doi:10.1016/j.jmb.2011.12.029 J. Mol. Biol. (2012) 421, 548–560

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Local Unfolding of Cu, Zn Superoxide Dismutase Monomer Determines the Morphology of Fibrillar Aggregates

Feng Ding [1], Yoshiaki Furukawa [2], Nobuyuki Nukina [3] and Nikolay V. Dokholyan [1]*

[1]*Department of Biochemistry and Biophysics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA*
[2]*Department of Chemistry, Keio University, Yokohama, Kanagawa 223-8522, Japan*
[3]*Laboratory for Structural Neuropathology, RIKEN Brain Science Institute, Wako, Saitama 351-0198, Japan*

Received 19 November 2011; accepted 14 December 2011
Available online
21 December 2011

Edited by D. Case

Keywords:
SOD1 misfolding and aggregation;
fibrillar aggregate;
aggregation building block;
molecular dynamics;
multiscale modeling

Aggregation of Cu, Zn superoxide dismutase (SOD1) is often found in amyotrophic lateral sclerosis patients. The fibrillar aggregates formed by wild type and various disease-associated mutants have recently been found to have distinct cores and morphologies. Previous computational and experimental studies of wild-type SOD1 suggest that the apo-monomer, highly aggregation prone, displays substantial local unfolding dynamics. The residual folded structure of locally unfolded apoSOD1 corresponds to peptide segments forming the aggregation core as identified by a combination of proteolysis and mass spectroscopy. Therefore, we hypothesize that the destabilization of apoSOD1 caused by various mutations leads to distinct local unfolding dynamics. The partially unfolded structure, exposing the hydrophobic core and backbone hydrogen bond donors and acceptors, is prone to aggregate. The peptide segments in the residual folded structures form the "building block" for aggregation, which in turn determines the morphology of the aggregates. To test this hypothesis, we apply a multiscale simulation approach to study the aggregation of three typical SOD1 variants: wild type, G37R, and I149T. Each of these SOD1 variants has distinct peptide segments forming the core structure and features different aggregate morphologies. We perform atomistic molecular dynamics simulations to study the conformational dynamics of apoSOD1 monomer and coarse-grained molecular dynamics simulations to study the aggregation of partially unfolded SOD1 monomers. Our computational studies of monomer local unfolding and the aggregation of different SOD1 variants are consistent with experiments, supporting the hypothesis of the formation of aggregation "building blocks" via apo-monomer local unfolding as the mechanism of SOD1 fibrillar aggregation.

Introduction

The misfolding and aggregation of Cu, Zn superoxide dismutase (SOD1) is associated with amyotrophic lateral sclerosis (ALS).[1,2] More than 100 mutations in SOD1 have been identified in familial ALS patients. Both wild-type and mutant

*Corresponding author. E-mail address: dokh@unc.edu.

Abbreviations used: SOD1, superoxide dismutase; ALS, amyotrophic lateral sclerosis; DMD, discrete molecular dynamics; WHAM, weighted histogram analysis method; RMSF, root-mean-square fluctuation.

SOD1 can form insoluble fibrillar aggregates with a common cross-β amyloid structure,[3,4] as observed in many other amyloidogenic proteins with distinct primary, secondary, and tertiary structures.[5] Many of the research efforts in SOD1 misfolding have been focused on finding a general mechanism for how mutations promote SOD1 misfolding and aggregation, under the assumption of a common "mutation-independent" aggregation pathway and similar aggregate structures.[6] However, a recent study of SOD1 aggregates formed by wild type and various mutants revealed distinct fibrillar core compositions and aggregate morphologies.[7] Accordingly, phenotypic heterogeneity has been reported in familial ALS patients with different SOD1 mutations.[8] Increasing evidence suggests that the structures and morphologies of protein aggregates affect their respective disease phenotypes and that polymorphism in protein aggregates associates with phenotypic heterogeneity.[9–12] Hence, uncovering the molecular mechanism governing the formation of polymorphic amyloid aggregates is important for gaining an understanding of ALS phenotypic heterogeneity.

SOD1 forms a stable dimer in solution, with each SOD1 monomer binding one copper and one zinc ion and forming one intra-monomer disulfide bond. Various biochemical and biophysical studies have suggested that wild-type SOD1 dimer is exceptionally stable because of the coordination of metal ions.[13] Mounting experimental and computational evidence suggest that apoSOD1 monomer is the most aggregation prone species.[14–19] The loss of the coordinated metal ions destabilizes the protein with a significant population unfolded at physiological conditions.[14] Although apoSOD1 is native like in the crystal structure,[20] the protein in solution features significant structural disorder and conformational flexibility.[21] Both experimental[22,23] and computational[19] studies suggest that the apoSOD1 monomer features frequent local unfolding. In addition to the two long loops, the metal-coordinated (in the native state) strands 4, 5, and 7 feature a high level of local unfolding, and the N-terminal β-sheet is the most stable structural element (Fig. 1a). Interestingly, the same regions that are stable in the locally unfolded apoSOD1 also correspond to those regions that participate in the fibrillar core of wild-type SOD1 aggregates, having been identified as proteolysis-resistant peptides.[7] This observation is consistent with the generic aggregation mechanism proposed earlier,[24,25] where the residual structural elements in the partially unfolded protein interact with each other and serve as "building blocks" for the formation of fibrillar amyloid aggregates. Recently, local unfolding induced by mutations has also been found to play an important role in the aggregation of γ-crystallin in human cataracts.[26] Therefore, we hypothesize that the various disease-causative mutations in SOD1 have different impacts on apoSOD1 conformational dynamics, which in turn lead to distinct patterns of local unfolding and thus the varied morphologies of the resulted aggregates.

To test this hypothesis, we apply a multiscale molecular dynamics approach to study the local unfolding of SOD1 monomer and the aggregation dynamics of multiple monomers. In the previous experimental study of SOD1 aggregates, Furukawa *et al.* discovered that three major regions comprise the fibril aggregate cores, including the N-terminal β-sheet (strands 1–3), the middle strand 6, and the C-terminal strand 8 (Fig. 1b). The N-terminal strands are observed in the aggregates formed by all SOD1 variants. As a result, there are only three possible combinations of core-forming peptide patterns: all three segments, the N-terminal sheet plus the middle strand, and the N-terminal sheet plus the C-terminal strand. Therefore, beside the wild-type SOD1 whose aggregation core is composed of all three segments, we also include two mutants, G37R and I149T, having representative aggregation core compositions.[7] In the core of G37R and I149T, the C-terminal strand and the middle strand 6 are observed, respectively, in addition to the N-terminal strands (Fig. 1b). In order to probe the monomer conformational dynamics, we perform atomistic discrete molecular dynamics (DMD) simulations.[19] In the atomistic simulations, we find that different mutations indeed result in different patterns of local unfolding and that the residual structures in the locally unfolded states are consistent with the core-forming segments, supporting the "building block" aggregation mechanism.[24] We further perform coarse-grained DMD simulations to study SOD1 monomer aggregation for each of the three SOD1 variants. To promote the formation of partially unfolded structures, we develop a hybrid structure-based interaction model where the interactions between the core-forming residues are enhanced. The reconstructed model structure of the amyloid aggregates is consistent with the experimentally observed morphology. Therefore, our multiscale simulations suggest a molecular mechanism of mutation-dependent SOD1 aggregation polymorphism.

Results and Discussion

We use a multiscale molecular dynamics approach to study the misfolding and aggregation of apoSOD1. We use the atomistic DMD simulations[27] to sample the conformational dynamics of the apoSOD1 monomer, the timescale of which is approximately microseconds to milliseconds. For monomer aggregation, which spans hours and days,

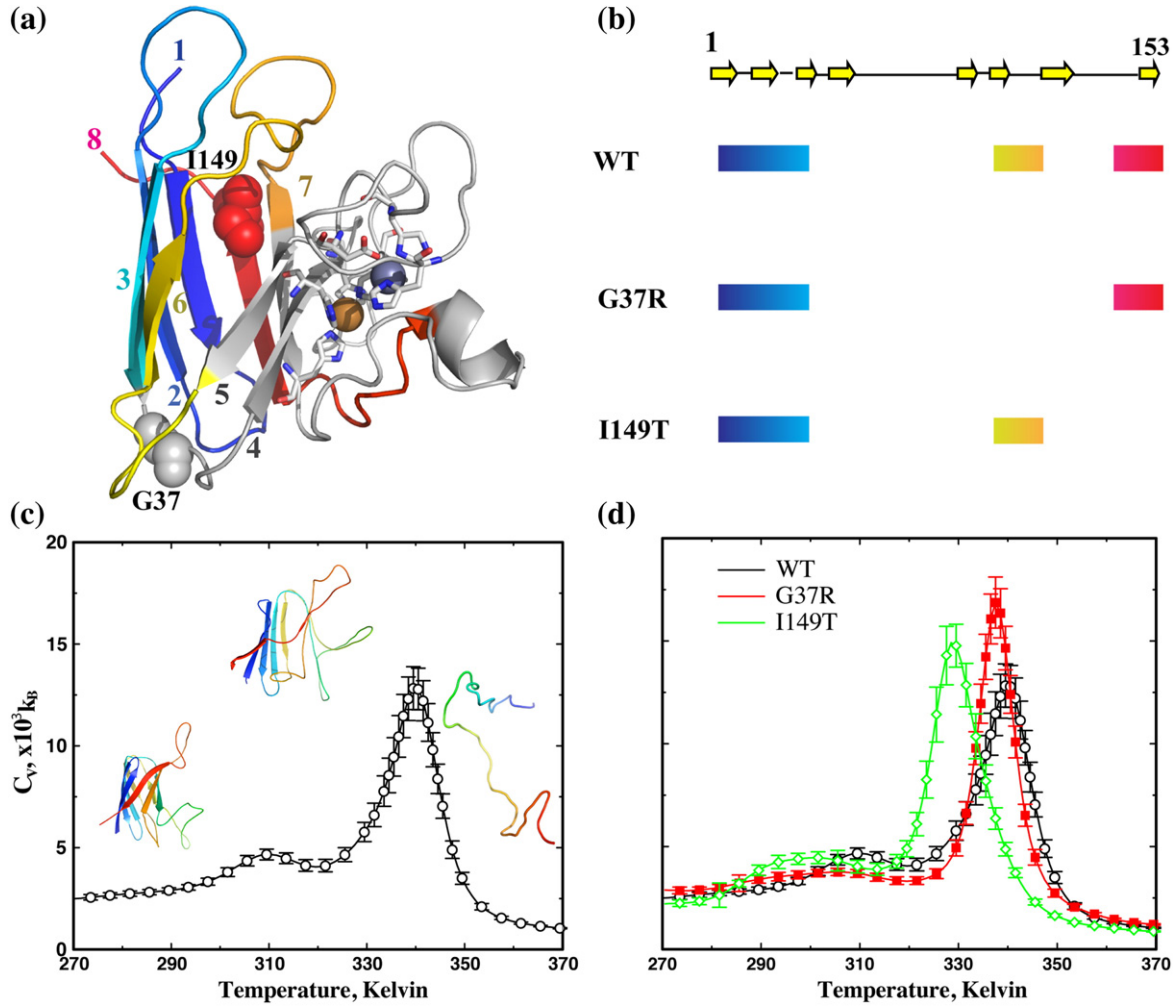

Fig. 1. The thermodynamics of apoSOD1 monomer. (a) In SOD1 monomer, the two metal ions (spheres) are coordinated by residues (sticks) in strands 4, 5, and 7. The amyloid core of wild-type (WT) SOD1 is composed of the N-terminal strands 1–3 (blue to cyan), the middle strand 6 (yellow), and the C-terminal strand 8 (red). Other regions are colored gray. The two mutated residues in this study, G37 and I149, are shown in sphere. (b) The regions forming the amyloid core of three typical SOD1 variants: WT, G37R, and I149T. The arrows illustrate the strands along the primary sequence. The thick lines highlight the regions found in the proteolysis-resistant amyloid core for each of the three SOD1 variants. (c) The specific heat of the WT apoSOD1 as a function of temperature. The specific heat values and corresponding statistical uncertainties (shown as error bars) are computed using WHAM analysis of the replica exchange simulation trajectories (Materials and Methods). At temperatures below the low-temperature peak, the protein is native like with the β-barrel intact. At temperatures higher than the high-temperature peak, the protein is unfolded, without persistent secondary and tertiary structures. At the intermediate temperature between the two peaks, the protein is partially unfolded. (d) The specific heat curves for all three variants feature similar local unfolding and global unfolding thermodynamics.

we use coarse-grained DMD simulations[24] with experimental constraints to enhance sampling of protein aggregation.

Conformational dynamics of apoSOD1 monomer

We perform all-atom DMD simulations[27] to sample the conformational dynamics of apoSOD1 monomers. The protein is in united-atom representation, with all heavy atoms and polar hydrogen explicitly modeled. The simulation is performed with implicit solvent, and interatomic interactions are modeled by a physical force field adapted from Medusa,[28] which includes van der Waals, solvation,[29] and hydrogen bonding potentials. In addition to the previous version of the all-atom DMD force field,[27] we also introduce screened electrostatic interactions between charged residues (Materials and Methods). Atomistic DMD simulations allow

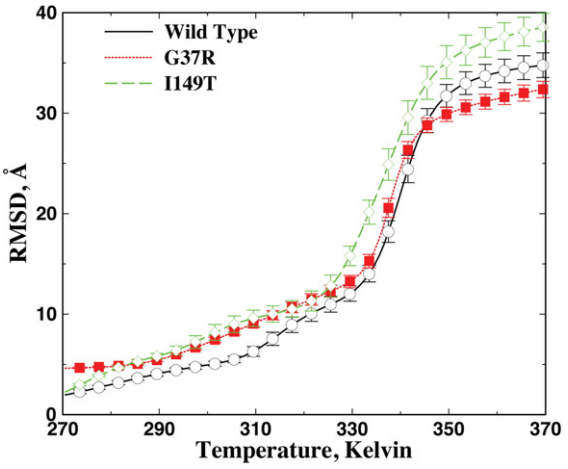

Fig. 2. Average RMSD as a function of temperature. Average RMSD is computed using WHAM analysis. Based on the self-consistently determined density of states $\rho(E)$ and the conditional probability to observe a structure with an RMSD of R at the given energy E, $P(R|E)$, the average RMSD and corresponding error bar can be computed accordingly (Materials and Methods).

efficient sampling of the large-scale conformational dynamics of proteins and protein complexes.[19,30,31]

In order to efficiently sample the conformational dynamics of apoSOD1, we utilize replica exchange[32] DMD simulations, which allow enhanced sampling of the conformational space. We allocate 12 replicas for each SOD1 variant, with each replica running at different temperatures from low to high (Materials and Methods). Periodically, replicas with neighboring temperature values exchange their simulation temperatures stochastically. A temporarily trapped state in a replica can be rescued by simulating at a higher temperature, thereby enhancing the sampling efficiency of DMD simulations. For each simulation, we start from the crystal structure conformation, with metal ions and the disulfide bond removed. The structures for wild type (Protein Data Bank ID: 1SPD) and G37R (Protein Data Bank ID: 1AZV) are known. The starting structure of I149T is modeled by amino acid substitution and rotamer optimization using Eris.[33] In the simulations, cysteine residues do not form disulfide bond mimicking a reducing condition. The total simulation time of each replica is 100 ns. Therefore, we perform a total of 1.2 μs of DMD simulations for each SOD1 variant (Materials and Methods).

ApoSOD1 features frequent local unfolding before global unfolding

Based on replica exchange simulations, we use the weighted histogram analysis method (WHAM; Materials and Methods) to compute the thermodynamics of apoSOD1 unfolding. In WHAM analysis, the density of states is self-consistently determined from overlapping energy histograms for the replica exchange simulations.[34] To reduce structural relaxation artifacts from the starting crystallography structure, we exclude the first quarter of each simulation trajectory from the WHAM analysis. As a simple approach to test convergence of the simulations, we divide the rest of the simulations into halves, and then compute and compare the histogram of potential energies for each of the half. We find that the histograms from both halves are close to each other, suggesting the convergence of the simulations (Fig. S1). Given the density of states, thermodynamic parameters such as specific heat and average root-mean-square deviation (RMSD) from the starting crystal structure can be calculated (details in Materials and Methods).

The specific heat of wild-type apoSOD1 monomer features a major peak at $T_f \sim 330$ K (Fig. 1c), which corresponds to the global unfolding or melting transition of the protein. Above the transition temperature $T > T_f$, the protein is random coil like with large RMSD (Fig. 2). At lower temperatures, the protein is folded or at least partially folded with lower RMSD. Interestingly, a minor peak appears at a lower temperature, $T_p \sim 310$ K, which corresponds to a partial unfolding transition. Such a weak structural transition has also been observed in earlier computational studies, although the global unfolding transition was the main focus of these studies, using the melting temperature as the measure of thermal stability.[19] In an earlier work by Zhou and Karplus, a similar transition was termed surface-molten to solid transition.[35] The lower value of the local unfolding temperature as compared to the global unfolding temperature suggests that the apoprotein undergoes significant local unfolding before it becomes globally unfolded (Fig. 2), which is consistent with the NMR study of the wild-type apoSOD1 monomer in solution.[21]

We similarly compute the specific heat of the two SOD1 mutants (Fig. 1d). Both mutants have a major and a minor peak in the specific heat, similar to that of the wild type. Because of the vast conformational space of SOD1, we do not expect DMD simulations to reach folding/unfolding equilibrium as observed in simulations of small fast-folding proteins.[27] For example, the melting temperature of wild-type apoSOD1 monomer in DMD simulations (~330 K) is higher than that measured in experiments (~315 K).[36] However, starting from the crystal structure, we expect that our microsecond-long replica exchange simulations would be able to sufficiently sample conformations in the folded and partially folded states (e.g., the convergence of simulations illustrated in Fig. S1 and Fig. 2), and the computed unfolding temperatures can be used as a qualitative measure of a protein's thermostability. Compared to the wild type, the I149T mutant has a

lower melting temperature and thus a weaker thermal stability, while G37R has a similar or slightly higher melting temperature. Interestingly, the two mutants have lower local unfolding temperatures (the low-temperature peak in Fig. 1d) than that of the wild type, suggesting that mutations enhance the local unfolding of apoSOD1.

Mutations lead to different local unfolding dynamics

In order to characterize the conformational dynamics of different SOD1 variants in the partially unfolded states, we reconstruct the conformational ensemble from the replica exchange simulation trajectories (see Materials and Methods) based on the RMSD ranges identified from WHAM calculations (Fig. 2). Given the partially unfolded structural ensemble for each SOD1 variant, we compute the root-mean-square fluctuation (RMSF) of each residue around its average positions (Fig. 3a and c). A larger RMSF value denotes higher conformational flexibility of the residue in the locally unfolded state. We also compute the average RMSD of each residue from the corresponding native structure (Fig. 3b and d). In the RMSD per residue calculation, we exclude the two major loops, residues 51–80 and 121–141, which are coordinated by metal ions in the crystal structure and are highly disordered without metals bound. For the wild type, the three N-terminal strands, the middle strand near residue 100, and the C-terminal strand all feature low RMSD and RMSF values. These results are consistent with previous experimental[22] and computational[6,19] studies of local unfolding in the apoSOD1 monomer. These same segments were also found to participate in the fibrillar core by wild-type apoSOD1.[7]

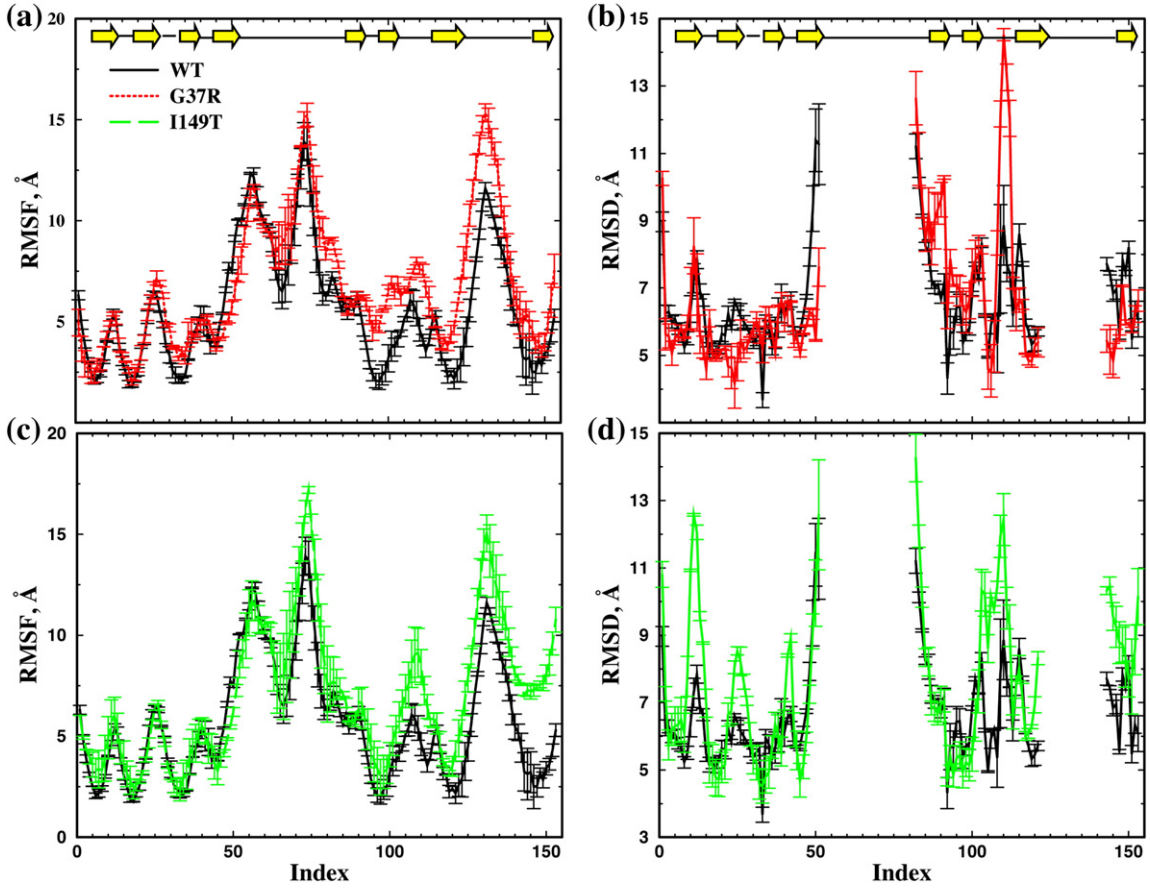

Fig. 3. The local unfolding conformational dynamics of apoSOD1 monomer. The locally unfolded conformational ensemble for each SOD1 variant is reconstructed from the simulation trajectories (Materials and Methods). The corresponding conformational dynamic parameters, including RMSF (a and c) and RMSD (b and d) per residue, are computed based on the reconstructed structural ensemble. In order to estimate the error bars, we split the simulation trajectory into halves and compute the dynamic parameters independently. To avoid overlapping lines, we compare the RMSF and RMSD between WT and G37R (a and c) and between WT and I149T (b and d) separately. The average RMSD per residue is computed with respect to the native states. Due to their large conformational flexibility, we do not include the loops in the RMSD calculation.

G37R displays larger RMSD and RMSF in the middle strand near residue 100 compared to the wild type, while the N- and C-terminal strands of G37R have low RMSD and RMSF similar to that of the wild type (Fig. 3a and b). In the I149T mutant, the C-terminal strand has larger RMSF and RMSD values than those of the wild type. The N-terminal and middle strands of I149T

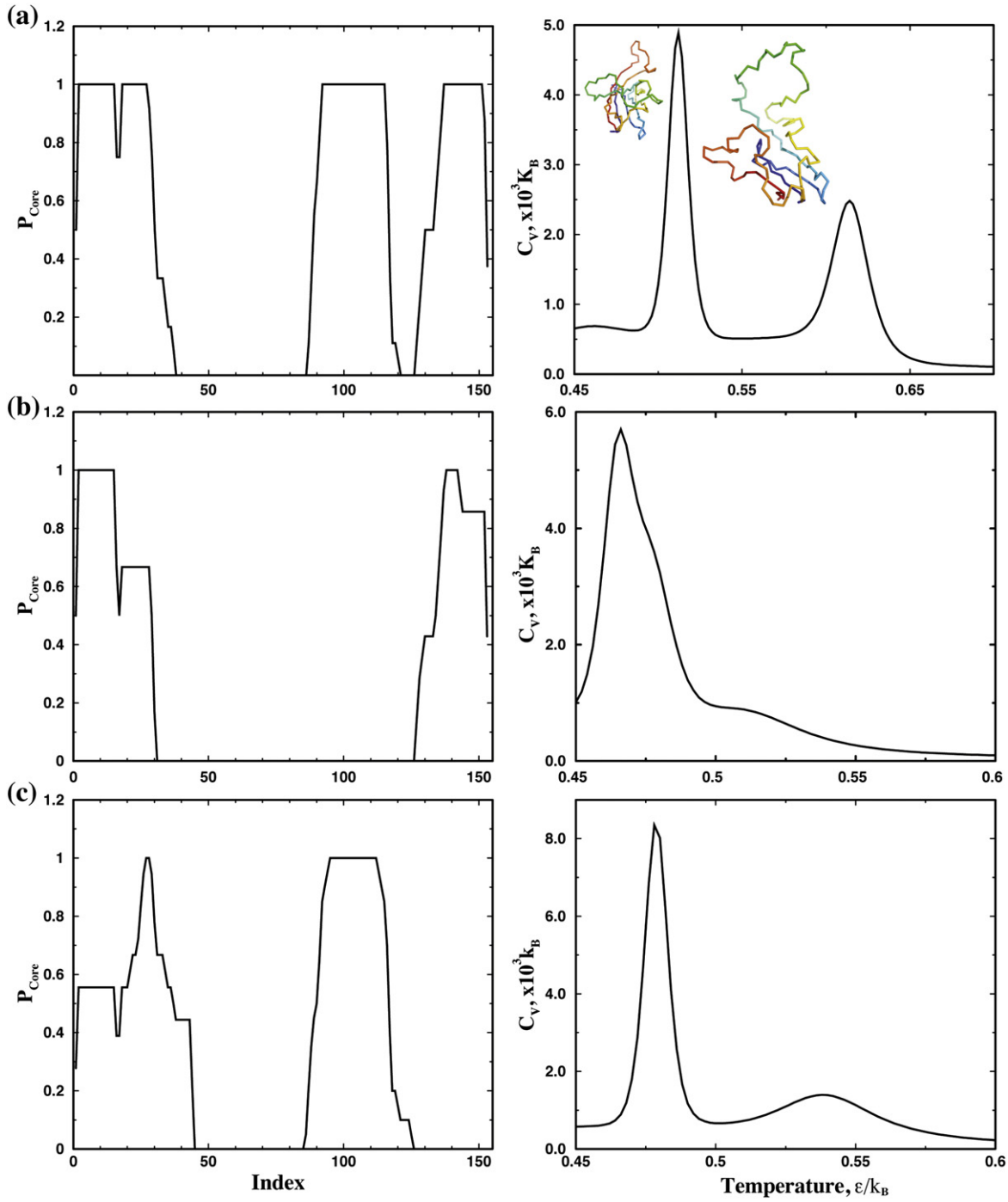

Fig. 4. Thermodynamics of coarse-grained SOD1 monomer: (a) WT, (B) G37R, and (c) I149T. On the left column, the probability of each residue to form the amyloid core (P^{Core}) is computed from the experimentally identified, proteolysis-resistant peptides in the amyloid core (Materials and Methods). On the right column, the specific heat of the coarse-grained SOD1 monomer is computed from replica exchange simulations of monomer folding. Due to the experimentally based bias potential, the protein features a stable partially folded intermediate state.

feature low RMSD and RMSF as is found in the wild type, although the RMSD of the turns in I149T is large (Fig. 3c and d). These regions with low RMSD and RMSF values in mutant SOD1 correspond to the core-forming peptides in the respective aggregates[7] (Fig. 1). Therefore, there is a correlation between the computationally identified stable, native-like regions in the locally unfolded apoSOD1 monomer and the proteolysis-resistant peptide segments in the corresponding amyloid core. Taken together, these results suggest that mutations affect the local unfolding dynamics and result in different patterns of local unfolding. The peptide segments in the well-defined residual structure in the locally unfolded apoSOD1 can serve as a "building block" for fibril aggregates, which interact with each other to form the fibrillar core.

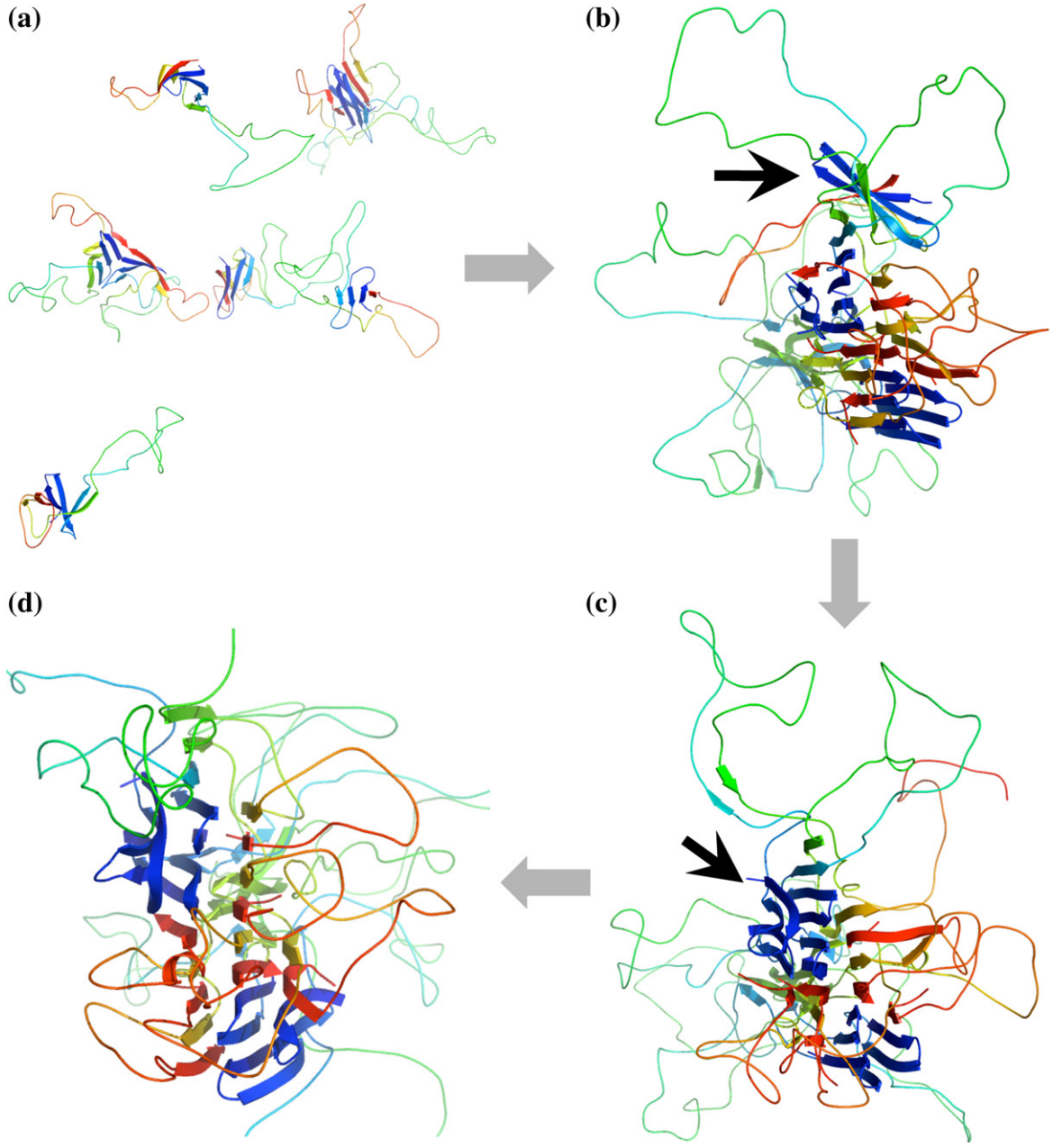

Fig. 5. The aggregation process of WT SOD1 monomers. (a) The initial configuration of eight SOD1 monomers in the simulation box. The monomer concentration is as high as 1 mM. The cartoon representation is assigned and illustrated by using MOLSCRIPT[40] and PyMOL, respectively. (b) One monomer (indicated by the arrow) associates with the end of the nascent amyloid-like aggregate and forms the cross-β core. (c) The associated monomer is incorporated into the aggregate after structural rearrangement. (d) The resulting stable aggregate structure with well-formed cross-β core.

Aggregation of apoSOD1 monomers

The aggregation of SOD1 is a long timescale process, which takes days, weeks, and months under the quiescent condition[37] and hours under the agitated condition.[7] Modeling this process with high-resolution DMD simulations is computationally challenging. Instead, we use a coarse-grained two-bead protein model in DMD simulations[24] to study the aggregation of apoSOD1 monomers, where a structure-based interaction potential for aggregation[24,38] is used (Materials and Methods). To promote the formation of partially unfolded SOD1 as observed in both computation and experiments,[7] we assign stronger attractive interactions between residues forming the amyloid core. The core-forming residues are identified from proteolysis and mass spectroscopy of the aggregates of SOD1 variants (Materials and Methods; Fig. 4).

Experimental constraints promote partial unfolding in SOD1 monomers

We first characterize the coarse-grained monomer folding dynamics of the three SOD1 variants. For each variant, we perform replica exchange DMD simulations of the monomer and compute the specific heat using WHAM (Materials and Methods). Due to the two types of interactions (structure-based interactions and experimentally derived core interactions), the coarse-grained SOD1 monomers feature three-state folding dynamics with two distinct peaks in specific heat, which corresponds to folded intermediate and intermediate unfolded transitions, respectively[39] (Fig. 4a). Strong attractions between the experimentally determined core-forming residues stabilize the partially unfolded intermediates, the structures of which are consistent with the experimentally derived input constraints with the core-forming segments folded (Fig. 4). Although the specific heat plots between coarse-grained and atomistic simulations of corresponding proteins are different (Fig. 1 and Fig. S1), which is expected due to different models and also different types of interaction potentials, they all display two peaks featuring a partially unfolded intermediate. To model the aggregation of SOD1 via the association of the partially unfolded intermediates, we perform equilibrium simulations of multiple SOD1 monomers at the average temperature between the two transitions.

Formation of amyloid-like SOD1 aggregates

For each SOD1 variant, we perform DMD simulations with eight SOD1 monomers in a cubic box with dimensions of 227 Å, corresponding to a high concentration of approximately 1 mM. At a temperature that promotes local unfolding, each isolated monomer remains in the intermediate state. For example, in wild-type SOD1 monomers, the N-terminal strands (blue strands in Fig. 5a), the central strand 6 (yellow strand in Fig. 5a), and the C-terminal strand (red strand in Fig. 5a) are all folded. The monomers associate with each other in simulations and form amyloid-like oligomers (Figs. 5 and 6), where the residual strands in each monomer assemble into β-sheets via inter- and intra-monomer hydrogen bonds. These β-sheets face each other to form high-order "cross-β" structures as shown by the computed fibrillar diffraction pattern (Fig. 6). There are strands from the same protein incorporated into neighboring sheets, which further stabilize the aggregates in addition to the side-chain–side-chain interactions. Our aggregation simulations of three SOD1 variants demonstrate the formation of amyloid-like aggregates by association of the partially unfolded SOD1 monomers. The amyloid-like aggregate structure of SOD1 is not a simple stacking of the residual folded structure but requires a major rearrangement of each monomer.

The ends of each fibril core expose unsatisfied hydrogen bond donors and acceptors, which allow for further fibril growth. For instance, in Fig. 5b–d, we illustrate one monomer incorporation event from the wild-type simulation. An SOD1 monomer initially associates with one end of the amyloid-like aggregate by diffusion (Fig. 5b) and undergoes structural rearrangement in order to be incorporated into the ordered aggregate (Fig. 5c and d). In the final aggregate structures, we notice that both wild type (Fig. 5d) and I149T mutant (Fig. 6c) form a single-core aggregate, while G37R forms an aggregate with two cores during the course of DMD simulations (Fig. 6b). This result is consistent with the experimentally observed morphology of G37R fibrillar aggregate, which is much thinner, branched, and less ordered compared to the fibrils formed by

Fig. 6. The fibrillar aggregates of SOD1: (a) WT, (b) G37R, and (c) I149T. The first and the second columns correspond to the aggregates formed in simulations. Two views are shown by a 90° rotation along the axis of the amyloid fibril. The aggregate of G37R contains two cores. The computed fibril diffraction patterns of the aggregates[24] feature the typical "cross-β" characteristics (third column). The peak along the fibril axis corresponds to the hydrogen bonds between strands, and the peak perpendicular to the axis corresponds to the separation between the adjacent β-sheet. The fourth column corresponds to the electron microscopic images of aggregates formed *in vitro*. The SOD1 fibrils were prepared by shaking disulfide-reduced apoSOD1 at 37 °C, 1200 rpm for 50 h (Materials and Methods). A bar in each panel represents 0.1 μm.

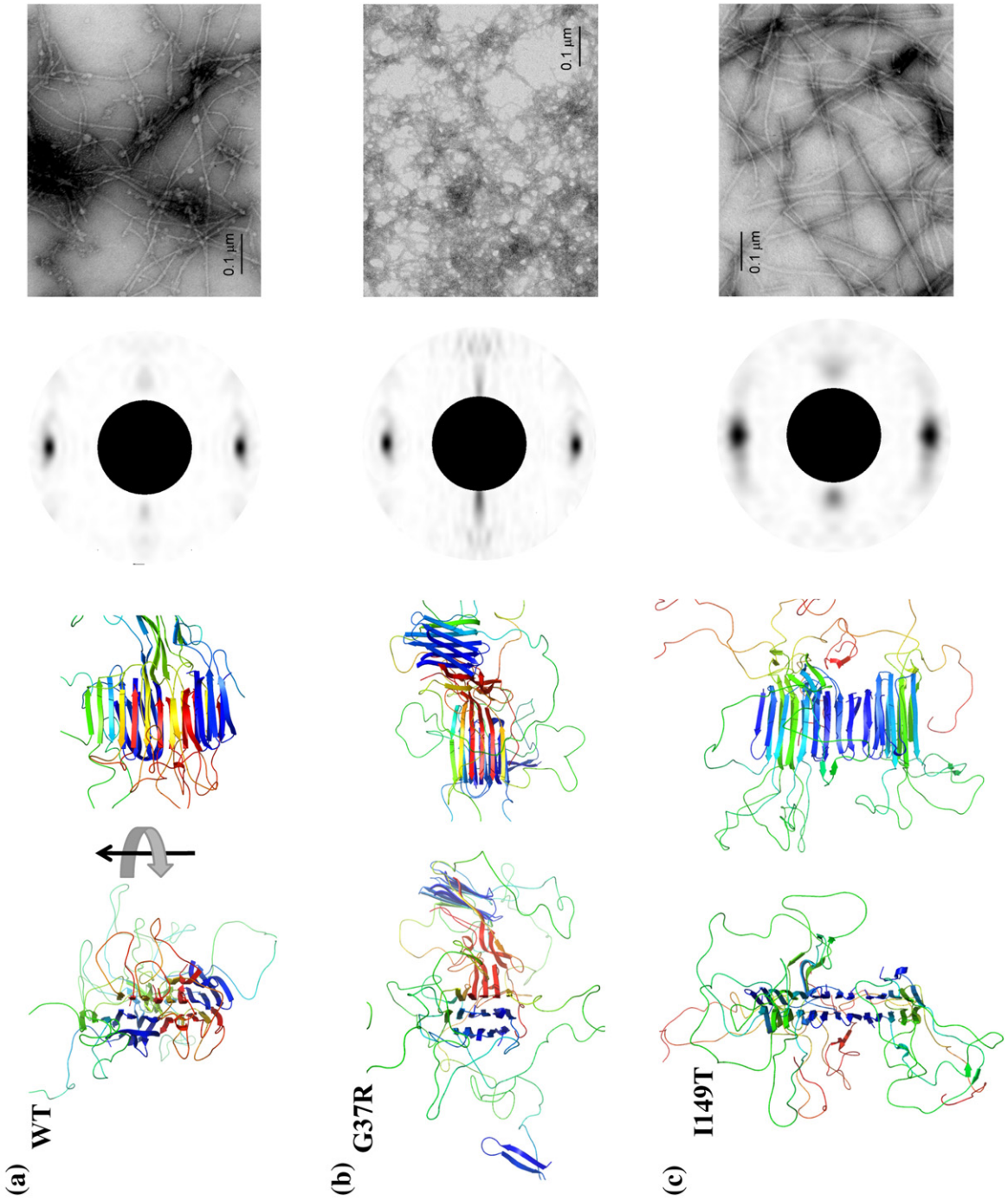

Fig. 6 *(legend on previous page)*

wild type and I149T (Fig. 6). Therefore, we postulate that the abundance of flexible loops in partially unfolded G37R results in the formation of many smaller β-rich aggregates and inhibits the formation of the long and discrete fibrillar aggregates observed in wild type and I149T.

In summary, we performed multiscale molecular dynamics simulations describing monomer conformational dynamics and oligomer formation of apoSOD1 in wild type and two mutants, G37R and I149T, each of which has distinct aggregation morphology and peptide segments forming the fibrillar core. Our simulation results suggested a generic SOD1 aggregation mechanism. After the loss of the stabilizing metal ions and disulfide bond, apoSOD1 monomer undergoes significant local unfolding before global unfolding. Mutations affect the conformational dynamics and result in distinct local unfolding patterns. The residual structure of the partially unfolded protein exhibits an exposed hydrophobic core and unsatisfied hydrogen bond donors and acceptors, which is prone to aggregate. At aggregation-favoring conditions, including high concentration and temperatures that promote local unfolding, the partially unfolded apoSOD1 monomers associate into amyloid-like aggregates with "cross-β" characteristics. The aggregate structure is not formed by a simple stacking of the persistent residual structure of the monomer but, rather, requires structural rearrangement to form the ordered structure. We also found a correlation between the aggregate structures in simulations and the mesoscopic aggregate morphologies observed in experiments. Mutant G37R, which forms less ordered fibrillar aggregates *in vitro*, is found to form two smaller amyloid-like cores due to the large portion of unstructured segments inhibiting the formation of a single ordered "cross-β" core as observed in the two other variants (Fig. 5). Therefore, mutations affect the residual structure of the locally unfolded apoSOD1, where the peptide segments serve as the "building block" of SOD1 aggregation.[24,25] The structured and unstructured regions of the partially folded state determine the morphology of the aggregates.

Materials and Methods

Atomistic DMD simulations

DMD is a special type of molecular dynamics simulation where pairwise interaction potentials are modeled with discontinuous functions.[41] The algorithm for DMD can be found in Refs. 42 and 43. We use an atomistic DMD force field introduced in Ref. 27 to study apoSOD1 monomer dynamics. Briefly, we use the united-atom model to represent the protein, where all heavy atoms and polar hydrogen atoms are explicitly modeled. The bonded interactions include covalent bonds, bond angles, and dihedrals. We include van der Waals, solvation, and environment-dependent hydrogen bonding interactions in the nonbonded interactions. The solvation energy is modeled using the Lazaridis–Karplus implicit solvation model with the fully solvated conformation as the reference state.[29] The hydrogen bond interaction is modeled using a reaction algorithm.[44] In addition to the previous version of the atomistic DMD force field,[27] we also add electrostatic interactions between charged residues, including the basic and acidic residues. We assign integer charges to the central atoms of charged groups: CZ for arginine, NZ for lysine, CG for aspartic acid, and CD for glutamic acid. We use the Debye–Hückel approximation to model the screened charge–charge interactions. The Debye length is set at approximately 10 Å by assuming water relative permittivity of 80 and a monovalent electrolyte concentration of 0.1 mM. We discretize the continuous electrostatic interaction potential with an interaction range of 30 Å, where the screened potential approaches 0 (Fig. 7). We use the constant volume DMD simulations with period boundary conditions and control the simulation temperature using the Anderson thermostat.[45]

Replica exchange simulations and WHAM analysis

We use the replica exchange method to perform simulations of multiple copies of the same system in parallel at various temperatures. At given time intervals, replicas with neighboring temperatures exchange temperature values according to a Metropolis-based stochastic algorithm. We set the temperature exchange interval as 50 ps. Exchange between replicas increases the sampling efficiency in that energetic barriers can be overcome more easily and, in less time, with exposure to higher

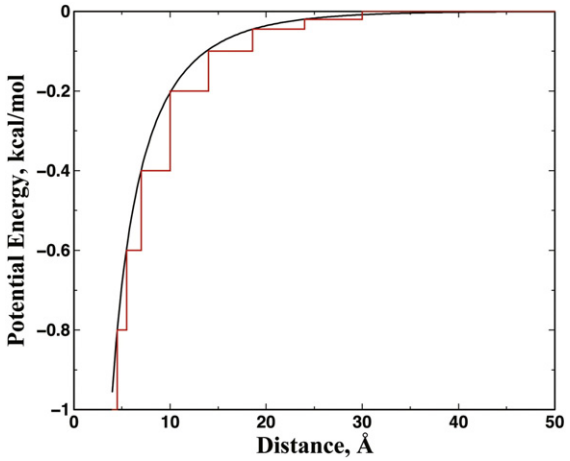

Fig. 7. Screened Debye–Hückel potential function between two opposite single charges. The continuous potential has a Debye length of ~10 Å, assuming the relative permittivity of water of 80 and a monovalent salt concentration of 0.1 mM. The step function in red is the discretized step function utilized in DMD simulations. For two atoms of the same charge, we change the sign of the potential to indicate a repulsive.

temperatures. In our simulations, we allocate 12 replicas with temperatures of 270, 282, 294, 303, 312, 321, 330, 339, 345, 352.5, 365, and 377.5 K for the simulations of apoSOD1 monomers.

We perform WHAM analysis using the trajectories from the replica exchange simulations. The WHAM method utilizes multiple simulation trajectories with overlapping sampling along the reaction coordinates to self-consistently compute the density of states ρ(*E*) by combining histograms from different simulation trajectories.[34] Given the density of states, the folding specific heat (C_v) can be computed at various temperatures according to the partition function $Z = \int \rho(E) exp(-E/k_BT) dE$. Here, k_B is the Boltzmann constant. To compute the average RMSD as a function of temperature, we compute the conditional probability $P(A \mid E)$ by observing a structure with an RMSD of A at the given energy E, evaluated from all simulation trajectories. The average RMSD as a function of temperature can be computed as $\langle A(T) \rangle = 1/Z \int A \cdot P(A|E) \rho(E) exp(-E/k_BT) dE dA$. We also estimate the error bars as statistical uncertainty[46] in the WHAM estimation of specific heat and average RMSD. The temporal correlation in sequentially generated configurations is obtained by autocorrelation analysis.

Reconstruction of the locally unfolded states

We use an RMSD range [D_{min}, D_{max}] to identify the locally unfolded structures from the trajectories. In order to identify the cutoff values, we compute average RMSD as a function of temperature using WHAM analysis (Fig. 2). From the specific heat, we identify the transition temperature at low temperatures corresponding to local unfolding, T_l. We assign the average RMSD at T_l as D_{min} for each of the three SOD1 variants. For all three SOD1 variants, the melting or global unfolding starts from a state with an average RMSD of approximately 10 Å (Fig. 2). Therefore, we use a D_{max} of 10 Å for all three variants.

Coarse-grained DMD simulations

We use a two-bead protein model to study apoSOD1 aggregation.[24,47] In the two-bead model, each amino acid is represented by only the α-carbon (backbone) and β-carbon (side chain). The bonded interactions between neighboring atoms along the peptide chain are assigned so as to mimic peptide geometry.[47] We use a structure-based potential to model the side-chain–side-chain packing interactions, where native interactions in the native state as observed in the crystal structure are favored. Two interacting residues can form either intra- or inter-monomer contacts, in order to promote protein–protein association.[24,38] The attractions between β-carbons are assigned with a hardcore distance of 3 Å and an interaction range of 7.5 Å. We also model the backbone–backbone hydrogen bond interaction as in Ref. 24.

To stabilize the partially unfolded state, where the amyloid-core-forming residues remain folded, we assign a strong attraction between the core residues. In the proteolysis/mass spectroscopy study of aggregates, there are overlapping peptides observed for each core-forming region (e.g., the N-terminal region for wild-type SOD1).[7] For all of the *overlapping peptides* observed experimentally, we assume that each peptide has an equal chance to participate in the amyloid core. Therefore, we can compute the probability for each residue in a given region to be observed in the amyloid core, $P^{Core}(i)$ (Fig. 4). By combing all regions, we obtain the core-forming probability for all residues. We therefore introduce an experimentally derived bias potential to the structure-based potential:

$$E_{ij} = \left[\varepsilon^{G\bar{o}} + \left(\varepsilon^{Core} - \varepsilon^{G\bar{o}} \right) P^{Core}(i) P^{Core}(j) \right] \Delta_{ij}$$

Here, Δ_{ij} is the native interaction matrix, which equals 1 if the two residues i and j are in contact in the native state and 0 otherwise. $\varepsilon^{G\bar{o}}$ and ε^{Core} are the energy scales for the structure-based and experimentally determined bias potentials, correspondingly. In our simulations, we assign $\varepsilon^{G\bar{o}} = 0.5\ \varepsilon_0$ and $\varepsilon^{Core} = 1.0\ \varepsilon_0$ with ε_0 as the energy unit. The hydrogen bond interaction strength is 2.0 ε_0.

Electron microscopic observation of *in vitro* SOD1 fibrils

Preparation and observation of SOD1 fibrils (WT, G37R, and I149T) were performed as described previously.[7] Briefly, following constant agitation of 100 mM SOD1 at 37 °C, 1200 rpm for 50 h, we collected insoluble aggregates by ultracentrifugation at 110,000*g* for 30 min. In order to avoid shearing fibrillar structures during agitation, we further performed a seeding reaction by adding 10 mM SOD1 aggregates (monomer based) to 100 mM soluble SOD1 and incubating the mixture at 37 °C for 3 days in the absence of agitation. Insoluble pellets were again collected by ultracentrifugation, resuspended in pure water, and adsorbed on 400-mesh grids coated by a glow-charged supporting membrane. After negative staining with 1% uranyl acetate, images were obtained using an electron microscope (1200EX, JEOL).

Supplementary materials related to this article can be found online at doi:10.1016/j.jmb.2011.12.029

Acknowledgements

We thank Elizabeth A. Proctor and Rachel Redler for helpful discussions and critical reading of the manuscript. Calculations are performed on the topsail high-performance computing cluster at the University of North Carolina at Chapel Hill. This work was supported by the National Institutes of Health grant R01 GM080742.

References

1. Rosen, D. R., Siddique, T., Patterson, D., Figlewicz, D. A., Sapp, P., Hentati, A. *et al.* (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature*, **362**, 59–62.
2. Bruijn, L. I., Miller, T. M. & Cleveland, D. W. (2004). Unraveling the mechanisms involved in motor neuron degeneration in ALS. *Annu. Rev. Neurosci.* **27**, 723–749.

3. Furukawa, Y., Kaneko, K., Yamanaka, K., O'Halloran, T. V. & Nukina, N. (2008). Complete loss of post-translational modifications triggers fibrillar aggregation of SOD1 in the familial form of amyotrophic lateral sclerosis. *J. Biol. Chem.* **283**, 24167–24176.
4. Oztug Durer, Z. A., Cohlberg, J. A., Dinh, P., Padua, S., Ehrenclou, K., Downes, S. *et al.* (2009). Loss of metal ions, disulfide reduction and mutations related to familial ALS promote formation of amyloid-like aggregates from superoxide dismutase. *PLoS One*, **4**, e5004.
5. Jahn, T. R., Makin, O. S., Morris, K. L., Marshall, K. E., Tian, P., Sikorski, P. & Serpell, L. C. (2010). The common architecture of cross-β amyloid. *J. Mol. Biol.* **395** , 717–727.
6. Khare, S. D. & Dokholyan, N. V. (2006). Common dynamical signatures of familial amyotrophic lateral sclerosis-associated structurally diverse Cu, Zn superoxide dismutase mutants. *Proc. Natl Acad. Sci. U. S. A.* **103**, 3147–3152.
7. Furukawa, Y., Kaneko, K., Yamanaka, K. & Nukina, N. (2010). Mutation-dependent polymorphism of Cu, Zn-superoxide dismutase aggregates in the familial form of amyotrophic lateral sclerosis. *J. Biol. Chem.* **285**, 22221–22231.
8. Andersen, P. M., Nilsson, P., Keranen, M. L., Forsgren, L., Hagglund, J., Karlsborg, M. *et al.* (1997). Phenotypic heterogeneity in motor neuron disease patients with CuZn-superoxide dismutase mutations in Scandinavia. *Brain*, **120**, 1723–1737.
9. Bruce, M. E. & Fraser, H. (1991). Scrapie strain variation and its implications. *Curr. Top. Microbiol. Immunol.* **172**, 125–138.
10. Tanaka, M., Chien, P., Naber, N., Cooke, R. & Weissman, J. S. (2004). Conformational variations in an infectious protein determine prion strain differences. *Nature*, **428**, 323–328.
11. Jones, E. M. & Surewicz, W. K. (2005). Fibril conformation as the basis of species- and strain-dependent seeding specificity of mammalian prion amyloids. *Cell*, **121**, 63–72.
12. Nekooki-Machida, Y., Kurosawa, M., Nukina, N., Ito, K., Oda, T. & Tanaka, M. (2009). Distinct conformations of *in vitro* and *in vivo* amyloids of huntingtin-exon1 show different cytotoxicity. *Proc. Natl Acad. Sci. USA*, **106**, 9679–9684.
13. Shaw, B. F. & Valentine, J. S. (2007). How do ALS-associated mutations in superoxide dismutase 1 promote aggregation of the protein? *Trends Biochem. Sci.* **32**, 78–85.
14. Kayatekin, C., Zitzewitz, J. A. & Matthews, C. R. (2010). Disulfide-reduced ALS variants of Cu, Zn superoxide dismutase exhibit increased populations of unfolded species. *J. Mol. Biol.* **398** , 320–331.
15. Khare, S. D., Caplow, M. & Dokholyan, N. V. (2004). The rate and equilibrium constants for a multistep reaction sequence for the aggregation of superoxide dismutase in amyotrophic lateral sclerosis. *Proc. Natl Acad. Sci. USA*, **101**, 15094–15099.
16. Arnesano, F., Banci, L., Bertini, I., Martinelli, M., Furukawa, Y. & O'Halloran, T. V. (2004). The unusually stable quaternary structure of human Cu, Zn-superoxide dismutase 1 is controlled by both metal occupancy and disulfide status. *J. Biol. Chem.* **279**, 47998–48003.
17. Hough, M. A., Grossmann, J. G., Antonyuk, S. V., Strange, R. W., Doucette, P. A., Rodriguez, J. A. *et al.* (2004). Dimer destabilization in superoxide dismutase may result in disease-causing properties: structures of motor neuron disease mutants. *Proc. Natl Acad. Sci. USA*, **101**, 5976–5981.
18. Rakhit, R., Crow, J. P., Lepock, J. R., Kondejewski, L. H., Cashman, N. R. & Chakrabartty, A. (2004). 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**, 15499–15504.
19. Ding, F. & Dokholyan, N. V. (2008). Dynamical roles of metal ions and the disulfide bond in Cu, Zn superoxide dismutase folding and aggregation. *Proc. Natl Acad. Sci. USA*, **105**, 19696–19701.
20. Elam, J. S., Taylor, A. B., Strange, R., Antonyuk, S., Doucette, P. A., Rodriguez, J. A. *et al.* (2003). Amyloid-like filaments and water-filled nanotubes formed by SOD1 mutant proteins linked to familial ALS. *Nat. Struct. Biol.* **10**, 461–467.
21. Banci, L., Bertini, I., Cramaro, F., Del Conte, R. & Viezzoli, M. S. (2003). Solution structure of Apo Cu,Zn superoxide dismutase: role of metal ions in protein folding. *Biochemistry*, **42**, 9543–9553.
22. Nordlund, A. & Oliveberg, M. (2006). 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. USA*, **103**, 10218–10223.
23. Shaw, B. F., Durazo, A., Nersissian, A. M., Whitelegge, J. P., Faull, K. F. & Valentine, J. S. (2006). Local unfolding in a destabilized, pathogenic variant of superoxide dismutase 1 observed with H/D exchange and mass spectrometry. *J. Biol. Chem.* **281**, 18167–18176.
24. Ding, F., Dokholyan, N. V., Buldyrev, S. V., Stanley, H. E. & Shakhnovich, E. I. (2002). Molecular dynamics simulation of the SH3 domain aggregation suggests a generic amyloidogenesis mechanism. *J. Mol. Biol.* **324** , 851–857.
25. Sinha, N., Tsai, C. J. & Nussinov, R. (2001). A proposed structural model for amyloid fibril elongation: domain swapping forms an interdigitating β-structure polymer. *Protein Eng.* **14**, 93–103.
26. Das, P., King, J. A. & Zhou, R. (2011). Aggregation of γ-crystallins associated with human cataracts via domain swapping at the C-terminal β-strands. *Proc. Natl Acad. Sci. USA*, **108**, 10514–10519.
27. Ding, F., Tsao, D., Nie, H. & Dokholyan, N. V. (2008). *Ab initio* folding of proteins with all-atom discrete molecular dynamics. *Structure*, **16**, 1010–1018.
28. Ding, F. & Dokholyan, N. V. (2006). Emergence of protein fold families through rational design. *PLoS Comput. Biol.* **2**, e85.
29. Lazaridis, T. & Karplus, M. (2000). Effective energy functions for protein structure prediction. *Curr. Opin. Struct. Biol.* **10**, 139–145.
30. Karginov, A. V., Ding, F., Kota, P., Dokholyan, N. V. & Hahn, K. M. (2010). Engineered allosteric activation of kinases in living cells. *Nat. Biotechnol.* **28**, 743–747.
31. Proctor, E. A., Ding, F. & Dokholyan, N. V. (2011). Structural and thermodynamic effects of post-

translational modifications in mutant and wild type Cu, Zn superoxide dismutase. *J. Mol. Biol.* **408**, 555–567.
32. Sugita, Y. & Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**, 141–151.
33. Yin, S., Ding, F. & Dokholyan, N. V. (2007). Eris: an automated estimator of protein stability. *Nat. Methods*, **4**, 466–467.
34. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A. & Rosenberg, J. M. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. *J. Comput. Chem.* **13**, 1011–1021.
35. Zhou, Y. & Karplus, M. (1997). Folding thermodynamics of a model three-helix-bundle protein. *Proc. Natl Acad. Sci. USA*, **94**, 14429–14432.
36. Rodriguez, J. A., Shaw, B. F., Durazo, A., Sohn, S. H., Doucette, P. A., Nersissian, A. M. *et al.* (2005). Destabilization of apoprotein is insufficient to explain Cu,Zn-superoxide dismutase-linked ALS pathogenesis. *Proc. Natl Acad. Sci. USA*, **102**, 10516–10521.
37. Vassall, K. A., Stubbs, H. R., Primmer, H. A., Tong, M. S., Sullivan, S. M., Sobering, R. *et al.* (2011). Decreased stability and increased formation of soluble aggregates by immature superoxide dismutase do not account for disease severity in ALS. *Proc. Natl Acad. Sci. USA*, **108**, 2210–2215.
38. Yang, S., Cho, S. S., Levy, Y., Cheung, M. S., Levine, H., Wolynes, P. G. & Onuchic, J. N. (2004). Domain swapping is a consequence of minimal frustration. *Proc. Natl Acad. Sci. USA*, **101**, 13786–13791.
39. Khare, S. D., Ding, F. & Dokholyan, N. V. (2003). Folding of Cu, Zn superoxide dismutase and familial amyotrophic lateral sclerosis. *J. Mol. Biol.* **334** , 515–525.
40. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 4.
41. Rapaport, D. C. (1978). Molecular dynamics simulation of polymer chains with excluded volume. *J. Phys. A: Math. Gen.* **11**, L213–L217.
42. Allen, M. P. & Tildersley, D. J. (1989). *Computer Simulation of Liquids.* Clarendon Press, New York, NY.
43. Rapapport, D. C. (1997). *The Art of Molecular Dynamics Simulation.* Cambridge University Press, Cambridge, UK.
44. Ding, F., Borreguero, J. M., Buldyrey, S. V., Stanley, H. E. & Dokholyan, N. V. (2003). Mechanism for the α-helix to β-hairpin transition. *Proteins*, **53**, 220–228.
45. Andersen, H. C. (1980). Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* **72**, 2384–2393.
46. Chodera, J. D., Swope, W. C., Pitera, J. W., Seok, C. & Dill, K. A. (2007). Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *J. Chem. Theory Comput.* **3**, 26–41.
47. Ding, F., Dokholyan, N. V., Buldyrev, S. V., Stanley, H. E. & Shakhnovich, E. I. (2002). Direct molecular dynamics observation of protein folding transition state ensemble. *Biophys. J.* **83**, 3525–3532.