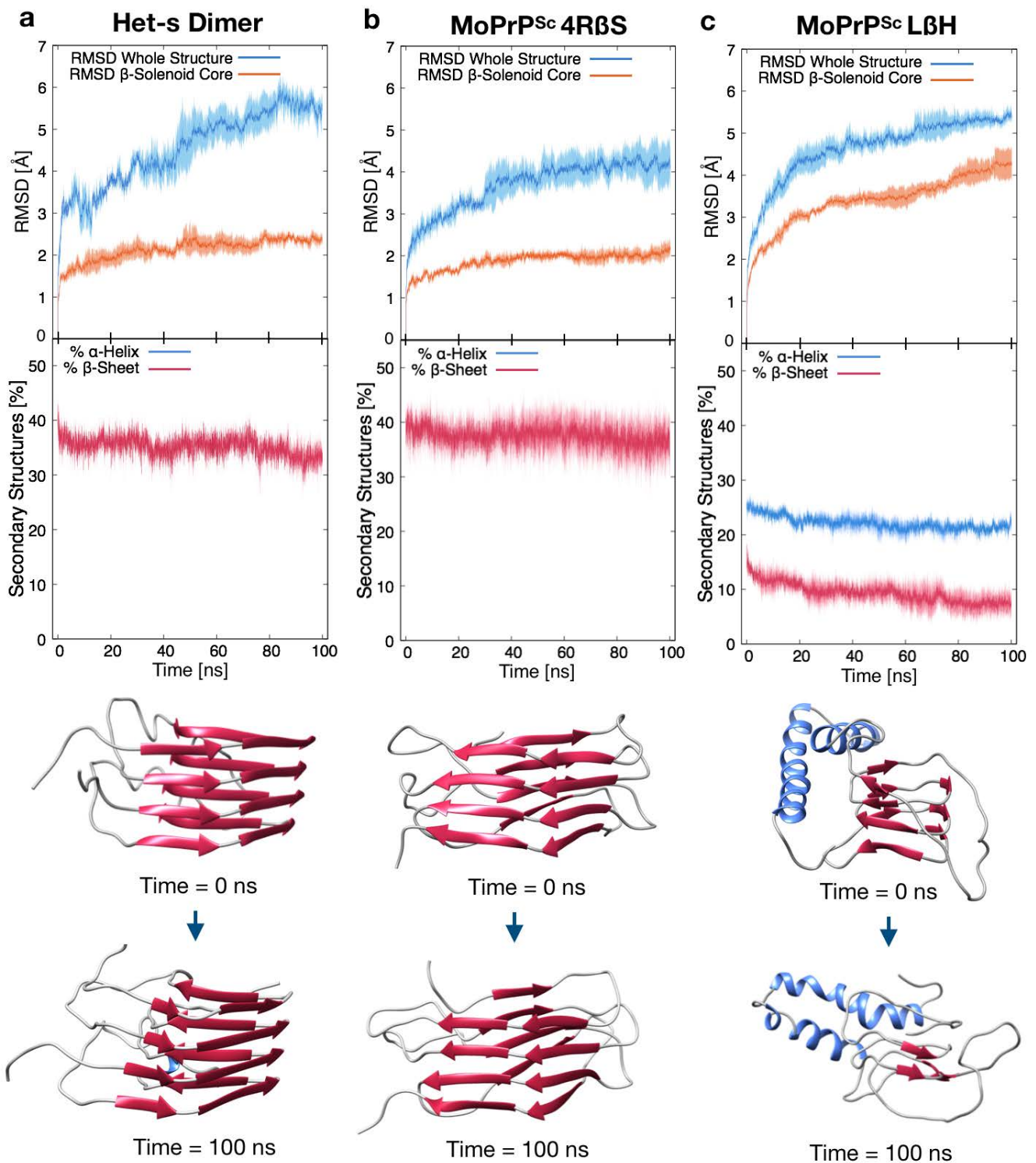


Figure 3

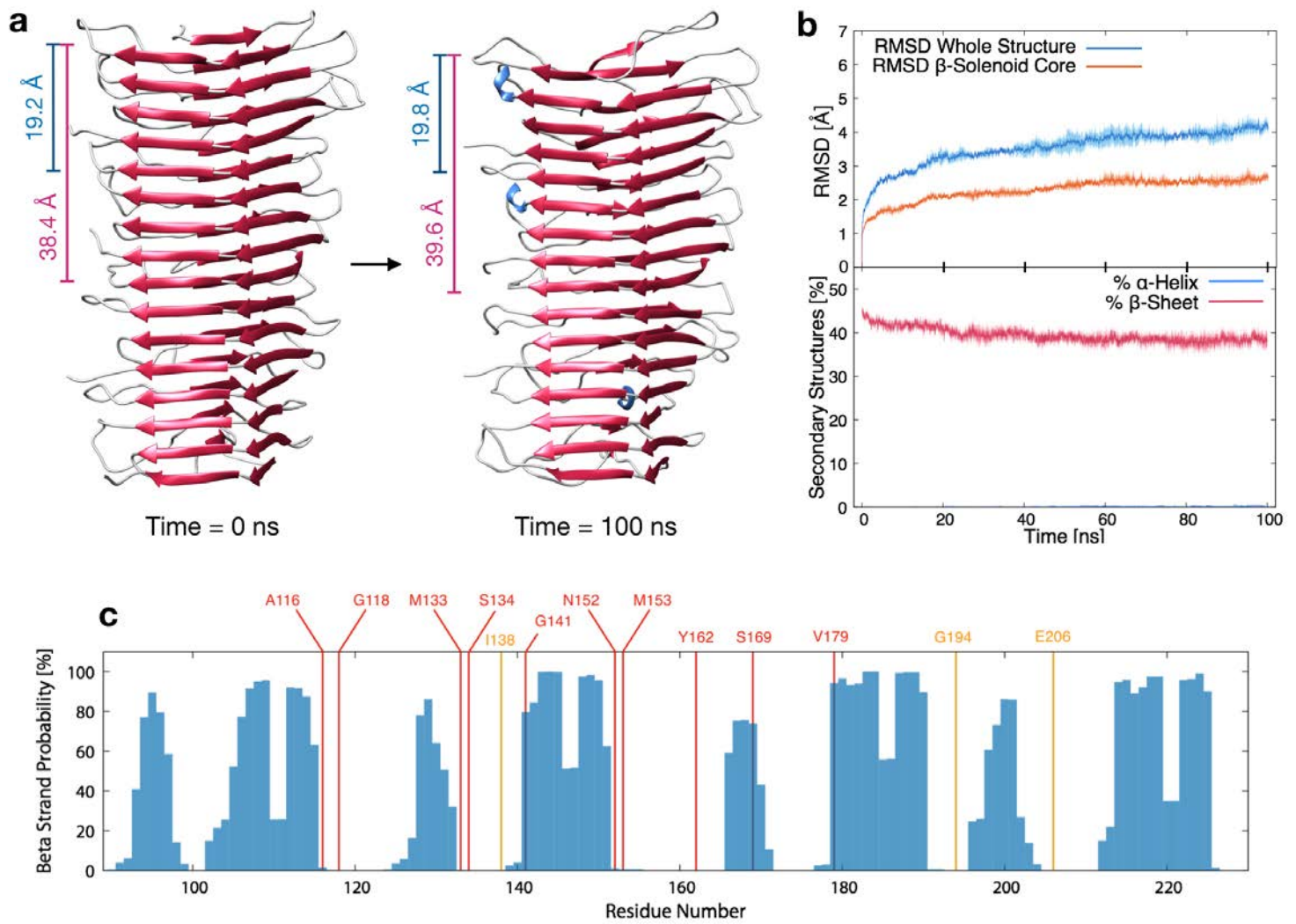
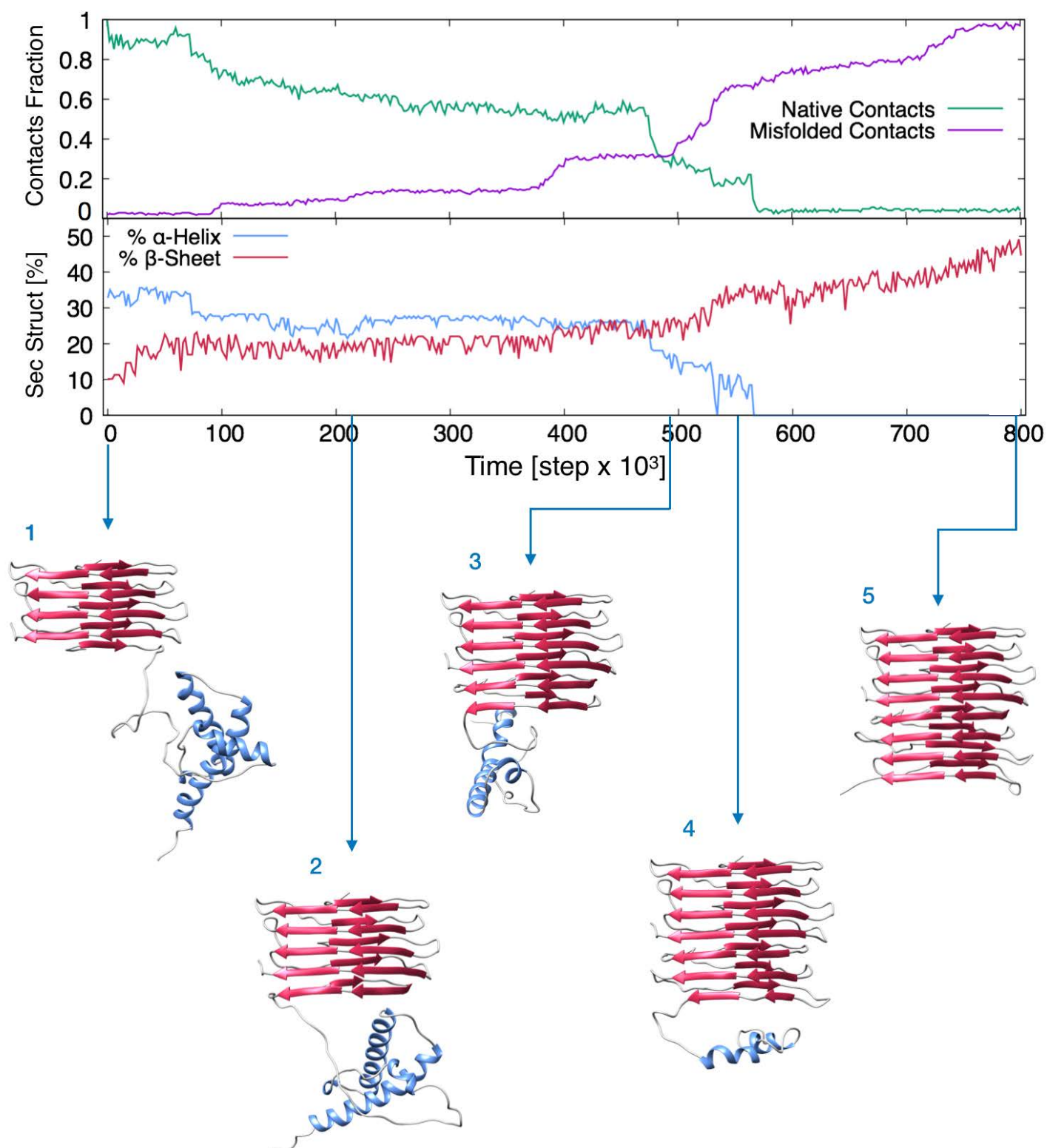
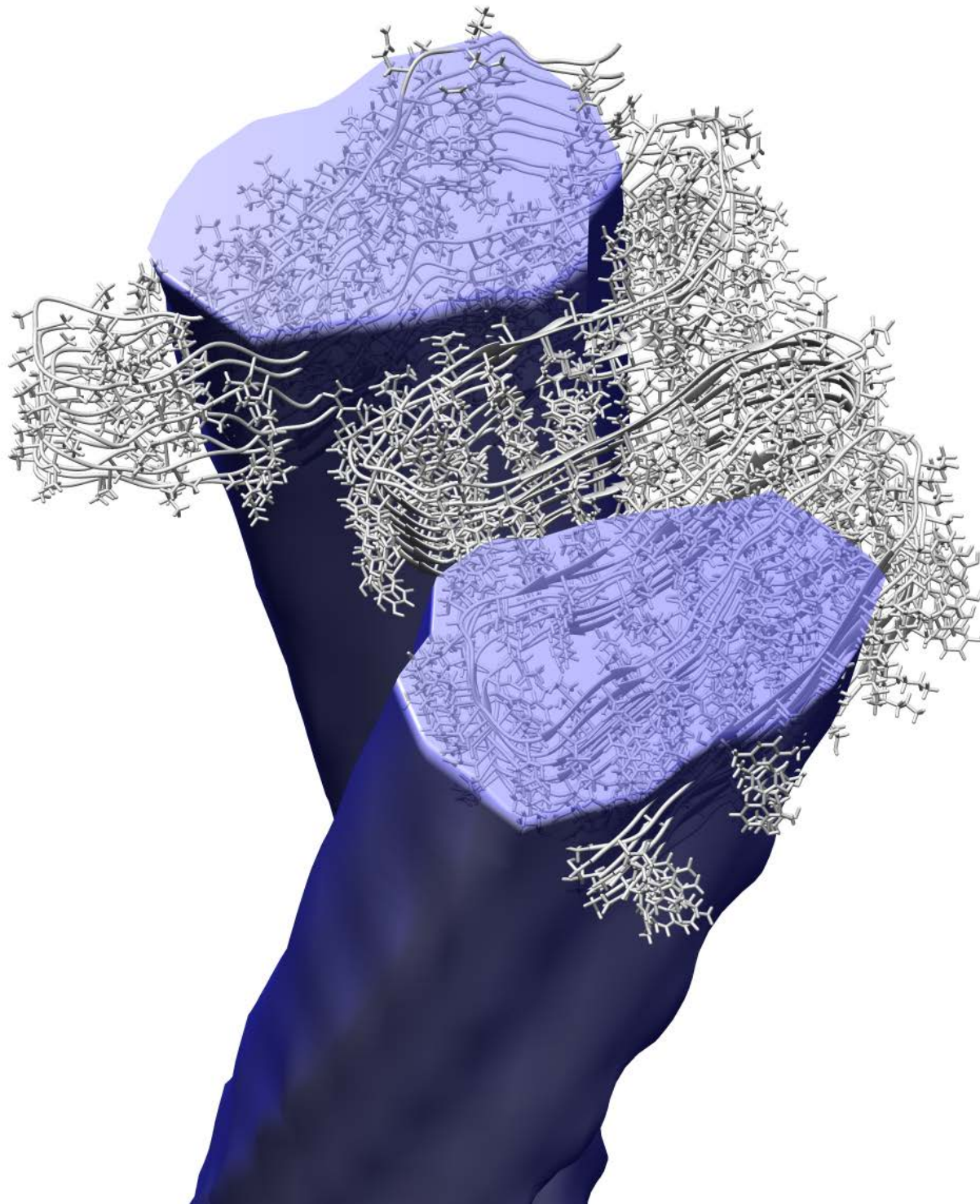




Figure 4

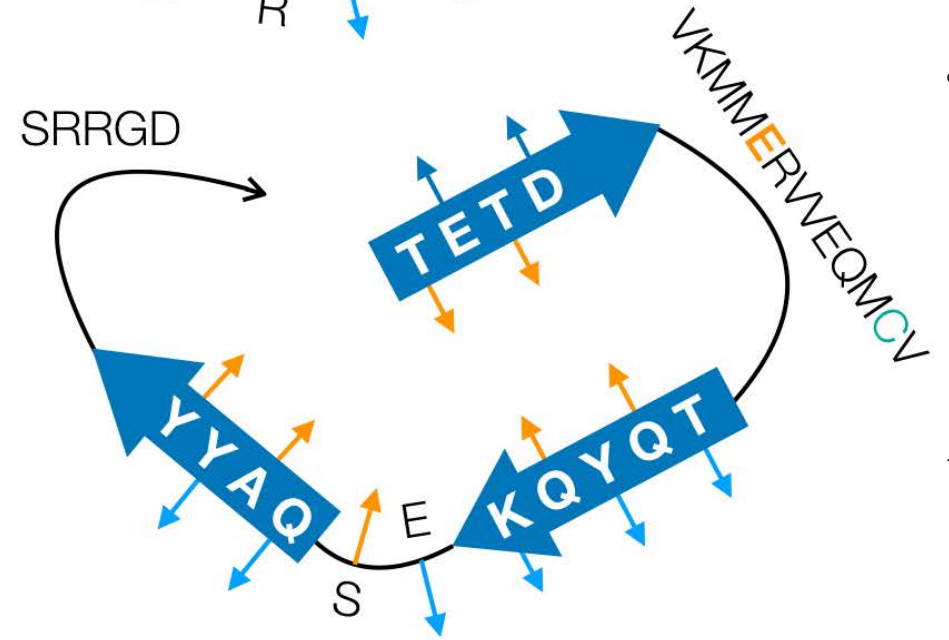
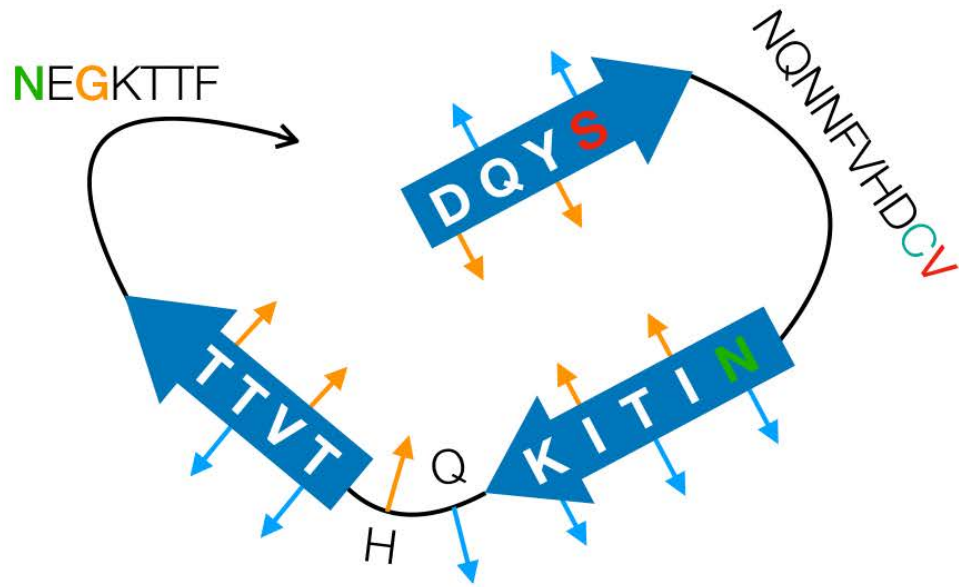
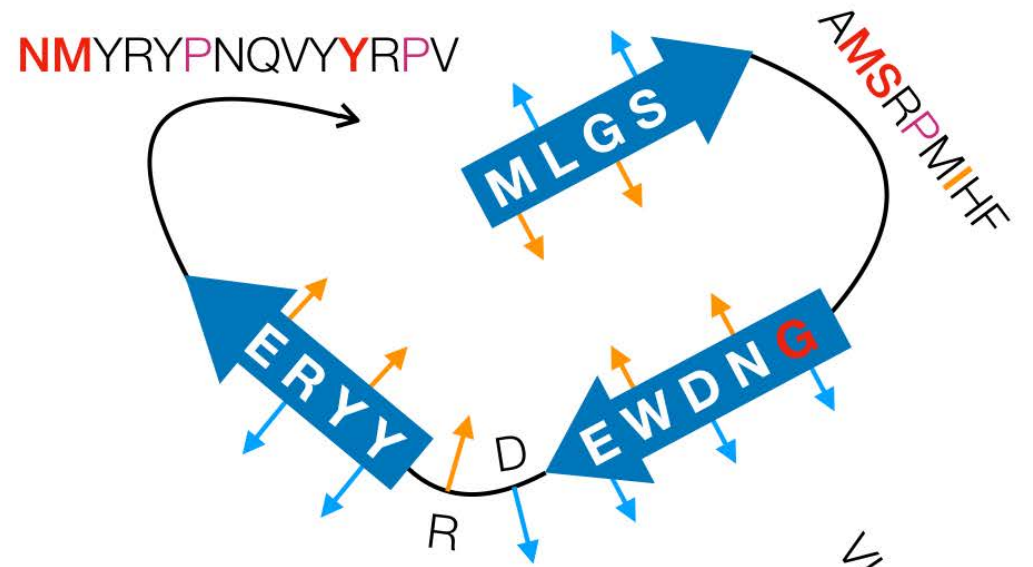
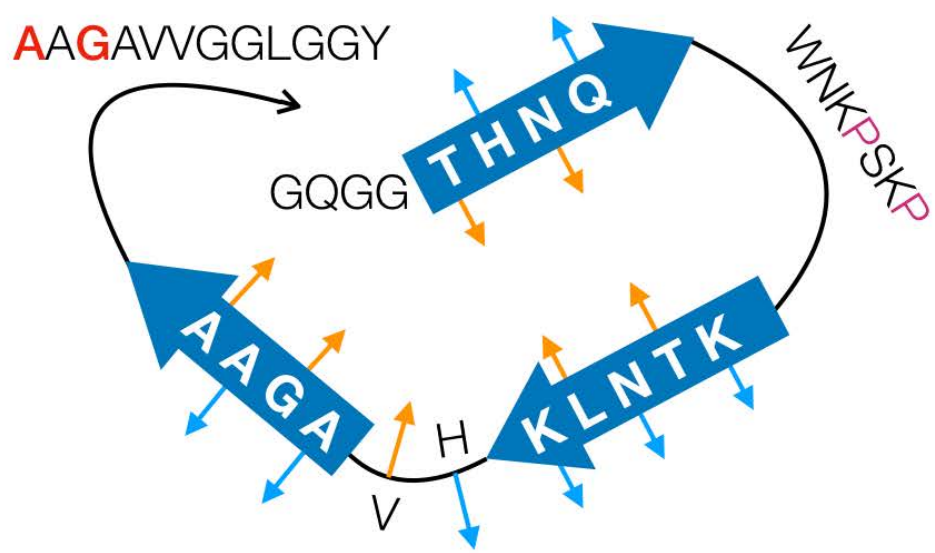


Supp Figure 1

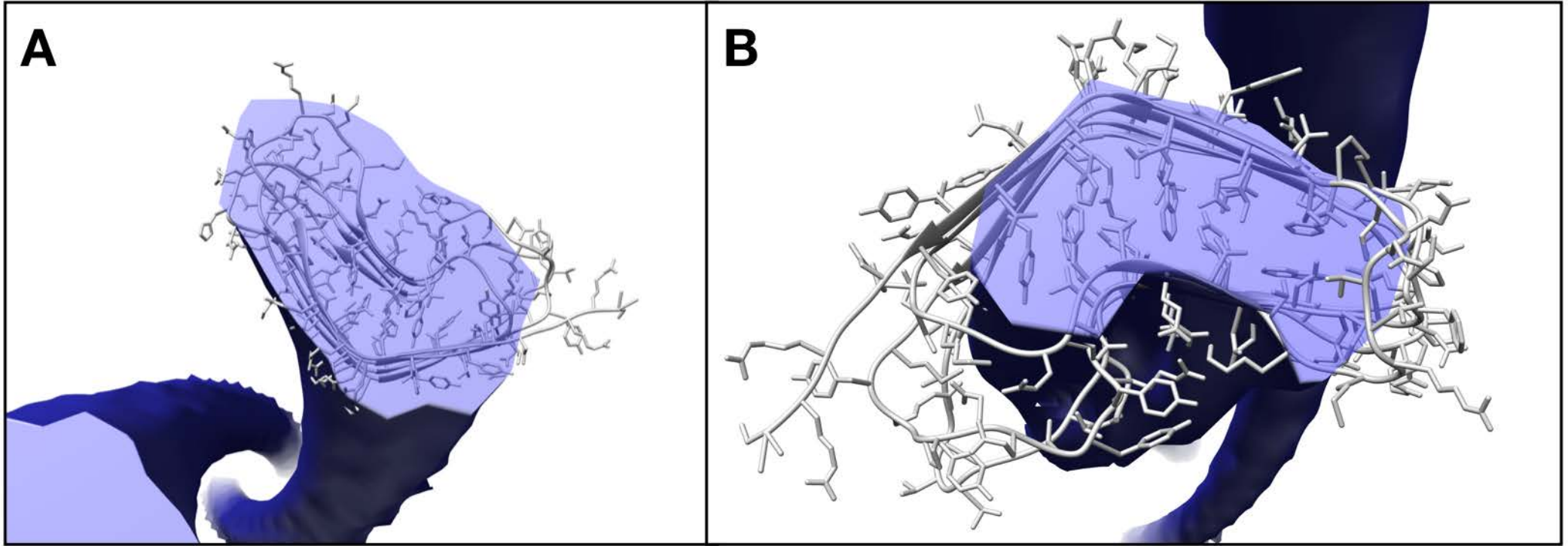




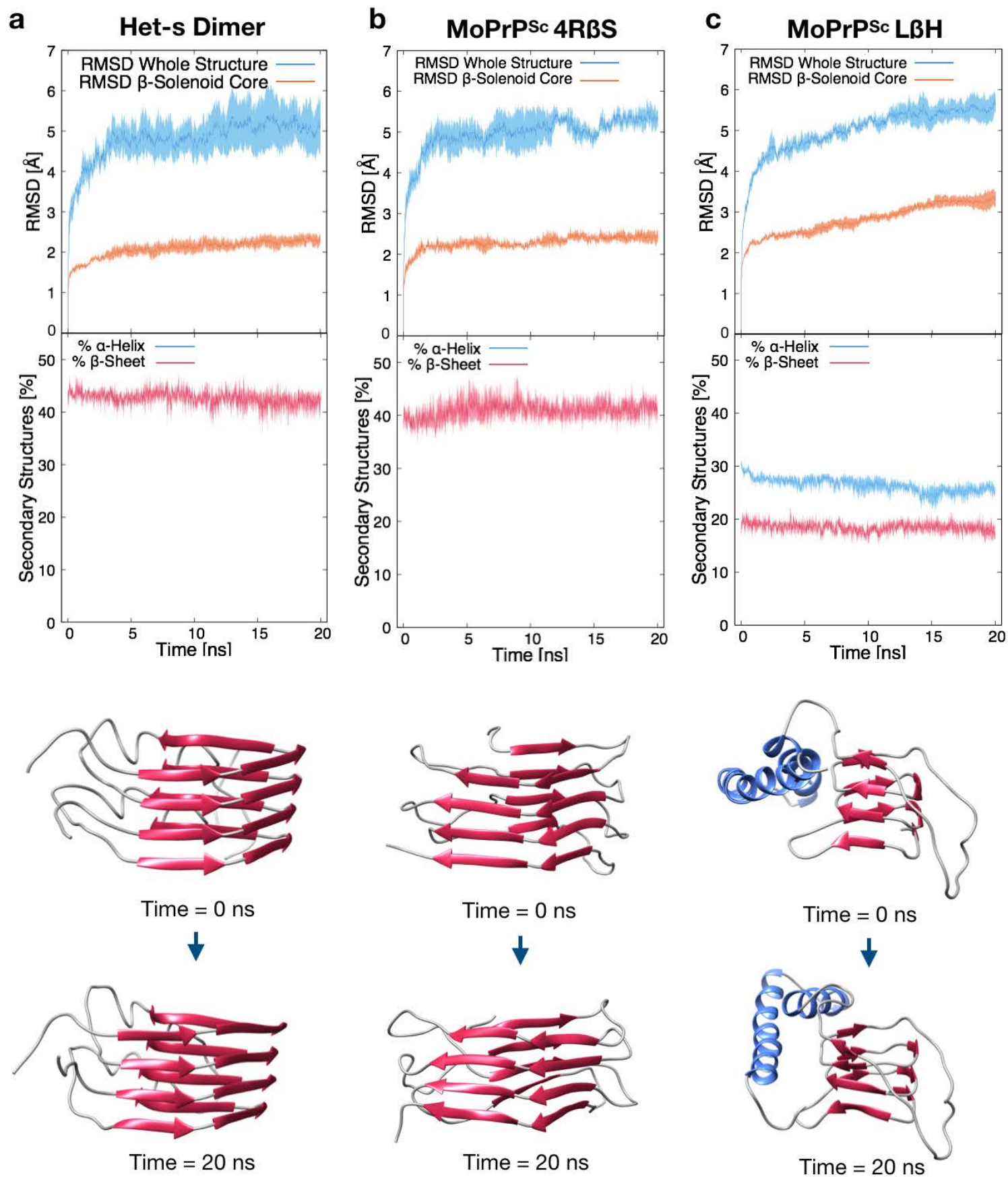
Supp Figure 2



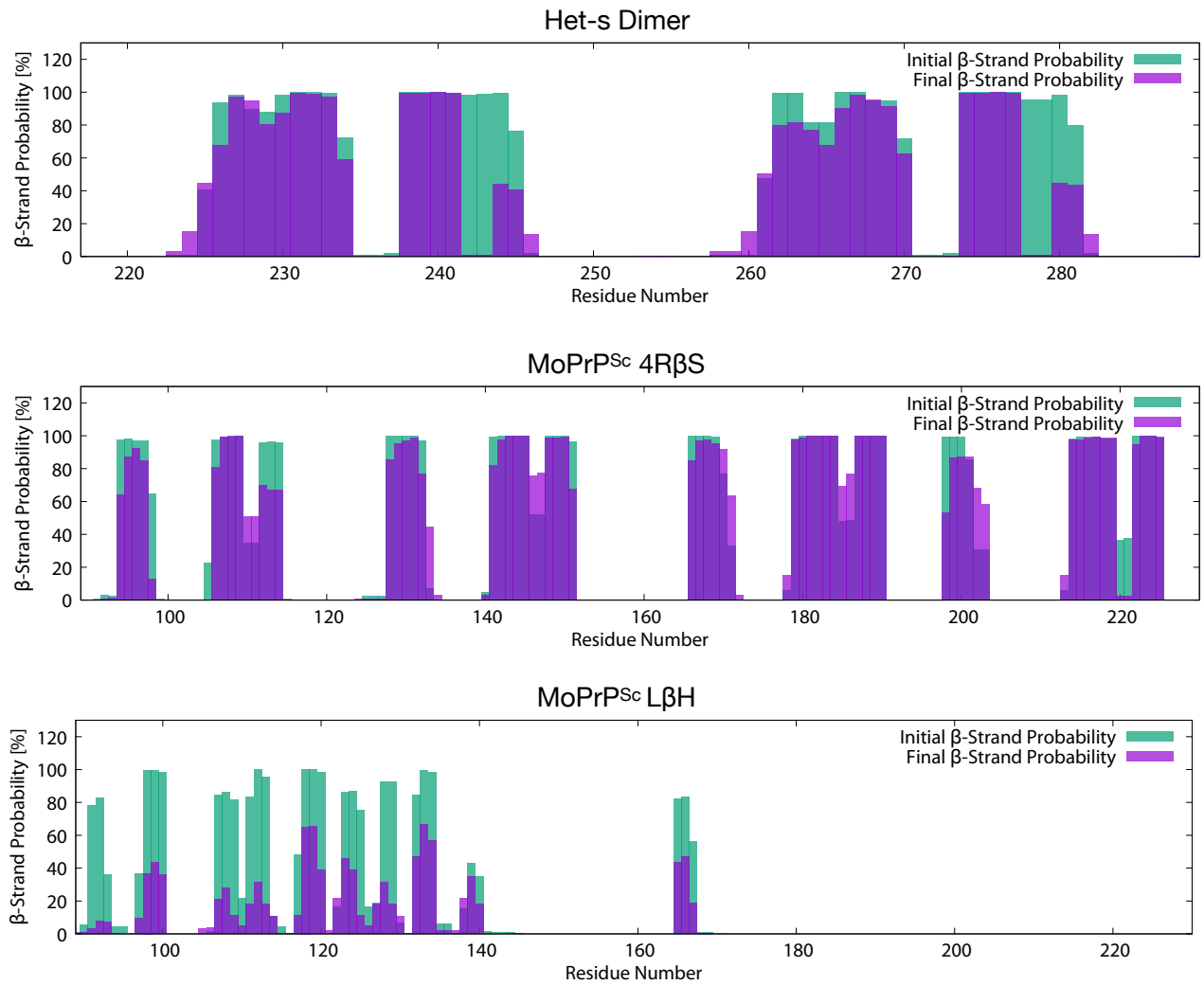
Supp Figure 3



Supp Figure 4

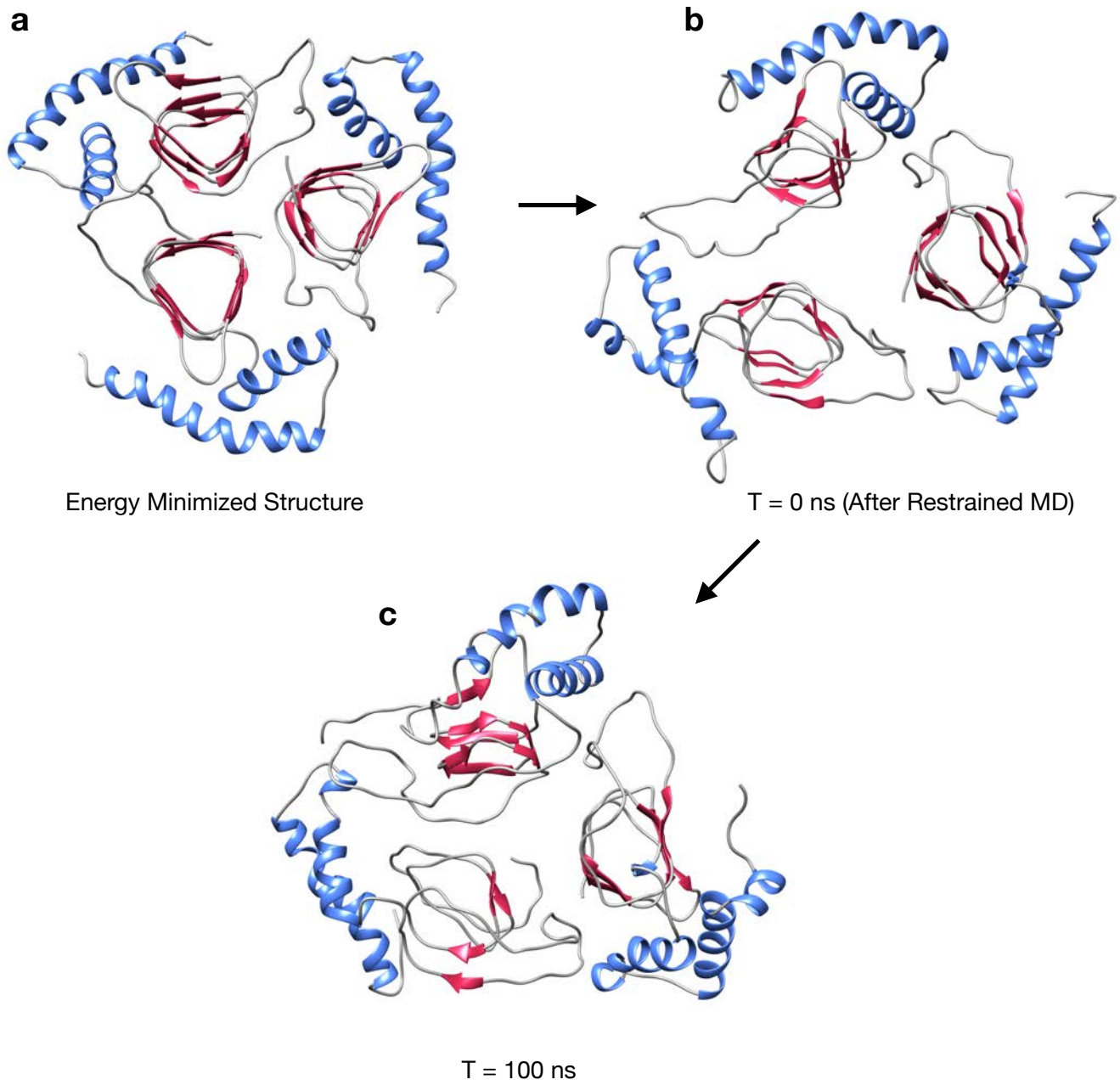


### Supp Figure 5

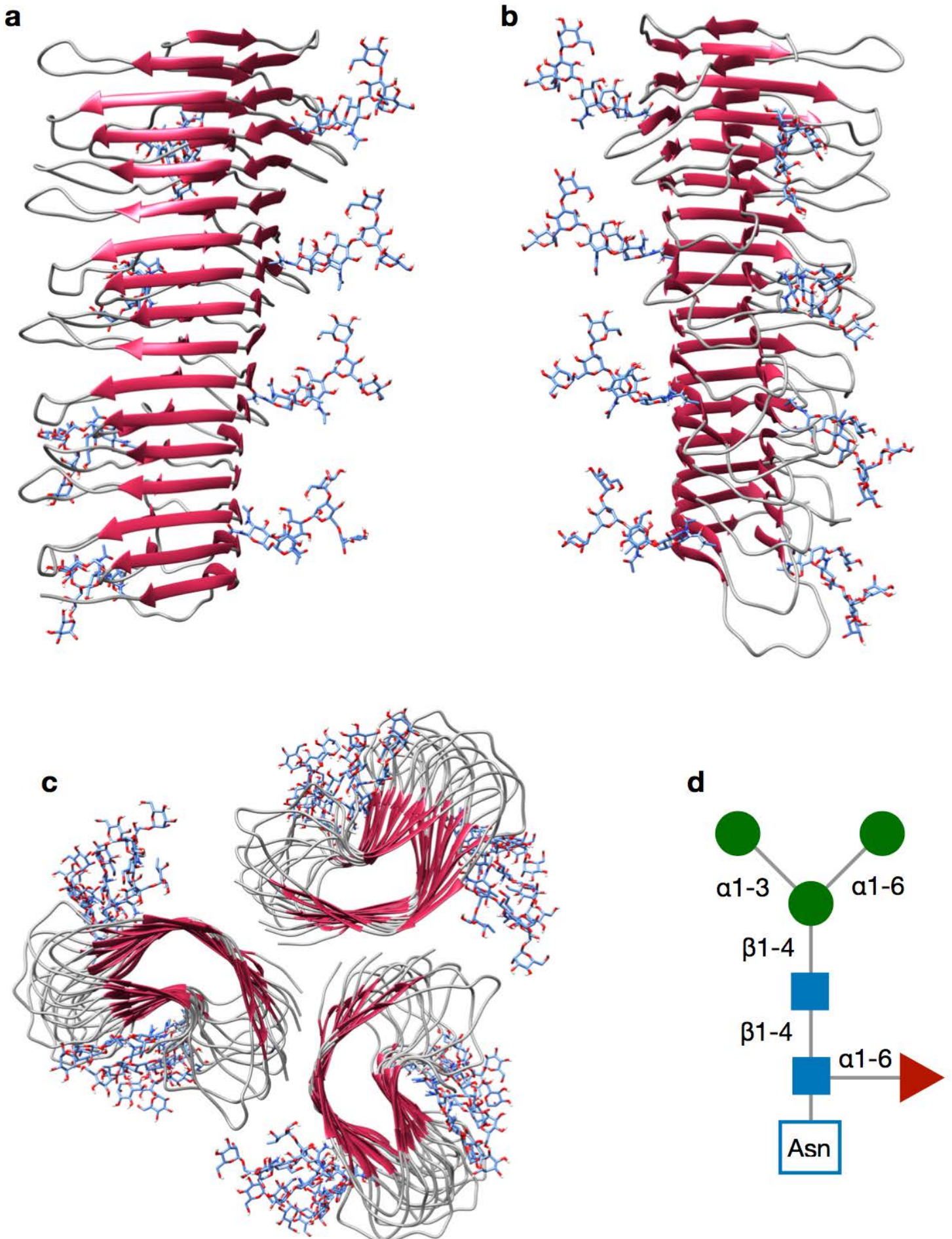




Supp Figure 6



Supp Figure 7





### Supp Fig 8

