BBA - Proteins and Proteomics 1865 (2017) 1631–1642

Contents lists available at ScienceDirect

BBA - Proteins and Proteomics

journal homepage: www.elsevier.com/locate/bbapap

The unfolding mechanism of monomeric mutant SOD1 by simulated force spectroscopy☆

Mona Habibi, Jörg Rottler, Steven S. Plotkin*

Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada

ARTICLE INFO

Keywords:
Protein misfolding
Molecular dynamics simulation
Superoxide dismutase 1
Amyotrophic lateral sclerosis
Single molecule force spectroscopy
Coarse-grained protein model

ABSTRACT

Mechanical unfolding of mutated apo, disulfide-reduced, monomeric superoxide dismutase 1 protein (SOD1) has been simulated via force spectroscopy techniques, using both an all-atom (AA), explicit solvent model and a coarse-grained heavy-atom Gō (HA-Gō) model. The HA-Gō model was implemented at two different pulling speeds for comparison. The most-common sequence of unfolding in the AA model agrees well with the most-common unfolding sequence of the HA-Gō model, when the same normalized pulling rate was used. Clustering of partially-native structures as the protein unfolds shows that the AA and HA-Gō models both exhibit a dominant pathway for early unfolding, which eventually bifurcates repeatedly to multiple branches after the protein is about half-unfolded. The force-extension curve exhibits multiple force drops, which are concomitant with jumps in the local interaction potential energy between specific β-strands in the protein. These sudden jumps in the potential energy coincide with the dissociation of specific pairs of β-strands, and thus intermediate unfolding events. The most common sequence of β-strand dissociation in the unfolding pathway of the AA model is β-strands 5, 4, 8, 7, 1, 2, then finally β-strands 3 and 6. The observation that β-strand 5 is among the first to unfold here, but the last to unfold in simulations of loop-truncated SOD1, could imply the existence of an evolutionary compensation mechanism, which would stabilize β-strands flanking long loops against their entropic penalty by strengthening intramolecular interactions. This article is part of a Special Issue entitled: Biophysics in Canada, edited by Lewis Kay, John Baenziger, Albert Berghuis and Peter Tieleman.

1. Introduction

Over 160 mutations throughout the homodimeric antioxidant protein Cu, Zn superoxide dismutase (SOD1) have been found to be associated with a familial form of amyotrophic lateral sclerosis (fALS), affecting about 1/5 of those with autosomal dominant inheritance [1,2]. Many of these mutations have been observed to show weakened dimer and/or folding stability [3–8], and accelerate fibril elongation (but not always fibril nucleation) [7,8]. In fALS cases, and in at least some of the more prevalent cases of sporadic ALS, patients display intraneuronal immunoreactivity to SOD1 misfolding-specific antibodies [9–11]. This suggests a potential fundamental role for the misfolding and propagation of SOD1 in the pathogenesis of at least some cases of ALS.

Many proteins implicated in misfolding-related degenerative disease form neurotoxic aggregates by intermolecular association of partially unfolded structures of monomer [12,13]. In the context of SOD1 misfolding related ALS, coarse-grained simulations of the unfolding pathway of WT and mutant SOD1 monomer exhibit multiple partially-unfolded intermediates that can determine fibril morphology [14]. The gain of transient oligomer interactions due to the partial disorder of the protein in the electrostatic loop region has been investigated computationally [15]. Thus, a molecular dissection of unfolding intermediate structures in more accurate, all-atom simulation models can provide insight into the conformational pathway involved in the seeding of multimeric oligomers that are involved in the infectious prion-like propagation [16,17] of this currently incurable disease.

SOD1 has a β barrel structure with two long loops dressing the core. Both of these loops have residues enabling metals to bind to the protein. A Cu ion imparts enzymatic activity to the protein to dismutate superoxide (O_2^-), while a Zn ion imparts thermodynamic and structural stability [18].

ALS-associated mutants of SOD1 have reduced affinity for metals [19,20], and loss of these metals renders the protein more susceptible to unfolding [21]. ALS mutations have the greatest destabilizing effect on the apo and reduced form of SOD1 leading to unfolding and oxidative aggregation [21] and monomerization [22].

Not all ALS-associated mutants show clear loss of metal affinity or reduced thermal stability for either holo or apo states in vitro [23]. However, in-cell NMR studies have shown that ALS-associated SOD1

☆ This article is part of a Special Issue entitled: Biophysics in Canada, edited by Lewis Kay, John Baenziger, Albert Berghuis and Peter Tieleman.

* Corresponding author.
E-mail address: steve@phas.ubc.ca (S.S. Plotkin).

http://dx.doi.org/10.1016/j.bbapap.2017.06.009
Received 24 April 2017; Received in revised form 9 June 2017; Accepted 9 June 2017
Available online 16 June 2017

mutants have significantly enhanced unfolded population and reduced metal affinity [24,25]. Studies of external challenges that induce unfolding of the apo state are thus interesting and potentially physiologically relevant. Because monomerization is a likely event in the disease pathway of ALS, we analyze here the molecular unfolding mechanism of apo, disulfide-reduced SOD1 (E,E(SH) SOD1). We find that this protein exhibits *partially folded states* that may partake in aberrant non-native intermolecular interactions.

Single molecule force spectroscopy experiments have been used to provide insight on the unfolding mechanism of SOD1 [26]; these experiments observed approximately 5 intermediates in the unfolding process, which could facilitate non-native interactions between unfolded or partially-folded monomers. Here we employ simulated force spectroscopy to elucidate the elusive intermediate structures on-route to unfolding. For this purpose we examine an obligate-monomer variant of SOD1 with partially-restored enzymatic activity, namely E133Q/C6A/C111S/F50E/G51E (PDB-ID 1RK7) [27]. Mutations C6A and C111S ablate potential intermolecular disulfide bonds, mutations F50E and G51E reside in the dimer interface and thus prevent dimer formation, and the mutation E133Q enhances superoxide binding rates without significant effects on SOD1 structure. Several previous studies have investigated the obligate SOD1 monomer through the mutants C6A/C111S/F50E/G51E [6,26,28-34]. In addition to being a more tractable system computationally [4], monomeric SOD1 is an important thermodynamic state on the pathway to oligomer formation and pathogenic aggregation [14,32,35,36]. The E133Q mutant on the pseudo-WT background enables additional analysis of enzymatic activity, does not significantly alter protein structure, and has also been studied previously [27,37-41,69].

2. Methods

2.1. Simulation models

We perform molecular dynamics to simulate the force-induced unfolding process of full-length monomeric SOD1 protein. We compare results for two different types of force fields: an all-atom (AA) model in explicit solvent [42,43], with a heavy-atom Gō model (HA-Gō) [44]. Note that in Gō models, only native interactions are attractive, while non-native interactions are purely repulsive. These models are described in detail below. Fig. 1 shows a representation of four amino acids in each of the models.

Force spectroscopy simulations are carried out by tethering both termini with a harmonic potential. The last residue (C-terminus) is then moved along the vector from C- to N-terminus with constant velocity. The stiffness of the spring that imparts the pulling force on the protein was set to 1000 kJ/(mol ·nm^2) for both AA and HA-Gō models.

All-atom pulling simulations were performed at velocity of v_o = 1 m/s at room temperature. Heavy-atom Gō (HA-Gō) simulations were performed at T = 100 K, which corresponds to 93% of the folding temperature. Pulling velocities of v_o = 1 m/s ("fast") and v_o = 0.1 m/s ("slow") were normalized across different computational models using

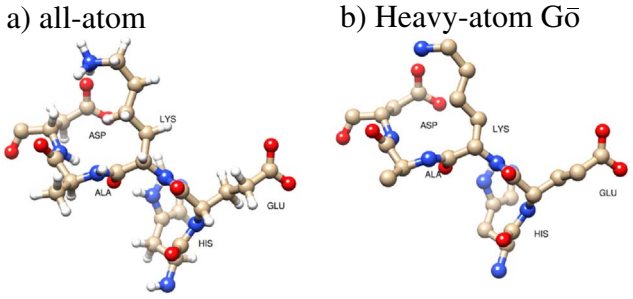

Fig. 1. Residues 120–124 in the structure of 1RK7 (HIS, GLU, LYS, ALA, and ASP) are shown in the all-atom model (a) and the Heavy atom-Gō model (b). The hydrogen atoms are not present in the HA-Gō model.

the method described in Habibi et al. [45]. Pulling simulations were repeated 10 times for the all-atom model, 100, 20 times for the fast, and slow pulling for the HA-Gō model, respectively.

2.1.1. All-atom (AA) model

We used the CHARMM22* force field [42] to model a monomer of SOD1 protein with TIP3P [46,47] water. AA simulations were carried out with the molecular dynamics code GROMACS-4.6 [48,49]. To obtain the initial configuration for the pulling simulation, the PDB structure was energy minimized and equilibrated for 20 ns in an isobaric ensemble (NPT) simulation with a salt concentration of 0.15 M. The average size of the simulation box is approximately $8 \times 8 \times 100\ \mathrm{nm}^3$ with 212,279 water molecules, 392 Na^+ ions, and 385 Cl^- ions. A time step of 2 fs was used with the LINCS algorithm [50]. All AA simulations were performed at constant temperature T = 300 K and pressure p = 1 atm. The temperature of the protein and the solvent were kept constant with two separate thermostats, as described in [51–53]. The velocity rescaling algorithm with a stochastic term was used as thermostat for both protein and solvent [54]. The pressure was kept constant using the Parrinello-Rahman algorithm with a weak coupling of 1 ps [55]. Lennard-Jones interactions (LJ) were truncated at 1.4 nm, and the particle-mesh Ewald method [56] was used for the electrostatic interactions. A representation of four amino acids in the AA model is shown in Fig. 1a).

2.1.2. Heavy atom-Gō model (HA-Gō)

In the HA-Gō model [44], all heavy atoms (non-hydrogen) are present and the potential function is only defined by the native state. A representation of four amino acids in the HA-Gō model is shown in Fig. 1b). Any two heavy atoms that are within a cut-off distance of 0.6 nm in the native state and are three or more residues apart are defined to form a native contact. In this system, the energy per contact for native interactions is $\epsilon_c = 0.4\ k_BT$. The interactions between these non-bonded atom pairs are modeled by a 6–12 Lennard Jones potential [44,57] and the separation corresponding to the potential energy minimum between pairs is set to the separation distance between pairs of atoms in the native PDB structure.

Pairs of atoms separated by a distance r that are not in contact in the native state are given a purely repulsive interaction of the form $U(r) = \sum_{non-native} \epsilon(2.5[\text{Å}]/r)^{12}$, with uniform values of $\epsilon = 0.01\ k_BT$ [44]. Bonded atoms are modeled by harmonic bond and angle potentials, along with dihedral potentials [44]. The HA-Gō simulations were carried out with GROMACS-4.5 [48].

GROMACS input files were generated from the PDB structure using the SMOG web server [58]. The time step was set at 2 fs. The simulations were performed at constant temperature of 100 K (see below) using a Langevin thermostat with time constant of 1 ps. The initial configuration of the pulling simulations was obtained after 1 ns equilibration at the desired temperature.

2.2. Native contacts

To compare mechanical unfolding pathway of SOD1 in the all-atom and HA-Gō model, we compute the number of native contacts of all configurations during each pulling simulation. We calculate native contacts using pairwise distances for all heavy atoms i and j in either AA or HA-Gō model for any protein structure. The fraction of the native contacts Q for any conformation α, $Q(\alpha)$, is defined as

$$Q(\alpha) = \frac{1}{|S|} \sum_{(i,j)\in S} \frac{1}{1+\exp\left[\beta^0(r_{ij}(\alpha)-\lambda r_{ij}^0)\right]}, \quad (1)$$

where $r_{ij}(\alpha)$ is the distance between heavy atoms i and j in conformation α, r_{ij}^0 is the distance between the corresponding heavy atoms i to j in the native state conformation, and S is the set of all pairs of native contacts (i,j) belonging to structure α. Amino acids having a native contact must be separated by four or more residues in the primary sequence, and $r_{ij}^0 < r_{cut}$ (r_{cut} = 0.48 nm) in the native state [59], $\beta^0 = 50\ \mathrm{nm}^{-1}$ is a

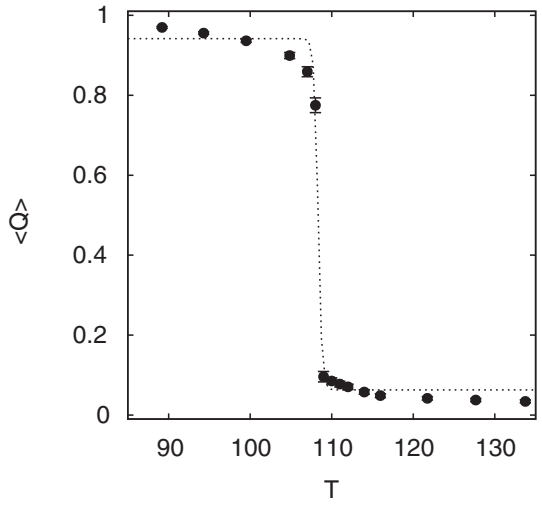

Fig. 2. Melting curve for HA-Gō. The transition region is fit to Eq. (3), the dotted lines show the fitted curve. For the HA-Gō model the melting temperature is T_f = 108 K. Error bars are obtained from the standard error in the data.

smoothing parameter and the factor λ = 1.8 takes into account contact fluctuations in the equilibrium ensemble. Since contacts are only among heavy atoms, both models have the same number of contacts for the same structure.

2.3. Non-native contacts

Any contacts formed during the simulations between heavy atoms *i* and *j* are considered a non-native contact if the distance between *i,j* in the PDB structure is larger than 0.48 nm. The total number of non-native contacts in configuration *α* is a function of the set of distances between all heavy atoms *i* and *j*, and is given by

$$N_{nn}(\alpha) = \sum_{(i,j)}{}' \frac{1}{1+\exp\left[\beta^0(r_{ij}(\alpha)-\lambda R_0)\right]}, \tag{2}$$

where R_0 = 0.24 nm is the mean of the distances between native atom pairs in the PDB structure: $R_0 = \langle r_{ij}^0 \rangle$. The smoothing parameter β_0 and the factor λ are the same as above for native contacts. The prime in the sum in Eq. (2) indicates that all possible contacts other than native contacts should be summed over. The ensemble of non-native contacts at a partial degree of nativeness Q is defined as those conformations within the bin ($Q - \delta Q, Q + \delta Q$, δQ = 0.002).

2.4. Normalization temperature/energy scales across models

The interpretation of "time" and "energy" in a coarse-grained (CG) model must be carefully considered. The smoother potential landscapes of CG models generally result in faster dynamics in comparison to all-atom forcefields. Therefore, the meaning of time in the HA-Gō model is not the same as time in the AA model.

To be able to compare the AA and HA-Gō at the same effective temperature, we wish to perform all simulations at 93% of the folding temperature T_f of the protein in each model. To this end, the thermal melting curves for the HA-Gō model is shown in Fig. 2 as a function of *T*. The melting curve is obtained from replica-exchange molecular dynamics (REMD) simulations [60] over 60 ns. Eq. (3) is used to fit the melting curve on the data in the transition region 90 K < *T* < 120 K.

$$\langle Q(T)\rangle = p_f\left(1-\frac{1}{1+\exp(\Delta G(T)/RT)}\right) + p_U\left(\frac{1}{1+\exp(\Delta G(T)/RT)}\right) \tag{3}$$

where *R* is the ideal gas constant and $\Delta G(T) = \Delta H - T\Delta S$. The melting temperature T_f is defined as the temperature where $\Delta G(T_f) = 0$. We assume here that the heat capacity is temperature-independent, an often-used assumption in calorimetry analyses, but one which is prone to extrapolation errors if the stability at other temperatures is desired [61]. Here, however, we are primarily interested in obtaining a good estimate for the melting temperature.

The melting temperature of the all-atom model is approximated as the experimental value for pseudo-WT SS-reduced, which is about 321 K [33], since the computational effort for performing either direct MD or REMD simulations on such a large protein in explicit solvent is prohibitive. Comparisons between experimental and computational melting temperatures by the Shaw group show good agreement for two all-*β* proteins WW domain and protein G [62]. The unfolding mechanism of the AA model has been previously shown by the authors to be relatively insensitive to changes in temperature between 290 K and 310 K [45].

2.5. Normalization of time scales across models

The rate of pulling in each model system depends on each system's internal time scale. Here, we follow the normalization scheme in Habibi et al. [45], wherein the relevant time scales for the AA and HA-Gō models were calculated based on the characteristic relaxation time in each model for loop-truncated SOD1 (tSOD1) [63]. tSOD1 has residues 49–81 and 124–139 replaced with GLY-ALA-GLY linkers but retains the 8-stranded *β* barrel, and has mutations C6A/C111S to avoid intermolecular disulfide bonds.

We use the same timescales as derived previously in tSOD1 for the AA and HA-Gō model [45]. In summary, time-scales and thus pulling rates were normalized across the AA and HA-Gō models by rescaling the time by the characteristic relaxation time in each model, as obtained from the mean relaxation time after mechanically perturbing the system. The mean relaxation time is calculated from the decay of the autocorrelation function for the fraction of native contacts, $\langle (Q(t) - \overline{Q})(Q(0) - \overline{Q})\rangle$, which is fit to a double exponential decay $A_1 \exp(-\kappa_1 t) + A_2 \exp(-\kappa_2 t)$. To normalize pulling rates, we take the relevant time-scale in each model to be the inverse of the slower relaxation rate $t_{CG} = \langle \kappa_1 \rangle^{-1}$ to the folded state.

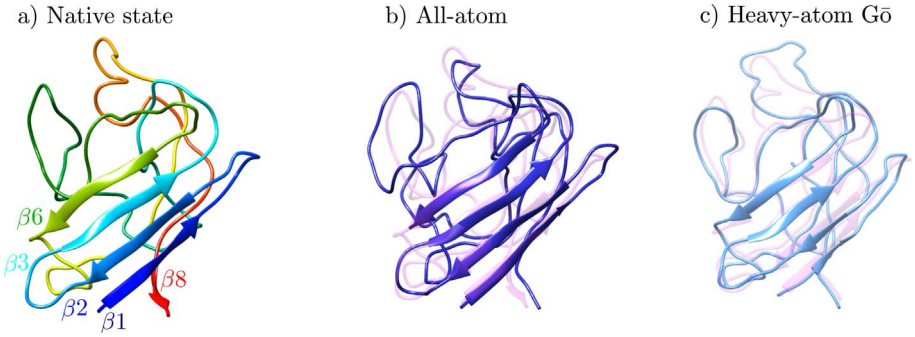

Fig. 3. Representative structures of E,E(SH) SOD1. a) Model 1 of PDB ID 1RK7 shows a partially-disordered native state of the protein (compared to the holo dimeric structure). The structure is color coded by primary sequence: blue to red corresponds to N to C terminus. b) Representative equilibrium structure taken from the last structure frame of the simulation of SOD1 after 20 ns equilibration in explicit solvent (blue), superimposed on the structure in panel a) (pink). c) Representative equilibrium structure of the protein in the HA-Gō model (cyan), again superimposed on the structure of panel a).

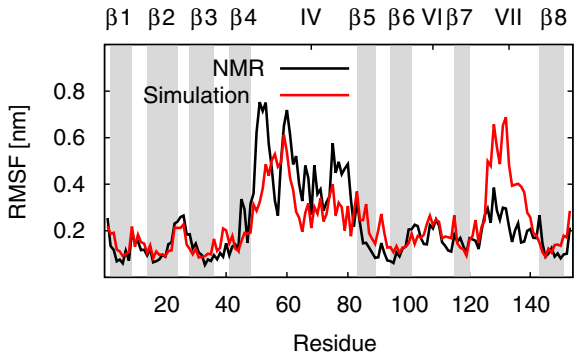

Fig. 4. RMSF of 1RK7 from all-atom equilibrium simulations (red) of the native structure, and for 30 structural models (black). The shaded regions show the β-strands. The labels for the shaded regions and for loops IV and VII are at the midpoints of each secondary structure element.

2.6. Unfolding taxonomy

To determine whether there exists a well-defined unfolding pathway for SOD1, and test its robustness between the AA and HA-Gō models, we use the template modeling score (TM-score) [64] to compare the similarity between the protein structures of different pulling trajectories at the same Q. The TM-score for the alignment of two structures is defined as:

$$\mathrm{TM}=\frac{1}{L}\sum^{N}\frac{1}{1+\left(\frac{d_i}{d}\right)^2},\ d=1.24\sqrt[3]{(N-15)}-1.8, \tag{4}$$

where N is the number of residue pairs, d_i is the distance between identical residues i in two structures, and L is the number of the residues in the reference structure. The TM-score lies between zero and one; a TM-score of one indicates that the two protein structures are perfectly matched. Usually, two structures with TM-score higher than 0.5 are considered to have essentially the same conformation, while uncorrelated protein structures have a TM-score of less than 0.2 [64]. Measuring the TM-alignment, as well as clustering of structures by TM-score, was performed by using Maxcluster (http://www.sbg.bio.ic.ac.uk/maxcluster) [65].

2.7. Calculating β-sheet content

β-strands are identified by the DSSP method of Kabsch & Sander [66], using the program Chimera [67]. Residues with hydrogen-bond energy cutoff of $-0.5\ \mathrm{kcal.mol^{-1}}$ and with minimum strand length of 3 are rendered as β-strands. Renderings of the secondary structures as ribbons (see Fig. 7) are made using JPred4 (A Protein Secondary Structure Prediction Server) [68] for PDB structures 1RK7 and 1HL5. For the schematic ribbon-plots of various partially-folded structures, we rendered these in Chimera [67] using a hydrogen bond cut-off for β strand identification of $-0.5\ \mathrm{kcal.mol^{-1}}$ in the AA model, and $-0.12\ \mathrm{kcal.mol^{-1}}$ in the HA-Gō model. The different cut-offs were used to best represent the apparent β sheet content based on visual inspection of β-strands and Ramachandran angles, given that the HA-Gō model does not contain hydrogen atoms.

3. Results

3.1. Mechanical unfolding of SOD1 (1RK7)

Fig. 3a) shows a representative structure of the 1RK7 protein in the native state. The last structure of the equilibrium MD trajectory of this protein in the AA and HA-Gō models are shown in Fig. 3 b) and c), respectively. After 20 ns of equilibration in the AA simulation, residues in β-strands 4 and 7, the C-terminus of β8, loop IV (residues 42–85), and loop VII (residues 121–148) have significantly increased in disorder. Overall, the all-atom protein loses 25% of its native contacts, so in

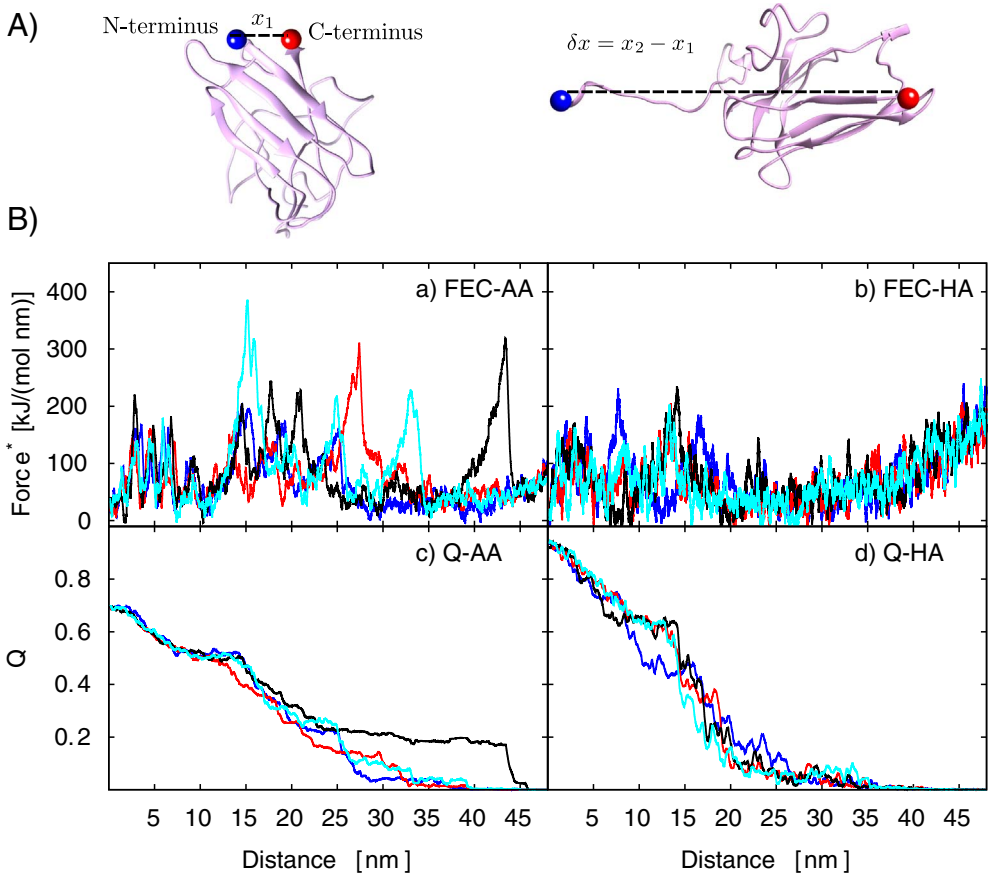

Fig. 5. A) Snapshots of the protein initially before extension, and at 10 nm extension. x_1 is the distance between N- and C- termini at the native state. $\delta x = x_2 - x_1$ is the reported distance value during pulling. B) Force and Q *vs.* distance for 4 pulling simulations of the AA and HA-Gō models. Each significant force drop in the force-extension curve (FEC) corresponds to an unfolding event. HA-Gō force values in panel b) are shown 10 times larger than they actually are. The notation Force* indicates that the force is not rescaled across the different models (see ref. [45] where force renormalization is considered).

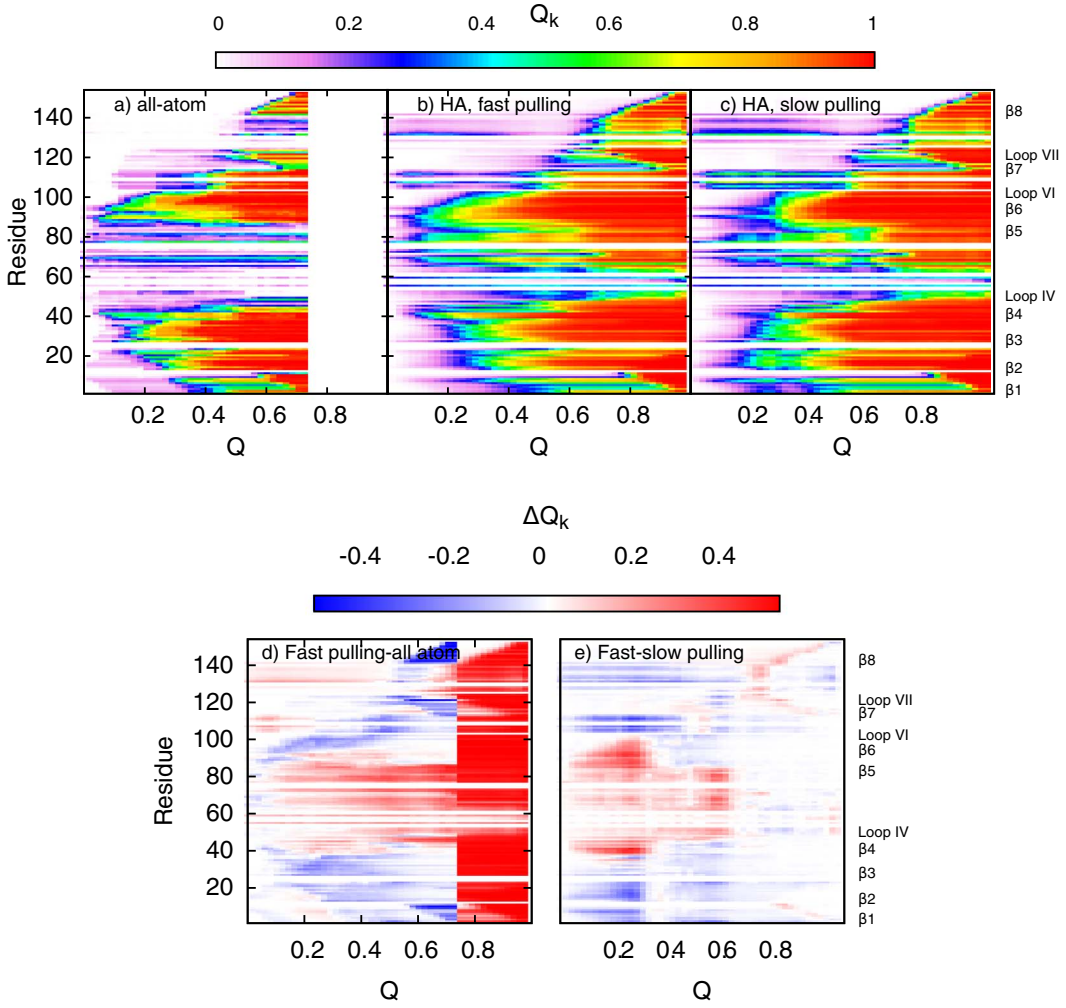

Fig. 6. Panels a)–c) show the fraction of native contacts for each residue $Q_k(Q)$ vs. total number of native contacts Q. Red color indicates the presence of nearly all native contacts, while white represents a residue that has essentially no native contacts remaining. Panel a) AA model results, averaged over 10 simulations, at a pulling rate of 1 m/s. Panel b) HA-Gō model results averaged over 100 simulations runs, at a pulling rate of 1 m/s. Panel c) HA-Gō simulations averaged over 20 simulations, at a slower pulling speed of 0.1 m/s. The AA model in panel a) only has data for $Q < 0.75$, since about 25% of the native contacts present in the PDB structure are lost upon equilibration; for the HA-Gō model this number is only about 5%. Panel d) shows the difference between $Q_k(Q)$ in the HA-Gō model and the AA model. Panel e) shows difference between the fast-pulling and slow-pulling HA-Gō model. Residue contacts that are enhanced in the HA-Gō model relative to the AA model or slower pulling simulations are colored red in panels d)–e). Residue contacts that are reduced are colored blue in panels d)–e).

native equilibrium $Q = 0.75$ (Fig. 3b). In the HA-Gō simulations, the native contacts in the equilibrium structure of the protein are mostly intact, $Q = 0.95$ (Fig. 3c).

The 25% of total contacts lost during equilibration in the AA model are primarily in the loop region. The entropic penalty of the disordered loop IV may induce disorder in β5 [70]. The behavior of loops IV and VII in the AA simulations is consistent with observations in the NMR solution structure of 1RK7 [27], and other structural studies of apo SOD1 monomer [34,70-73].

In Fig. 4, we plot the root mean-squared fluctuations (RMSF) of 1RK7 taken from both our equilibrium simulations (red) of the native structure, and for the 30 structural models in PDB 1RK7 (black). We can see that both computationally and experimentally, loops IV and VII are highly dynamic and flexible, retaining little native structure.

The force extension curves and fraction of native contacts (Q) for four runs of HA-Gō simulations are shown in Fig. 5. It is evident that the force extension curves (FECs) from these stochastic trajectories vary widely. Even within the same model, we observe variability in the FECs for different trajectories. FECs by themselves do not provide explicit information regarding the intermediate structures of the protein; in simulations we thus look at various metrics such as residue contacts and partially-folded structures to analyze the unfolding mechanism.

3.2. Primary sequence and temporal resolution of the loss of residue contacts upon unfolding

3.2.1. Native structure "fireplots"

To determine the sequence of the unfolding both spatially and temporally, we monitored the number of native contacts for each residue during the pulling simulations. Fig. 6a)–c) plots the average fraction $Q_k(Q)$ of a given residue k ($1 \leq k \leq 153$) as a function of total Q for the AA model and the HA-Gō (fast and slow pulling), respectively. To calculate $Q_k(Q)$, we normalize the number of contacts that residue k possesses at Q by the number of contacts that residue k possesses in the native structure where $Q = 1$. Red color in the figure corresponds to $Q_k(Q) \approx 1$, and white color indicates $Q_k(Q) \approx 0$, i.e. the residue has lost essentially all of its native contacts.

Panel a) of Fig. 6 shows that the disorder present in loop IV of the apo protein induces β5 to lose structure. The loss of structure between amino acids 51–84 and 126–138 is distinctly enhanced in the AA model

compared to the HA-Gō model (Panels b) and c) of Fig. 6). As well, β-strands 7 and 8 show less persistent structure in the AA model than in the HA-Gō model. Fig. 6c) shows that reducing the pulling speed by an order of magnitude in the HA-Gō model, from 1 m/s to 0.1 m/s, has apparently modest effects on the unfolding mechanism.

To quantify this further, Fig. 6d)–e) show *differences* in contact formation probability rendered in Fig. 6a)–c). Fig. 6d) shows the difference in native contact formation between the AA and HA-Gō models (both at a pulling rate of 1 m/s), i.e. the difference between panels a) and b) in Fig. 6. The dark red region for $Q > 0.75$ corresponds to the increase in native structure for the HA-Gō model relative to the AA model in this region. An increase in native structure of the HA-Gō model relative to the AA model persists over most of the range of Q for residues 40 to 90, a region corresponding to β4, loop IV (the Zn binding loop), and β5. Interestingly, there is a blue "halo" of decreased native structure for other parts of the protein, indicating that the native structure that is present in the AA model is generally more persistent than the native structure in the HA-Gō model.

Fig. 6e) shows the difference in native contact formation between the HA-Gō model at a pulling rate of 1 m/s, vs 0.1 m/s, i.e. the difference between panels b) and c) in Fig. 6. The effects are small. The region consisting of residues 40 to 90 retains slightly more native structure, while residues outside this region lose native structure more readily as the protein is unfolded more quickly.

3.2.2. Native structure retention maps

Another method for depicting the unfolding mechanism is shown in Fig. 7. Here the color scheme represents the time sequence (or equivalently extension) at which residues lose more than 50% of their native contacts during unfolding. The most persistent residues are colored dark blue, and the residues that lose their native contacts first (at lowest extension) are colored white. At the top of the figure are two secondary structure maps: one for the β sheet content in the holo, disulfide bonded dimer PDB (1HL5) (cyan), and another for the β sheet content present in the equilibrated apo SS-unbonded monomer (purple) treated here. β-sheet content is calculated as described in the Methods section.

As can be seen in Figs. 6 and 7, the AA model predicts large disordered regions involving loop IV and what was β5, as well as loop VII and what was β8; these are indicated as white regions in the corresponding panel of Fig. 7.

One discrepancy of the models is apparent: While β5 is already dissociated during equilibrium of the AA model, in the HA-Gō model at fast pulling rate, β5 is the 4th strand to unfold sequentially. The order that the strands are most likely to be dissociated in each model are given in Table 1. It is apparent from Table 1 that β4 (specifically the C-terminal half) and β5, flanking the long Zn-binding loop IV in the apo protein, are the first to unfold in the AA model, but are among the last to unfold in the loop-truncated AA SOD1 model. The implications of this dramatic shift in unfolding mechanism upon truncation of the loop is discussed further in the Conclusions.

In Table 1, last two β-strands are put in parentheses in each model, since β-strands require a partner to dissociate from; i.e. It is not sensible to say that a single β-strand dissociates last.

The above mechanisms indicate that, as a general rule, the C-terminus of the protein is the first to unfold in the force spectroscopy assay, followed by the N-terminus. The above mechanisms are based on the average over all unfolding trajectories, and so constitute general trends rather than a mandatory mechanism. We will see explicitly below (cf. Fig. 10) that there are multiple pathways for unfolding.

The most stable core of the protein also appears to be largely common to both models. This core involves β-strands 2, 3, and 6 (residues 17–36 and 90–103), which together form a β-sheet in the native structure (see Fig. 3). This result is consistent with the observed stable core of the protein during chemical denaturation [40], which consisted of β-strands 1, 2, 3, and 6—the last 4 strands to unfold in our AA model. The order that the strands in this core dissociate depends on the pulling speed more sensitively than the model: For both AA and HA-Gō models, β2 dissociates first, then β3 and β6 dissociate afterwards, while in the HA-Gō model at slow pulling, β6 dissociates first, followed by β2 and β3. The enhanced stability of β5 in the HA-Gō model is due to the bias towards the native state, while the destabilizing effect of the loops is reduced.

The similarity between unfolding events depicted in Fig. 7 may be quantified by computing the correlation coefficient between the degree of remaining structure for individual β-strands (the similarity of the darkness of the bands for each model in Fig. 7). This gives a correlation coefficient of 0.91 between the AA and HA-Gō models. The correlation coefficient is

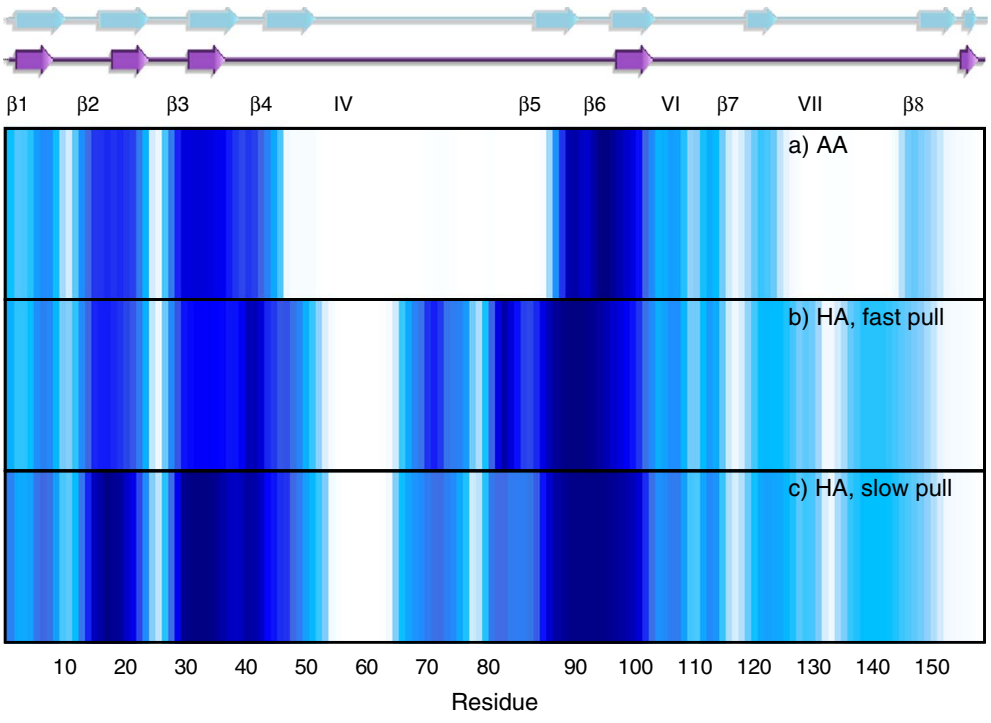

Fig. 7. The loss of native structure as the protein is mechanically unfolded. The color scheme represents the value of Q where each residue loses more than 50% of its contacts ($Q_k < 0.5$) first. The white color shows the least stable residues and the dark blue represents the most persistent residues. The schematic at the top shows the secondary β structure in 1HL5 (cyan) and 1RK7 (purple).

calculated here only for residues that are ordered in the equilibrated structures of both the AA and HA-Gō models, i.e. residues 1–51, 85–125, and 139–153. The correlation coefficient when all the residues are included is 0.68, $p < 10^{-21}$. The correlation coefficient between slow pulling rate and fast pulling rate for all the residues in the HA-Gō model is 0.95.

The native structural persistence of each residue may be mapped onto the PDB structure to illustrate which parts of the protein are most or least robust to unfolding. This is shown for both the AA and HA-Gō models in Fig. 8. It is noteworthy that both models predict the same native protein core to be most resilient against forced unfolding, involving β-strands 2, 3 and 6.

3.3. Native and non-native contacts in the partially-unfolded ensemble

Fig. 9 shows the contact-map for both the AA model and HA-Gō at half-nativeness ($Q = 0.5$). The probability of formation of native contacts is shown in the upper left triangle, and the fraction of conformations having a given non-native contact is shown in the bottom right triangle. Both ensembles are chosen from those configurations within the bin $Q - \delta Q$, $Q + \delta Q$, where $Q = 0.5$ and $\delta Q = 0.002$. The HA-Gō contacts are averaged over 20 (out of the 100) pulling trajectories, while the AA contacts are averaged over 10 trajectories. Five frames from each trajectory within the bin ($Q - \delta Q$, $Q + \delta Q$) were chosen, giving 50 frames total for the AA ensemble and 100 frames total for the HA-Gō ensemble. In the native contact-map, the color scheme is defined as red if nearly all native contacts are present between residues k,l in nearly all conformations of the ensemble at $Q = 0.5$.

The native contact maps are approximately the same for both models. As well, the non-native contacts in both models are, to first approximation, very similar to the native contacts. This is because non-native contacts consist largely of two native-like components [45]:

1.) “Near-native contacts”, which are between atoms not in contact in the PDB structure, but between amino acids that have other contacting atoms in the PDB structure.
2.) “Off-native contacts”, which are contacts between atoms in amino acids that are shifted by about one amino acid from interacting amino acids in the PDB structure. That is, if amino acids (α,β) are in contact in the PDB structure, these contacts are between ($\alpha \pm 1$,β), or (α,$\beta \pm 1$). These contacts are a consequence of taking one PDB structure to determine the native contact map, along with the presence of structural fluctuations in the native ensemble at room temperature. Structural determinations at cryogenic temperatures tend to be smaller, overpacked, and unrealistically unique. Near-native and off-native contacts are consistent with observations of room-temperature crystallographic ensembles, which undergo substantial conformational redistribution from the cryogenic ensemble for about a third of the side chains [74,75].

The non-native contact map for the HA-Gō model is sparser than the AA contact map, and non-native contacts are made with lower probability. Moreover, in the AA simulation, there are non-native contacts between residues 110–120 and residues 50–40, corresponding to an antiparallel β sheet between β7 and β4, which is not present in the ensemble at $Q = 0.5$, and is also not present in the native PDB structure 1RK7. These contacts are present however in the holo, disulfide bonded structure (PDB ID 1HL5).

3.4. Unfolding pathways and unfolding taxonomy

In order to render the unfolding pathways predicted by each model, we clustered the protein conformations based on the TM-scores during the unfolding at several different Q-values, see Fig. 10. The structures shown are centroids of the corresponding clusters that emerge from the clustering analysis. A TM-score cut-off of 0.5 is used to define when configurations no longer belong to a given cluster. The coloring is based on the residue index, where the C-terminus of the structured protein is in red and the N-terminus is colored blue. The thickness of the lines in the unfolding pathway tree is proportional to the fraction of total trajectories that fall into each cluster.

To compare conformations at various stages of unfolding, we considered C_α-atoms in the backbone for the remaining folded region of the protein. This folded region at each Q-value was defined as the sequence of n residues with residue index $i \le j \le i + n$, where the end points i and $i + n$ are defined as having sufficient degree of native structure: $\langle Q_i(Q)\rangle > 0.5$ and $\langle Q_{i+n}(Q)\rangle > 0.5$. The average $\langle\cdots\rangle$ corresponds to the ensemble of states within a window centered at Q, such that $Q - 0.01 < Q < Q + 0.01$ taken from all pulling trajectories. For the AA, HA-Gō, and HA-Gō slow pulling trajectories, TM-score clustering was performed over 50, 100, and 100 conformations respectively, chosen from a window centered at the same value of Q in each case.

Since loop IV and VII are disordered in the early stage of the unfolding, these loops (residue index of 51–84 and 126–139) are not included in the conformation of the folded segments for any structures. The TM-score was thus calculated with the gap penalty of zero.

All models in Fig. 10 show a dominant pathway between $Q = 0.75$ and $Q = 0.5$, denoted by the thick black horizontal line connecting the structures in the respective panels. The fraction of the trajectories participating in this pathway in each model are 1, 0.99, and 1 for panels a), b), and c), respectively. Along this pathway, β8 at the C-terminus loses structure. In the HA-Gō model at fast pulling rate, in one trajectory out of 100, the N-terminus unfolds first. For the AA model one dominant pathway persists until $Q \approx 0.35$.

As the protein continues to unfold, we observe multiple unfolding pathways in all models (Fig. 10), consistent with a funneled energy landscape [76,77]. The protein structures from different pulling simulations are distributed across multiple diverse conformations. A similar taxonomy was observed for a loop-truncated variant of SOD1 consisting only of the β-barrel portion, wherein at low degrees of native structure, multiple pathways were observed [45]. While for the AA model there are 3 main clusters at $Q = 0.3$, for the HA-Gō model there are 81 and 41 clusters for panels b) and c), respectively. For clarity, we only show a few representative structures at $Q = 0.35$, 0.3 for panels b) and c).

For the AA and HA-Gō models, the N-terminal β1 most often dissociates directly after the C-terminal β8 dissociates (in 80% and 94% of the trajectories respectively). However, in 2 out of 10 trajectories of the AA model, and in 5% of the HA-Gō trajectories, β-strands 1, 2 and 3 were the last to unfold. This mechanism with β-strands 1 and 2 unfolding last has been observed as the dominant pathway in tSOD1

Table 1
orce-induced unfolding mechanisms of full-length apo, disulfide reduced SOD1 (current paper), and loop-truncated SOD1 (ref. [45]).

Protein construct	Simulation model	Unfolding order of β-strands
Full-length SOD1	All-atom (AA)	5, 4(C-term), 8, 7, 1, 4(N-term), 2, (3, 6)
	HA-Gō	8, 7, 1,4, 5, 2, (3, 6)
	HA-Gō (slow pulling)	8, 7, 1, 4, 5, 6, (2, 3)
Truncated SOD1	All-atom (AA)	8, 1, 7, 2, 3, 4, (6,5)
	HA-Gō	8, 1, 2, 7, 3, 4, (6,5)

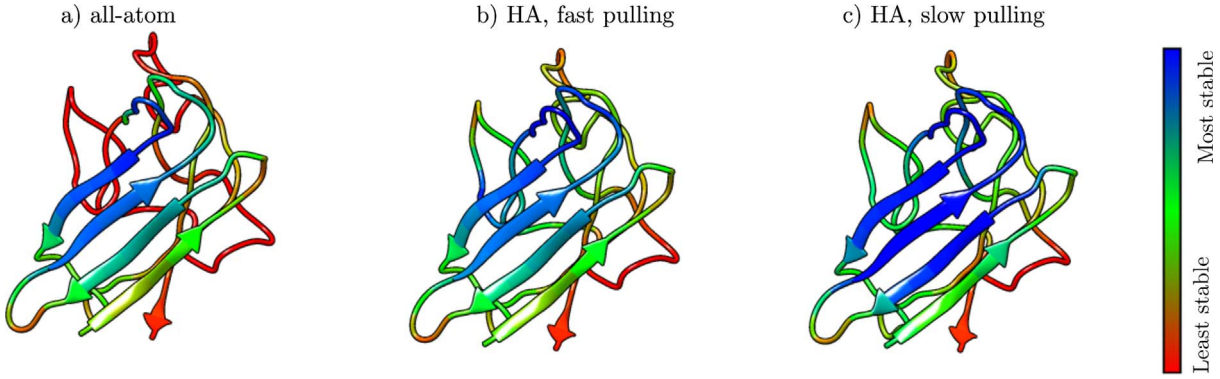

Fig. 8. The degree of native structure retention for each residue as shown in Fig. 7, now superimposed on the native structure of the protein, representing the order in which residues lose more than 50% of their contacts. Panel a): AA structure retention, panel b) HA-Gō structure retention, panel c) HA-Gō (slow-pulling) structure retention.

monomer [63] in coarse-grained (AWSEM [78]) model [45], and the same mechanism was also noted as a less-common unfolding pathway in AA models of tSOD1 [45]. Most often, for 80%, 94%, and 95% of the unfolding trajectories in panels a), b), and c), respectively, the last unfolding events involve breakage of contacts in β-strands 2, 3, and 6.

The dominant unfolding pathway has already been mentioned above and is reviewed here in the context of the unfolding taxonomy. In the AA model, the sequence of unfolding events along the main forced unfolding pathway is β8, then β1 and 7, β2, and finally β3-6. In the HA-Gō model, the main sequence of unfolding of events is β8 and β1, then β2, then β3 and β6 is the last domain to unfold, which is similar to the AA model. In the HA-Gō model at slow pulling rate, β8, part of β1-2, β6 and the remaining β2-3 is the last domain to unfold.

3.5. Force drops are concomitant with loss of interaction potential energy between β-strands

In this subsection we examine the correlation between force drops and the evolution of the local potential energy between several pairs of β-strands (strands 1-2, 1-8, 2-3, and 3-6) *vs.* time. Since we expect this correlation to be dominated by the short-ranged contributions, we include only the non-bonded electrostatic and VDW interactions within a cut-off range of 1.4 nm in the AA model. This neglects solvent effects, but should capture the relative proximity of β-strands. Figures 11 and 12 show two example trajectories of the AA and HA-Gō-model, respectively. The sudden jumps in the potential energy curves coincide with unfolding events, either as observed by force drops or by native structure drops. When the potential approaches zero, there are no contacts between the β-strands, which have then unzipped.

For the above-mentioned β-strands that were examined, the sequence of unfolding events corresponding to the trajectory shown for the force *vs.* time curve in blue in Fig. 11a-1), which represents the majority of the trajectories, are β8, β1, part of β2 and at the end the remaining β2, β3, and β6, see Fig. 11a-2). This is the same sequence of unfolding that we have shown in the rest of the figures in this paper. One can see from the figure that the force drops correlate precisely with losses in potential interaction energy between the corresponding β-strands.

For the force *vs.* time curve in black in Fig. 11b-1), which represents a minority of 2 trajectories out of 10 in the AA model, β8 again unfolds first. At a later time, β3-6 lose structure. For this trajectory, the most stable core is β1,2,3-strands, see the sharp jump in the blue and cyan curves in Fig. 11b-2).

For the HA-Gō model, the force *vs.* time along with the potential interaction energy between the above-mentioned β-strands is shown in Fig. 12, for two different trajectories with fast (a) and slow pulling rate (b). Here we see that force drops often correspond to loss of potential energy between strands, however, several β-strands often lose potential energy at the same time and are difficult to separate. For a pulling rate of 1 m/s in panels a-1), a-2), the time sequence of unfolding events is β8, β1, β2, and β6,3 are the last strands to dissociate.

A representative trajectory is also shown for the HA-Gō model with

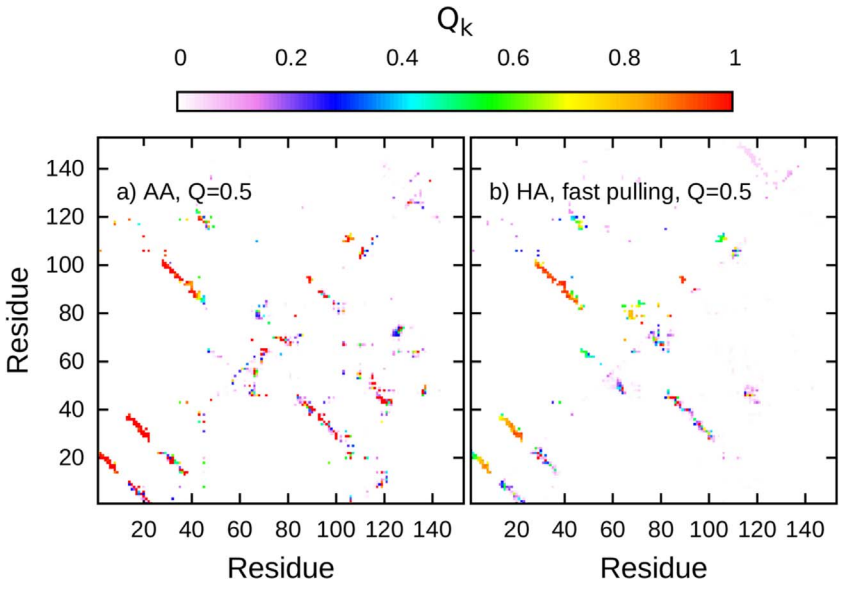

Fig. 9. Contact-maps for AA model (a) and HA-Gō (b) at half-nativeness (Q = 0.5). Probability of formation of native contacts is shown in the upper left triangle, and fraction of conformations having a given non-native contact is shown in the bottom right triangle. The non-native contacts are normalized by the frequency of occurrence, with red contacts corresponding to those occurring in nearly all conformations in the ensemble at Q = 0.5.

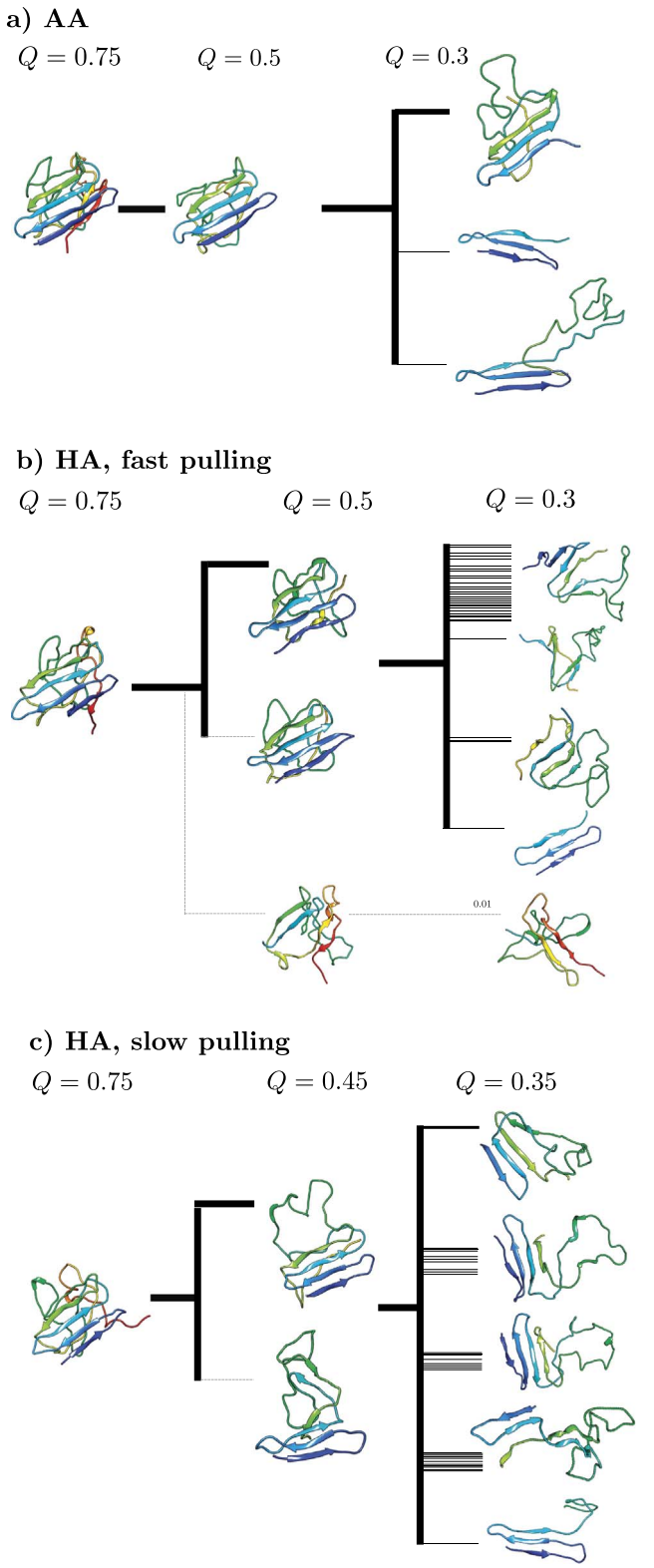

Fig. 10. Centroid conformations obtained from clustering, at different Q values, for a) AA, b) HA-Gō, and c) HA-Gō slow pulling. The dashed line in panel b) shows a single size cluster consisting of 1% of the conformations, where unfolding initiates from the N-terminus.

slower pulling rate of 0.1 m/s in Figs. 12b-1), b-2). Here, β8 again dissociates first, but along with the N-terminal part of β1. Then β6 unfolds, and shortly after that the remaining core of β1, 2, 3 unfolds fairly cooperatively.

We lastly mention some non-trivial aspects of the unfolding trajectories that are evident from the interaction potential energy *vs.* time plots. The interaction energy between two strands may be non-monotonic. Note for instance the increased interaction energy between β1 and β2 at $\approx$ 3 ns in trajectory in Fig. 11a), and the flickering between these two potential energy values between β1 and β2 in trajectory 11b). Similarly, refolding of β1-β2 and β2-β3 occurs after the strands have unfolded in the slow pulling trajectory of Fig. 12b).

Finally, we note that the potential energy between β-strands in the initial equilibrium ensemble does not necessarily predict its persistence during the unfolding trajectory. For example, in trajectory 11b), the initial equilibrium interaction energy between β3 and β6 is stronger than the equilibrium interaction energy between β2 and β3, but β3-β6 unfolds long before β2-β3, while β2-β3 persists until the last unfolding events. Similarly, the equilibrium energy between β1-β2 exceeds that in β3-β6 for the HA-Gō model, but in trajectory 12a), β1-β2 unfolds first.

4. Conclusion and discussion

We explored the unfolding mechanism of a monomeric variant of SOD1, namely the mutant E133Q/C6A/C111S/G50E/F51E, when the protein is subjected to force-induced unfolding. We used both an all-atom (AA), explicit-solvent model, and a HA-Gō model, wherein all non-hydrogen atoms are present, and interact with a short-range Lennard-Jones potential for contacts present in the native state. To facilitate the present comparison between the two models, time and temperature scales were normalized between them.

We analyzed several different metrics for the unfolding process: force-extension curves; native structure *vs.* extension; native structure "fireplots", where sequence of unfolding is captured via loss of contacts involving a particular residue; native structure retention maps, where the most resilient and fragile regions of the proteins are identified; time-resolving the potential interaction energy between specific β-strands; and clustering analysis of actual partially-folded structural conformations.

We found that the force-induced unfolding mechanism shows general agreement between the AA and HA-Gō models; both models capture the same overall sequence of unfolding events. The HA-Gō model captures some of the effects of heterogeneous interaction potential energy, since it counts multiple contacts between amino acids involving large side-chains, but otherwise it is an uncontrolled approximation that may return erroneous conclusions, particularly when electrostatic effects and solvation are important [4]. The HA-Gō model also captures entropic heterogeneity due to the variable backbone polymer length between residues participating in native contacts [79].

It is possible that the strong agreement between the AA and HA-Gō model that we observe here is due to the fast pulling speed. Slower pulling speeds may kinetically allow for alternate non-native hairpins or sheets to form between β-strands. In the predominantly β-sheet protein DDFLN4, AA forced-unfolding simulations were observed to exhibit a force peak specifically due to non-native interactions [80], which recapitulated experimental observations [81], but was not observed in structure-based Gō models. The pulling speeds in AA simulations by Kouza et al. ranged from 1 to 25 m/s, i.e. comparable to or faster than the speeds we used here, so it may be that DDFLN4 forms a non-native intermediate faster than SOD1 kinetically.

By examining unfolding taxonomies, we found that both models initially exhibited a single dominant unfolding pathway until the protein was about half-unfolded, which then bifurcated repeatedly to multiple alternative unfolding pathways. Both mechanisms are consistent with a funnel picture of folding, e.g. the HA-Gō model is an explicitly funneled model.

We have had to assume in this study that the melting temperature of the AA model was equivalent to the experimental melting temperature of apo SS-reduced pseudo-WT SOD1, which may be a severe approximation. We have found previously, however, that the unfolding

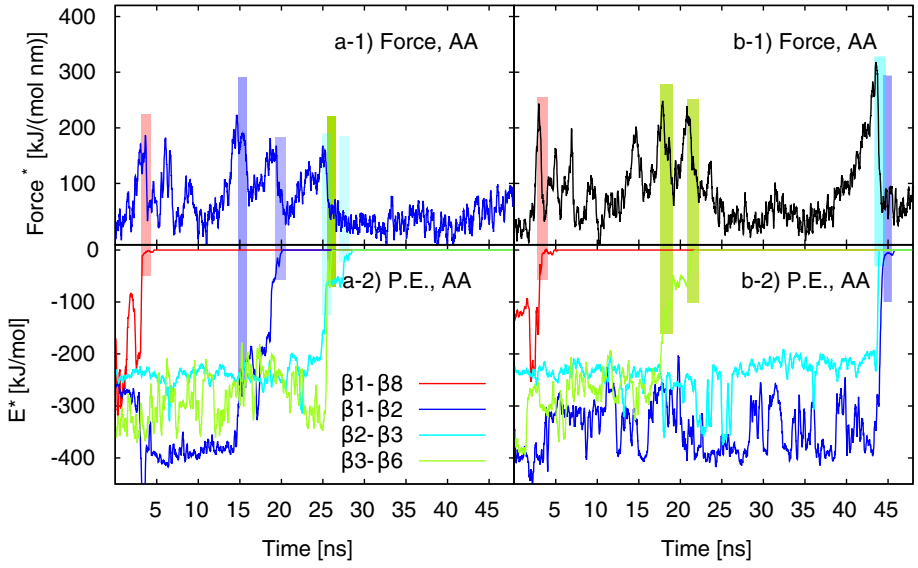

Fig. 11. Force *vs.* time (a-1,b-1) and interaction potential energy between pairs of β-strands *vs.* time (a-2,b-2), for two different unfolding trajectories of the AA model. The force vs time curves in the top panels are the same trajectories similarly colored in the force-extension curves of the AA model in Fig. 5. Bottom panels show the interaction potential energy for all electrostatic and VDW interactions within 1.4 nm, between β1-β8 (red), β1-β2 (blue), β2-β3 (cyan), β3-β6 (green). The sudden losses in interaction potential energy between strands correspond to the force drops in the pulling trajectory (colored vertical bars in the figure).

mechanism for a loop truncated SOD1 (tSOD1) [63] in the AA model is insensitive to variations in temperature of ± 10K [45]. We make the above assumption because of the difficulty in effective sampling necessary to fold AA models of large proteins. We used the above experimental temperature to normalize the temperature scale for the HA-Gō model to its corresponding melting temperature.

We should finally mention that there is nothing absolute about the force-induced unfolding mechanism, which may differ from the unfolding mechanism in either thermal or chemical denaturation. It is known for example that the force-induced unfolding mechanism is linkage dependent [82,83]. That said, the stable core that we observed here in all-atom simulations of force-induced unfolding was consistent with that in chemical denaturation studies of the same construct [40].

An interesting future study, both experimentally and computationally, may be to use either WT or mutant Lysines in SOD1 to apply tether linkages to the interior of the protein chain rather than the ends. Alternatively, one may study which regions of the protein are resilient or fragile when the protein is circularly permuted, for example to place the chain termini near the robust core involving β-strands 2 and 3. It will also be of significant interest to examine the molecular unfolding and refolding of tandem constructs of SOD1, where simulations may help to disentangle the intermolecular interactions that enable proper folding and dimerization from those non-native interactions that induce misfolding.

We finally note an interesting contrast between the unfolding mechanism that we observed here for full-length SOD1, and the mechanism that we had observed previously for loop-truncated SOD1 [45]. Specifically, β-strand 5 was among the first strands to unfold here, while it was among the last strands to unfold in the loop-truncated construct. This implies two compensating effects—one where the entropic penalty of loop IV destabilizes the flanking β5, and another, where in the absence of the long loop, apparently strong interactions stabilize the strand sufficiently so that it is among the last to unfold. Such an evolutionary compensation mechanism may be operating in general for proteins with large loops, and would be an interesting topic of future study.

Transparency document

The http://dx.doi.org/10.1016/j.bbapap.2017.06.009 associated with this article can be found, in online version.

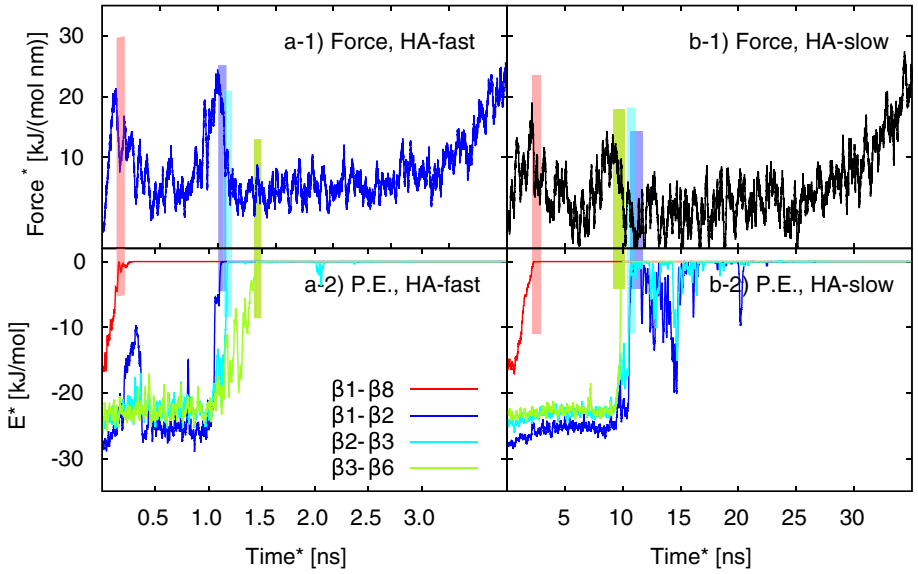

Fig. 12. Force *vs.* time (a-1,b-1) and interaction potential energy between pairs of β-strands *vs.* time (a-2,b-2), for two different unfolding trajectories of the HA-Gō model. Notation and coloring is the same as in Fig. 11 for the AA model. Panels a-1), a-2) correspond to "fast" pulling rate of 1 m/s; panels b-1), b-2) correspond to "slow" pulling rate of 0.1 m/s. Bottom panels show the Lennard-Jones potential energies between β1-β8 (red), β1-β2 (blue), β2-β3 (cyan), β3-β6 (green). The notation * indicates that the force, energy, and time are not rescaled across the different models.

Acknowledgments

This work was supported by the Alberta Prion Research Institute, the Canadian Institutes of Health Research, and the Natural Sciences and Engineering Research Council of Canada. We also acknowledge WestGrid (www.westgrid.ca) and Compute Canada/Calcul Canada (www.computecanada.ca) for providing computing resources. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] D.R. Rosen, T. Siddiquef, D. Patterson, D.A. Figlewicz, P. Sapp, A. Hentatif, J.P. O'Regan, H.X. Dengf, D. McKenna-Yasek, A. Cayabyabi, et al., Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic, Nature 362 (1993) 4, http://dx.doi.org/10.1038/362059a0.

[2] J. Tainer, E. Getzoff, H. Parge, et al., Amyotrophic lateral sclerosis and structural defects in cu, Zn superoxide dismutase, Biol 84 (1980) 151.

[3] S.D. Khare, M. Caplow, N.V. Dokholyan, FALS mutations in cu, Zn superoxide dismutase destabilize the dimer and increase dimer dissociation propensity: a large-scale thermodynamic analysis, Amyloid 13 (4) (2006) 226–235, http://dx.doi.org/10.1080/13506120600960486.

[4] A. Das, S.S. Plotkin, SOD1 Exhibits allosteric frustration to facilitate metal binding affinity, Proc. Natl. Acad. Sci. 110 (10) (2013) 3871–3876, http://dx.doi.org/10.1073/pnas.1216597110.

[5] A. Hörnberg, D.T. Logan, S.L. Marklund, M. Oliveberg, The coupling between disulphide status, metallation and dimer interface strength in Cu/Zn superoxide dismutase, J. Mol. Biol. 365 (2) (2007) 333–342, http://dx.doi.org/10.1016/j.jmb.2006.09.048.

[6] J.A. Rumfeldt, J.R. Lepock, E.M. Meiering, Unfolding and folding kinetics of amyotrophic lateral sclerosis-associated mutant cu, Zn superoxide dismutases, J. Mol. Biol. 385 (1) (2009) 278–298, http://dx.doi.org/10.1016/j.jmb.2008.10.003.

[7] L. Lang, P. Zetterström, T. Brännström, S.L. Marklund, J. Danielsson, M. Oliveberg, SOD1 Aggregation in ALS mice shows simplistic test tube behavior, Proc. Natl. Acad. Sci. 112 (32) (2015) 9878–9883, http://dx.doi.org/10.1073/pnas.1503328112.

[8] A. Abdolvahabi, Y. Shi, S. Rasouli, C.M. Croom, A. Aliyan, A.A. Martí, B.F. Shaw, Kaplan-Meier meets chemical kinetics: intrinsic rate of SOD1 amyloidogenesis decreased by subset of ALS mutations and cannot fully explain age of disease onset, ACS Chem. Neurosci. http://dx.doi.org/10.1021/acschemneuro.7b00029.

[9] A. Gruzman, W.L. Wood, E. Alpert, M.D. Prasad, R.G. Miller, J.D. Rothstein, R. Bowser, R. Hamilton, T.D. Wood, D.W. Cleveland, et al., Common molecular signature in SOD1 for both sporadic and familial amyotrophic lateral sclerosis, Proc. Natl. Acad. Sci. 104 (30) (2007) 12524–12529, http://dx.doi.org/10.1073/pnas.0705044104.

[10] D.A. Bosco, G. Morfini, N.M. Karabacak, Y. Song, F. Gros-Louis, P. Pasinelli, H. Goolsby, B.A. Fontaine, N. Lemay, D. McKenna-Yasek, et al., Wild-type and mutant SOD1 share an aberrant conformation and a common pathogenic pathway in ALS, Nat. Neurosci. 13 (11) (2010) 1396–1403, http://dx.doi.org/10.1038/nn.2660.

[11] K. Forsberg, P.A. Jonsson, P.M. Andersen, D. Bergemalm, K.S. Graffmo, M. Hultdin, J. Jacobsson, R. Rosquist, S.L. Marklund, T. Brännström, Novel antibodies reveal inclusions containing non-native SOD1 in sporadic ALS patients, PloS one 5 (7) (2010) e11552, http://dx.doi.org/10.1371/journal.pone.0011552.

[12] F. Chiti, C.M. Dobson, Amyloid formation by globular proteins under native conditions, Nat. Chem. Biol. 5 (1) (2009) 15–22, http://dx.doi.org/10.1038/nchembio.131.

[13] Y.M. Hwang, P.B. Stathopulos, K. Dimmick, H. Yang, H.R. Badiei, M.S. Tong, J.A. Rumfeldt, P. Chen, V. Karanassios, E.M. Meiering, Nonamyloid aggregates arising from mature copper/zinc superoxide dismutases resemble those observed in amyotrophic lateral sclerosis, J. Biol. Chem. 285 (53) (2010) 41701–41711, http://dx.doi.org/10.1074/jbc.M110.113696.

[14] F. Ding, Y. Furukawa, N. Nukina, N.V. Dokholyan, Local unfolding of Cu, Zn superoxide dismutase monomer determines the morphology of fibrillar aggregates, J. Mol. Biol. 421 (4) (2012) 548–560, http://dx.doi.org/10.1016/j.jmb.2011.12.029.

[15] E.F. Healy, A prion-like mechanism for the propagated misfolding of SOD1 from in silico modeling of solvated near-native conformers, PloS one 12 (5) (2017) e0177284, http://dx.doi.org/10.1371/journal.pone.0177284.

[16] L.I. Grad, W.C. Guest, A. Yanai, E. Pokrishevsky, M.A. O'Neill, E. Gibbs, V. Semenchenko, M. Yousefi, D.S. Wishart, S.S. Plotkin, et al., Intermolecular transmission of superoxide dismutase 1 misfolding in living cells, Proc. Natl. Acad. Sci. 108 (39) (2011) 16398–16403, http://dx.doi.org/10.1073/pnas.1102645108.

[17] C. Münch, J. O'Brien, A. Bertolotti, Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells, Proc. Natl. Acad. Sci. 108 (9) (2011) 3548–3553, http://dx.doi.org/10.1073/pnas.1017275108.

[18] J.P. Crow, J.B. Sampson, Y. Zhuang, J.A. Thompson, J.S. Beckman, Decreased zinc affinity of amyotrophic lateral sclerosis-associated superoxide dismutase mutants leads to enhanced catalysis of tyrosine nitration by peroxynitrite, J. Neurochem. 69 (5) (1997) 1936–1944, http://dx.doi.org/10.1046/j.1471-4159.1997.69051936.x.

[19] A. Tiwari, L.J. Hayward, Mutant SOD1 instability: implications for toxicity in amyotrophic lateral sclerosis, Neurodegener Dis. 2 (3-4) (2006) 115–127, http://dx.doi.org/10.1159/000089616.

[20] L.J. Hayward, J.A. Rodriguez, J.W. Kim, A. Tiwari, J.J. Goto, D.E. Cabelli, J.S. Valentine, R.H. Brown, Decreased metallation and activity in subsets of mutant superoxide dismutases associated with familial amyotrophic lateral sclerosis, J. Biol. Chem. 277 (18) (2002) 15923–15931 jbc.M112087200.

[21] Y. Furukawa, T.V. O'Halloran, Amyotrophic lateral sclerosis mutations have the greatest destabilizing effect on the apo-and reduced form of SOD1, leading to unfolding and oxidative aggregation, J. Biol. Chem. 280 (17) (2005) 17266–17274 jbc.M500482200.

[22] S.Z. Potter, H. Zhu, B.F. Shaw, J.A. Rodriguez, P.A. Doucette, S.H. Sohn, A. Durazo, K.F. Faull, E.B. Gralla, A.M. Nersissian, et al., Binding of a single zinc ion to one subunit of copper- zinc superoxide dismutase apoprotein substantially influences the structure and stability of the entire homodimeric protein, J. Am. Chem. Soc. 129 (15) (2007) 4575–4583, http://dx.doi.org/10.1021/ja066690.

[23] J.A. Rodriguez, B.F. Shaw, A. Durazo, S.H. Sohn, P.A. Doucette, A.M. Nersissian, K.F. Faull, D.K. Eggers, A. Tiwari, L.J. Hayward, et al., Destabilization of apoprotein is insufficient to explain cu, Zn-superoxide dismutase-linked ALS pathogenesis, Proc. Natl. Acad. Sci. 102 (30) (2005) 10516–10521, http://dx.doi.org/10.1016/j.bpj.2015.11.2657.

[24] E. Luchinat, L. Barbieri, J.T. Rubino, T. Kozyreva, F. Cantini, L. Banci, In-cell NMR reveals potential precursor of toxic species from SOD1 FALS mutants, Nat. Commun 5 doi:10.1038/ncomms6502.

[25] J. Danielsson, X. Mu, L. Lang, H. Wang, A. Binolfi, F.X. Theillet, B. Bekei, D.T. Logan, P. Selenko, H. Wennerström, et al., Thermodynamics of protein destabilization in live cells, Proc. Natl. Acad. Sci. 112 (40) (2015) 12402–12407, http://dx.doi.org/10.1073/pnas.1511308112.

[26] S.S. Mojumdar, D.R. Dee, L. Rouleau, U. Anand, C. Garen, M.T. Woodside, Multiple intermediates in the folding of superoxide dismutase 1 revealed by single molecule force spectroscopy, Biophys. J 110 (3) (2016) 497a, http://dx.doi.org/10.1016/j.bpj.2015.11.2657.

[27] L. Banci, I. Bertini, F. Cramaro, R. Del Conte, M.S. Viezzoli, Solution structure of apo Cu, Zn superoxide dismutase: role of metal ions in protein folding, Biochemistry 42 (32) (2003) 9543–9553, http://dx.doi.org/10.1021/bi034324m.

[28] P. Stathopulos, J. Rumfeldt, G. Scholz, R. Irani, H. Frey, R. Hallewell, J. Lepock, E. Meiering, Cu/Zn Superoxide dismutase mutants associated with amyotrophic lateral sclerosis show enhanced formation of aggregates in vitro, Proc. Natl. Acad. Sci. 100 (12) (2003) 7021–7026, http://dx.doi.org/10.1073/pnas.1237797100.

[29] M.J. Lindberg, J. Normark, A. Holmgren, M. Oliveberg, Folding of human superoxide dismutase: disulfide reduction prevents dimerization and produces marginally stable monomers, Proc. Natl. Acad. Sci. 101 (45) (2004) 15893–15898, http://dx.doi.org/10.1073/pnas.0403979101.

[30] A. Nordlund, M. Oliveberg, Folding of Cu/Zn superoxide dismutase suggests structural hotspots for gain of neurotoxic function in ALS: parallels to precursors in amyloid disease, Proc. Natl. Acad. Sci. 103 (27) (2006) 10218–10223, http://dx.doi.org/10.1073/pnas.0601696103.

[31] C. Kayatekin, J.A. Zitzewitz, C.R. Matthews, Zinc binding modulates the entire folding free energy surface of human Cu, Zn superoxide dismutase, J. Mol. Biol. 384 (2) (2008) 540–555, http://dx.doi.org/10.1016/j.jmb.2008.09.045.

[32] A. Nordlund, L. Leinartaitė, K. Saraboji, C. Aisenbrey, G. Gröbner, P. Zetterström, J. Danielsson, D.T. Logan, M. Oliveberg, Functional features cause misfolding of the ALS-provoking enzyme SOD1, Proc. Natl. Acad. Sci. 106 (24) (2009) 9667–9672, http://dx.doi.org/10.1073/pnas.0812046106.

[33] K.A. Vassall, H.R. Stubbs, H.A. Primmer, M.S. Tong, S.M. Sullivan, R. Sobering, S. Srinivasan, L.A.K. Briere, S.D. Dunn, W. Colón, et al., Decreased stability and increased formation of soluble aggregates by immature superoxide dismutase do not account for disease severity in ALS, Proc. Natl. Acad. Sci. 108 (6) (2011) 2210–2215, http://dx.doi.org/10.1073/pnas.0913021108.

[34] Y. Furukawa, I. Anzai, S. Akiyama, M. Imai, F.J.C. Cruz, T. Saio, K. Nagasawa, T. Nomura, K. Ishimori, Conformational disorder of the most immature Cu, Zn-superoxide dismutase leading to amyotrophic lateral sclerosis, J. Biol. Chem. 291 (8) (2016) 4144–4155, http://dx.doi.org/10.1074/jbc.M115.683763.

[35] R. Rakhit, J.P. Crow, J.R. Lepock, L.H. Kondejewski, N.R. Cashman, A. Chakrabartty, Monomeric Cu, Zn-superoxide dismutase is a common misfolding intermediate in the oxidation models of sporadic and familial amyotrophic lateral sclerosis, J. Biol. Chem. 279 (15) (2004) 15499–15504, http://dx.doi.org/10.1074/jbc.M313295200.

[36] M. Chattopadhyay, A. Durazo, S.H. Sohn, C.D. Strong, E.B. Gralla, J.P. Whitelegge, J.S. Valentine, Initiation and elongation in fibrillation of ALS-linked superoxide dismutase, Proc. Natl. Acad. Sci. 105 (48) (2008) 18663–18668, http://dx.doi.org/10.1073/pnas.0807058105.

[37] L. Banci, I. Bertini, C.Y. Chiu, G.T. Mullenbach, M.S. Viezzoli, Synthesis and characterization of a monomeric mutant Cu/Zn superoxide dismutase with partially reconstituted enzymic activity, FEBS J 234 (3) (1995) 855–860, http://dx.doi.org/10.1111/j.1432-1033.1995.855_a.x.

[38] M. Ferraroni, W. Rypniewski, K. Wilson, M. Viezzoli, L. Banci, I. Bertini, S. Mangani, The crystal structure of the monomeric human SOD1 mutant F50E/G51E/E133Q at atomic resolution. The enzyme mechanism revisited, J. Mol. Biol. 288 (3) (1999) 413–426, http://dx.doi.org/10.1006/jmbi.1999.2681.

[39] L. Banci, I. Bertini, F. Cramaro, R. Del Conte, A. Rosato, M.S. Viezzoli, Backbone dynamics of human Cu, Zn superoxide dismutase and of its monomeric F50E/G51E/E133Q mutant: the influence of dimerization on mobility and function, Biochemistry 39 (31) (2000) 9108–9118, http://dx.doi.org/10.1021/bi000067z.

[40] M. Assfalg, L. Banci, I. Bertini, P. Turano, P.R. Vasos, Superoxide dismutase folding/unfolding pathway: role of the metal ions in modulating structural and dynamical features, J. Mol. Biol. 330 (1) (2003) 145–158, http://dx.doi.org/10.

1016/S0022-2836(03)00533-3.
[41] J. Danielsson, M. Kurnik, L. Lang, M. Oliveberg, CuTting off functional loops from homodimeric enzyme superoxide dismutase 1 (SOD1) leaves monomeric β-barrels, J. Biol. Chem. 286 (38) (2011) 33070–33083.
[42] S. Piana, K. Lindorff-Larsen, D.E. Shaw, How robust are protein folding simulations with respect to force field parameterization? Biophys. J. 100 (9) (2011) L47–L49, http://dx.doi.org/10.1016/j.bpj.2011.03.051.
[43] K. Lindorff-Larsen, P. Maragakis, S. Piana, M.P. Eastwood, R.O. Dror, D.E. Shaw, Systematic validation of protein force fields against experimental data, PloS one 7 (2) (2012) e32131, http://dx.doi.org/10.1371/journal.pone.0032131.
[44] P.C. Whitford, J.K. Noel, S. Gosavi, A. Schug, K.Y. Sanbonmatsu, J.N. Onuchic, An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical forcefields, Proteins: Struct. Funct. Bioinf. 75 (2) (2009) 430–441, http://dx.doi.org/10.1002/prot.22253.
[45] M. Habibi, J. Rottler, S.S. Plotkin, As simple as possible, but not simpler: exploring the fidelity of coarse-grained protein models for simulated force spectroscopy, PLoS Comput. Biol 12 (11) (2016) e1005211, http://dx.doi.org/10.1371/journal.pcbi.1005211.
[46] E. Neria, S. Fischer, M. Karplus, Simulation of activation free energies in molecular systems, J. Chem. Phys. 105 (5) (1996) 1902–1921.
[47] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (2) (1983) 926–935.
[48] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J.C. Smith, P.M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29 (7) (2013) 845–854.
[49] D. van Der Spoel, E. Lindahl, B. Hess, GROMACS User Manual Version 4. 6. 7, (2014).
[50] B. Hess, H. Bekker, H.J. Berendsen, J.G. Fraaije, et al., Lincs: a linear constraint solver for molecular simulations, J. Comp. Chem. 18 (12) (1997) 1463–1472.
[51] A. Cheng, K.M. Merz, Application of the Nose-Hoover chain algorithm to the study of protein dynamics, J. Phys. Chem. 100 (5) (1996) 1927–1937, http://dx.doi.org/10.1021/jp951968y.
[52] M. Lingenheil, R. Denschlag, R. Reichold, P. Tavan, The "hot-solvent/cold-solute" problem revisited, J. Chem. Theory Comput. 4 (8) (2008) 1293–1306, http://dx.doi.org/10.1021/ct8000365.
[53] A. Mor, G. Ziv, Y. Levy, Simulations of proteins with inhomogeneous degrees of freedom: the effect of thermostats, J. Chem. Phys. 29 (12) (2008) 1992–1998, http://dx.doi.org/10.1002/jcc.20951.
[54] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys 126 (1)doi:10.1063/1.2408420.
[55] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, J. Appl. Phys. 52 (12) (1981) 7182–7190, http://dx.doi.org/10.1063/1.328693.
[56] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (19) (1995) 8577–8593, http://dx.doi.org/10.1063/1.470117.
[57] J.K. Noel, J.N. Onuchic, The many faces of structure-based potentials: from protein folding landscapes to structural characterization of complex biomolecules, Computational Modeling of Biological Systems, Springer, 2012, pp. 31–54.
[58] J.K. Noel, P.C. Whitford, K.Y. Sanbonmatsu, J.N. Onuchic, SMOG@ctbp: Simplified Deployment of Structure-Based Models in Nucleic Acids Research, doi:10.1093/nar/gkq498.
[59] R.B. Best, G. Hummer, W.A. Eaton, Native contacts determine protein folding mechanisms in atomistic simulations, Proc. Natl. Acad. Sci. 110 (44) (2013) 17874–17879, http://dx.doi.org/10.1073/pnas.1311599110.
[60] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chem. Phys. Lett. 314 (1) (1999) 141–151, http://dx.doi.org/10.1016/S0009-2614(99)01123-9.
[61] E.A. Mills, S.S. Plotkin, Protein transfer free energy obeys entropy-enthalpy compensation, J. Phys. Chem. B 119 (44) (2015) 14130–14144, http://dx.doi.org/10.1021/acs.jpcb.5b09219.
[62] K. Lindorff-Larsen, S. Piana, R.O. Dror, D.E. Shaw, How fast-folding proteins fold, Science 334 (6055) (2011) 517–520, http://dx.doi.org/10.1126/science.1208351.
[63] J. Danielsson, W. Awad, K. Saraboji, M. Kurnik, L. Lang, L. Leinartaitė, S.L. Marklund, D.T. Logan, M. Oliveberg, Global structural motions from the strain of a single hydrogen bond, Proc. Natl. Acad. Sci. 110 (10) (2013) 3829–3834, http://dx.doi.org/10.1073/pnas.1217306110.
[64] Y. Zhang, J. Skolnick, Scoring function for automated assessment of protein structure template quality, Proteins: Struct. Funct. Bioinf. 57 (4) (2004) 702–710, http://dx.doi.org/10.1002/prot.20264.
[65] N. Siew, A. Elofsson, L. Rychlewski, D. Fischer, Maxsub: an automated measure for the assessment of protein structure prediction quality, Bioinformatics 16 (9) (2000) 776–785, http://dx.doi.org/10.1093/bioinformatics/16.9.776.
[66] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (12) (1983) 2577–2637, http://dx.doi.org/10.1002/bip.360221211.
[67] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera-a Visualization system for exploratory research and analysis, J. Comp. Chem. 25 (13) (2004) 1605–1612, http://dx.doi.org/10.1002/jcc.20084.
[68] A. Drozdetskiy, C. Cole, J. Procter, G.J. Barton, JPRed4: a protein secondary structure prediction server, Nucleic Acids Res (2015) gkv332, http://dx.doi.org/10.1093/nar/gkv332.
[69] E.D. Getzoff, D.E. Cabelli, C.L. Fisher, H.E. Parge, M.S. Viezzoli, L. Banci, R.A. Hallewell, Faster superoxide dismutase mutants designed by enhancing electrostatic guidance, Nature 358 (6384) (1992) 347, http://dx.doi.org/10.1038/358347a0.
[70] A. Tiwari, A. Liba, S.H. Sohn, S.V. Seetharaman, O. Bilsel, C.R. Matthews, P.J. Hart, J.S. Valentine, L.J. Hayward, Metal deficiency increases aberrant hydrophobicity of mutant superoxide dismutases that cause amyotrophic lateral sclerosis, J. Biol. Chem. 284 (40) (2009) 27746–27758, http://dx.doi.org/10.1074/jbc.M109.043729.
[71] K. Teilum, M.H. Smith, E. Schulz, L.C. Christensen, G. Solomentsev, M. Oliveberg, M. Akke, Transient structural distortion of metal-free Cu/Zn superoxide dismutase triggers aberrant oligomerization, Proc. Natl. Acad. Sci. 106 (43) (2009) 18273–18278, http://dx.doi.org/10.1073/pnas.0907387106.
[72] A. Sekhar, J.A. Rumfeldt, H.R. Broom, C.M. Doyle, G. Bouvignies, E.M. Meiering, L.E. Kay, Thermal fluctuations of immature SOD1, lead to separate folding and misfolding pathways, elife, 4 (2015) e07296, http://dx.doi.org/10.7554/eLife.07296.
[73] L. Banci, I. Bertini, O. Blaževitš, F. Cantini, M. Lelli, C. Luchinat, J. Mao, M. Vieru, NMR characterization of a "fibrilready" state of demetalated wild-type superoxide dismutase, J. Am. Chem. Soc. 133 (2) (2010) 345–349, http://dx.doi.org/10.1021/ja1069689.
[74] J.S. Fraser, H. van den Bedem, A.J. Samelson, P.T. Lang, J.M. Holton, N. Echols, T. Alber, Accessing protein conformational ensembles using room-temperature x-ray crystallography, Proc. Natl. Acad. Sci. 108 (39) (2011) 16247–16252, http://dx.doi.org/10.1073/pnas.1111325108.
[75] D.A. Keedy, J.S. Fraser, H. van den Bedem, Exposing hidden alternative backbone conformations in x-ray crystallography using qfit, PLoS Comput. Biol 11 (10) (2015) e1004507, http://dx.doi.org/10.1371/journal.pcbi.1004507.
[76] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: a synthesis, Proteins: Struct. Funct. Bioinf. 21 (3) (1995) 167–195, http://dx.doi.org/10.1002/prot.340210302.
[77] S.S. Plotkin, J.N. Onuchic, Understanding protein folding with energy landscape theory part i: basic concepts, Q. Rev. Biophys. 35 (02) (2002) 111–167, http://dx.doi.org/10.1017/S0033583502003761.
[78] A. Davtyan, N.P. Schafer, W. Zheng, C. Clementi, P.G. Wolynes, G.A. Papoian, Awsem-md: protein structure prediction using coarse-grained physical potentials and bioinformatically based local structure biasing, J. Phys. Chem. B 116 (29) (2012) 8494–8503, http://dx.doi.org/10.1021/jp212541y.
[79] M. Rustad, K. Ghosh, Why and how does native topology dictate the folding speed of a protein? J. Chem. Phys 137 (20) (2012) 205104, http://dx.doi.org/10.1063/1.47675675.
[80] M. Kouza, C.-K. Hu, H. Zung, M.S. Li, Protein mechanical unfolding: importance of non-native interactions, J. Chem. Phys 131 (21) (2009) 215103, http://dx.doi.org/10.1063/1.3272275.
[81] I. Schwaiger, A. Kardinal, M. Schleicher, A.A. Noegel, M. Rief, A mechanical unfolding intermediate in an actin-crosslinking protein, Nat. Struct. Mol. Biol. 11 (1) (2004) 81–85, http://dx.doi.org/10.1038/nsmb705.
[82] M. Carrion-Vazquez, H. Li, H. Lu, P.E. Marszalek, A.F. Oberhauser, J.M. Fernandez, The mechanical stability of ubiquitin is linkage dependent, Nat. Struct. Mol. Biol. 10 (9) (2003) 738–743, http://dx.doi.org/10.1038/nsb965.
[83] D.J. Brockwell, E. Paci, R.C. Zinober, G.S. Beddard, P.D. Olmsted, D.A. Smith, R.N. Perham, S.E. Radford, Pulling geometry defines the mechanical resistance of a β-sheet protein, Nat. Struct. Mol. Biol. 10 (9) (2003) 731–737, http://dx.doi.org/10.1038/nsb968.