

Effects of Protein Crowders and Charge on the Folding of Superoxide Dismutase 1 Variants: A Computational Study

Published as part of *The Journal of Physical Chemistry virtual special issue "Jose Onuchic Festschrift"*.

Atrayee Sarkar, Andrei G. Gasic, Margaret S. Cheung, and Greg Morrison*



Cite This: *J. Phys. Chem. B* 2022, 126, 4458–4471



Read Online

ACCESS |



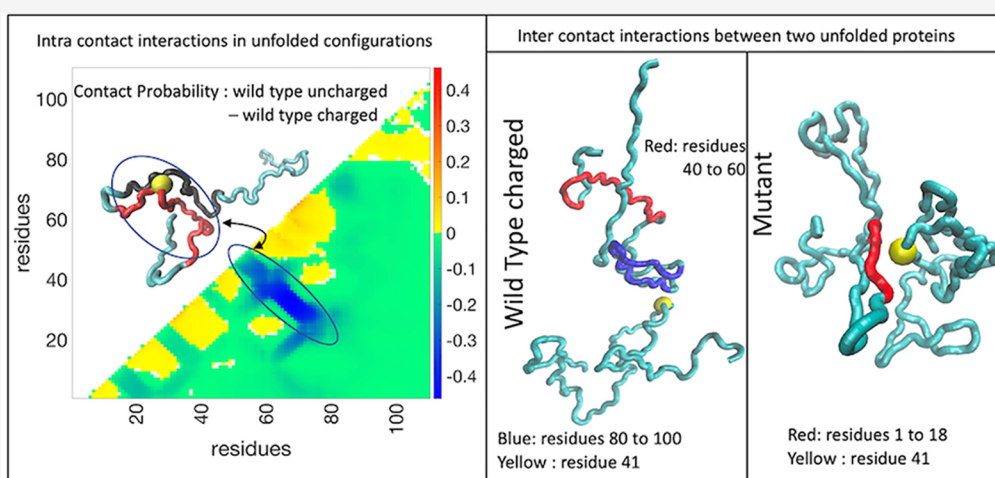
Metrics & More



Article Recommendations



Supporting Information



ABSTRACT: The neurodegenerative disease amyotrophic lateral sclerosis (ALS) is associated with the misfolding and aggregation of the metalloenzyme protein superoxide dismutase 1 (SOD1) via mutations that destabilize the monomer–dimer interface. In a cellular environment, crowding and electrostatic screening play essential roles in the folding and aggregation of the SOD1 monomers. Despite numerous studies on the effects of mutations on SOD1 folding, a clear understanding of the interplay between crowding, folding, and aggregation *in vivo* remains lacking. Using a structure-based minimal model for molecular dynamics simulations, we investigate the role of self-crowding and charge on the folding stability of SOD1 and the G41D mutant where experimentalists were intrigued by an alteration of the folding mechanism by a single point mutation from glycine to charged aspartic acid. We show that unfolded SOD1 configurations are significantly affected by charge and crowding, a finding that would be extremely costly to achieve with all-atom simulations, while the native state is not significantly altered. The mutation at residue 41 alters the interactions between proteins in the unfolded states instead of those within a protein. This paper suggests electrostatics may play an important role in the folding pathway of SOD1 and modifying the charge via mutation and ion concentration may change the dominant interactions between proteins, with potential impacts for aggregation of the mutants. This work provides a plausible reason for the alteration of the unfolded states to address why the mutant G41D causes the changes to the folding mechanism of SOD1 that have intrigued experimentalists.

1. INTRODUCTION

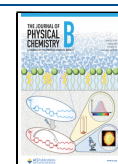
Superoxide dismutase 1 (SOD1) forms a homodimer that prevents oxidative cellular damage,^{1,2} an essential role that depends sensitively on proper folding and dimerization. However, its monomeric apoprotein is marginally stable and is prone to misfolding.^{2–7} Mutations in SOD1 further destabilize the monomer–dimer interface and may lead to prion-like aggregation and fibril formation.^{2–7} Previous studies^{8–11} have identified various mutations in SOD1 as a potential cause of the neurodegenerative disease familial amyotrophic lateral sclerosis (fALS).^{2–7} The stability of SOD1 and its folding

mechanism has been widely studied experimentally *in vitro* and *in vivo*.^{12–22} Previous experiments have shown that the folding (or misfolding) of SOD1 and aggregation of the monomers into dimers or fibrils are complex, multistep processes.^{23,24}

Received: February 2, 2022

Revised: May 26, 2022

Published: June 10, 2022



Gnutt et al.²⁵ showed that point mutations can reverse the stability effect of quinary interactions on SOD1 in the presence of macromolecular crowders. Experiments and all-atom simulations found the existence of multiple intermediate states and their effect on misfolding of SOD1.²⁶ Additionally, all-atom simulation studies of SOD1 in the presence of crowders predicted the unfolding pathway of SOD1 involving several intermediate states.²⁴ Despite these extensive experimental and theoretical studies, the details of how crowding and mutations alter the entire folding and aggregation processes remain unclear. It is imperative to have a clear understanding of the thermodynamic variables that govern protein stability, not only to find a treatment for the neurodegenerative diseases, but also to develop the principles of protein folding in living cells.

In the crowded environment of the cell, excluded volume effects and electrostatics become important, which are known to influence the onset of protein misfolding^{27–29} and aggregation.^{30,31} Inside a cell, the volume exclusion from macromolecules places constraints on the conformational space of proteins^{32,33} and complicates the folding process.^{34–36} Polymeric effects of crowders can further affect a protein's dynamics and function^{37,38} than hard-sphere crowders. SOD1 is a highly charged protein, and the effects of electrostatics on the stability of proteins like SOD1 can depend on both the net charge as well as its charge patterning.³⁷ These electrostatic effects are well-known for intrinsically disordered proteins³⁹ as well as polyampholytes.^{40,41} Factors such as pH, ionic concentration, and oxidative stress also contribute to the folding stability of highly charged proteins.⁴¹ Dimerization of SOD1 is the first step in the formation of the native tetramer or other misfolded aggregates, so the interaction of multiple SOD1 proteins with one another is important to understand.

In this paper, we study the interplay of electrostatics and self-crowding on SOD1 (PDB ID 4BCZ). Since most experimental and all-atom simulation investigations of SOD1 highlight the effects of perturbing the native state,^{14–18,25} we focus on understanding the entire folding process, particularly the unfolded state. As such, we use coarse-grained molecular dynamics simulations to ensure sufficient sampling and proper structural features.⁴² We study the equilibrium behavior of SOD1 under multiple conditions to highlight the effect of charge: an uncharged system (high ionic strength) and an unscreened Coulombic interaction (low ionic strength) for the Wild Type (WT). These conditions represent the extremes of electrostatic interaction and highlight the effects of charge most starkly. These simulations are performed for single proteins as well as a self-crowded state of 64 proteins. We further study the effect of G41D point mutation on SOD1 and how this mutation alters the interactions between nearby SOD1 pairs (which may have implications for the development of disease-causing aggregation). The residue 41 in SOD1 is known to have two mutations, G41D and G41S, related to ALS disease^{43–45} that showed that these mutations can lead to protein aggregation.⁴⁶ In the point mutation G41D, the neutral G amino acid at residue 41 of wild type SOD1 is replaced by a negatively charged D amino acid. We will show that residue 41 on the β -strand 4 takes part in the formation of an intermediate state and is an essential part of the folding pathway of SOD1 and that this mutation leads to interprotein interactions that differ from those observed for pairs of WT proteins.

The outline of this paper is as follows: In section 2 we describe the simulation and analysis methodology, and system

design. In section 3 we show that the crowding and electrostatics significantly affect the unfolded configurations in comparison to the folded ones by calculating the probability of intracontacts. We develop a new order parameter based on pairwise interactions that are more likely if charge is included (in comparison to the uncharged simulations) and identify an intermediate state for SOD1 folding. We also investigate the effect of a point mutation on SOD1 folding by computing the intercontact probabilities between proteins. The order parameter remains relevant for the G41D mutation, with changes to the location of interprotein interactions the primary effect of the mutation. Finally, in section 4 we discuss the implications of our results, highlight the effects of the mutation on interprotein interactions, and discuss the implications in aggregations and possible disease formation.

2. COMPUTATIONAL METHODS

2a. Coarse-Grained Molecular Dynamics (MD) Simulation. For our simulations investigating the mechanism of SOD1 folding dynamics, we use a structure-based model, which is a minimalist protein model (“beads on a chain”) that incorporates experimentally derived structural information.⁴⁸ Structure-based models provide a folding energy landscape with minimal frustration and contain a funneled landscape with a dominant basin of attraction corresponding to an experimentally determined configuration.^{49,50} The folded state bias dramatically reduces the complexity of the resulting force field and allows for a clear physical understanding of a system. These minimal models give access to biologically relevant time scales while retaining the essential dynamical features that are observed in experiments.⁵¹ Additionally, the polymeric scaling seen in unfolded proteins⁵² is preserved using structure-based models,⁴² whereas all-atom models tend to overestimate the conformational compactness of proteins.⁵³ Previous work has compared all-atom and structure-based models for the folding/unfolding of SOD1, showing that structure-based models capture the essential sequences of unfolding.⁵⁴ These benefits are key for understanding the unfolded ensemble of SOD1.

Each loop-depleted SOD1 protein variant is represented as 110 C α -atoms (residues). For an individual protein variant α , the Hamiltonian of the structure-based model, $\mathcal{H}_\alpha$, is

$$\begin{aligned} \mathcal{H}_\alpha(\Gamma_\alpha, \Gamma_0) = & \sum_{i < j} K_r (r_{ij}^\alpha - r_{ij}^0)^2 \delta_{j,i+1} + \sum_{i \in \text{angles}} K_\theta (\theta_i^\alpha - \theta_i^0)^2 \\ & + \sum_{i \in \text{dihedrals}} K_\phi \left(\{1 - \cos[\phi_i^\alpha - \phi_i^0]\} + \frac{1}{2} \{1 - \cos[3(\phi_i^\alpha - \phi_i^0)]\} \right) \\ & + \sum_{i < j} \frac{e^{-k r_{ij}^\alpha} q_i q_j}{4\pi\epsilon_0 r_{ij}^\alpha} + \sum_{i,j \in \text{native}} \epsilon \left[\left(\frac{r_{ij}^0}{r_{ij}^\alpha} \right)^{12} - 6 \left(\frac{r_{ij}^0}{r_{ij}^\alpha} \right)^{10} \right] + \sum_{i,j \notin \text{native}} \epsilon \left(\frac{\sigma}{r_{ij}^\alpha} \right)^{12} \end{aligned} \quad (1)$$

where Γ_α represents the configuration of protein α (including the set of pairwise distances r , the set of angles between three consecutive beads θ , and the set of dihedral angles defined over four sequential residues ϕ). The native structure values of r_{ij}^0 , θ_i^0 , and ϕ_i^0 were obtained from their crystal structure configuration (Γ_0). In the backbone terms, K_r , K_θ , and K_ϕ are force constants of the bond, bond-angle, and dihedral potentials, respectively. Setting the energy scale $\epsilon \equiv 1$, the coefficients are given the values: $K_r = 200\epsilon/\text{\AA}^2$, $K_\theta = 40\epsilon$, $K_\phi = \epsilon$.⁵⁵ The third to last term is a Debye–Hückel potential, $U_{\text{DH}}(r_{ij}^\alpha)$, where $4\pi\epsilon_0 = 1$. In this study, we explore the extreme limits of this interaction; that is, $\lim_{k \rightarrow \infty} U_{\text{DH}}(r_{ij}^\alpha) = 0$ and

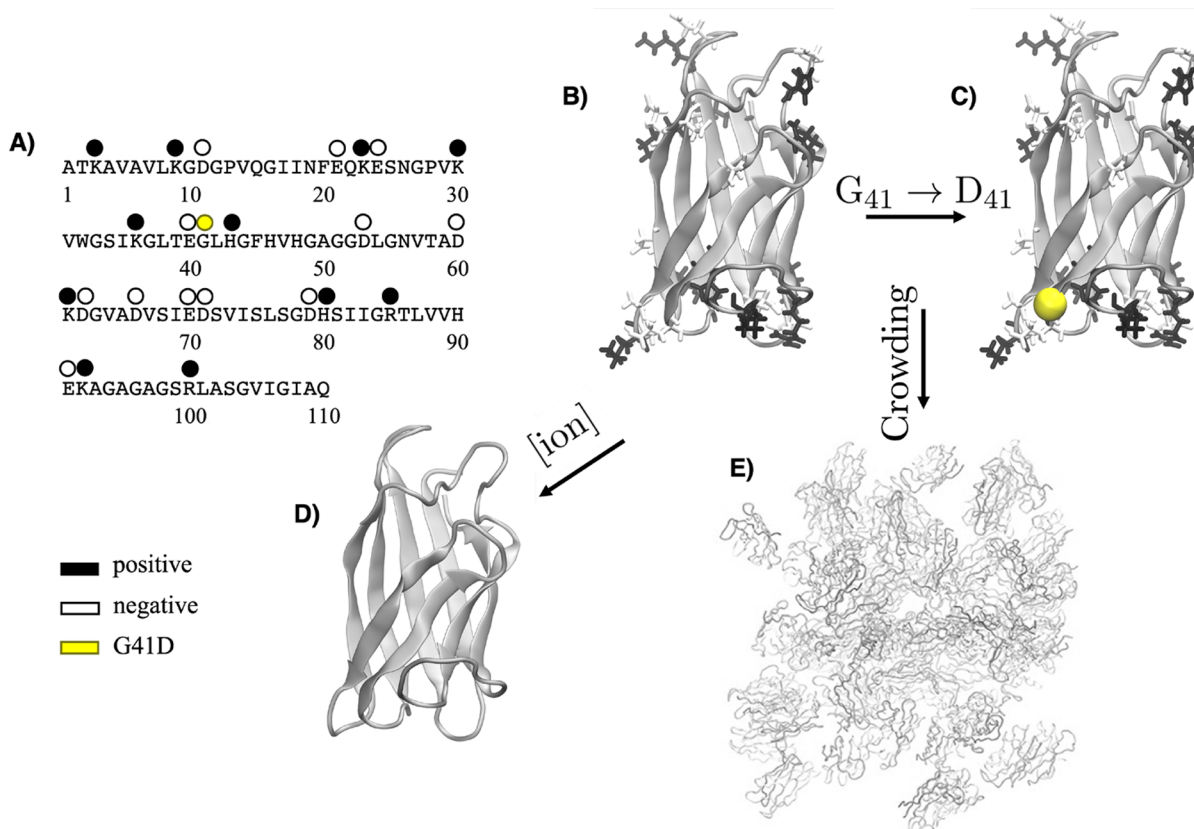


Figure 1. Sequence and structure of SOD1 with system perturbations. (A) Sequence of SOD1. Black-filled and empty circles indicate positively and negatively charged residues, respectively. The yellow circle indicates residue 41, which is the site of the point mutation in Mutant. (B) Native crystal structure of the loop-depleted SOD1_{WT} (PDB ID 4BCZ) with β -strands depicted in gray. Black and white side chains indicate positive and negative charged residues, respectively. (C) The point mutation in Mutant is shown as a yellow sphere and is located at the beginning of β -4. (D) At high ionic strength, screening effectively removes charges. (E) Illustration of a simulation box containing 64 identical SOD1_{WT} protein crowders excluding $\approx 20\%$ of the total volume when folded. These structures are used as the basis for charge (ion concentration) and identical protein crowders volume fraction perturbations to our coarse-grained C^α model system. Visualized using VMD.⁴⁷

$\lim_{k \rightarrow 0} U_{\text{DH}}(r_{ij}^\alpha) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}^\alpha}$. The last two terms are a 10–12 Lennard–Jones interaction for native contacts^{48,56} and a repulsive volume exclusion potential for non-native interactions ($\sigma = 4 \text{ \AA}$). In the crowded case, the total Hamiltonian of the system takes the following form:

$$\mathcal{H} = \sum_{\alpha=1}^N \mathcal{H}_\alpha + \sum_{\alpha=1}^{N-1} \sum_{\beta=\alpha+1}^N \sum_{i,j} \left[\epsilon \left(\frac{\sigma}{r_{ij}^{\alpha\beta}} \right)^{12} + \sum_{\alpha=1}^{N-1} \sum_{\beta=\alpha+1}^N \sum_{i,j} \frac{e^{-kr_{ij}^{\alpha\beta}}}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}^{\alpha\beta}} \right] \quad (2)$$

where $r_{ij}^{\alpha\beta}$ is the interaction distance between protein residues i and j of proteins α and β , and N is the total number of proteins. Specific interactions between different proteins were not included.

2b. Simulation Setup and Details. We performed all simulations using GROMACS⁵⁷ to integrate Langevin equations of motion in the underdamped limit. To prepare the computational model of SOD1 (PDB ID: 4BCZ) for the GROMACS simulations, we used the SMOG (Structure-based Models for Biomolecules⁵⁵) software, and we used a H++⁵⁸ server to predict the correct protonated state of the charged residues.

Using the natural time unit of a coarse-grained model $\tau = (m\sigma^2/\epsilon)^2$ ($\approx 1.9 \text{ ps}$) and an integration time step of $10^{-3}\tau$, we equilibrated the system for $10^4\tau$ over a wide range of temperatures. With the initial structures taken from the

equilibration step, we simulated trajectories over a temperature range of $120 \text{ K} < T < 220 \text{ K}$ and repeated this for all conditions (single and crowded cases for wild type, mutant, and uncharged). Since degrees of freedom are coarse-grained, the simulation temperature is colder than the experimentally observed temperature. We sampled the trajectories at every 5τ , resulting in a total number of approximately 6×10^5 samples for each single protein condition and 1×10^4 samples for each crowded protein condition. For the crowded case, we used a box of 64 proteins with a volume of 3160 nm^3 , resulting in an excluded volume of approximately $\phi = 0.2$ (calculated using the radius of gyration of the native protein structure in the folded state). Note that ϕ will slightly decrease as the proteins unfold since unfolded proteins usually have a lower volume than folded ones. Periodic boundary conditions are imposed for all simulations.

2c. System Design. To investigate the effect of charges on the folding stability, we have used three SOD1 variants: the wild type loop-depleted monomer with unscreened charges (WT_c), a wild-type uncharged monomer (WT_u), and the G41D mutant (*Mutant*). The native crystal structure of the SOD1 monomer with 110 residues forms a β -barrel with 8 β -strands: $\beta 1$ (residues: 3–9), $\beta 2$ (residues: 15–21), $\beta 3$ (residues: 29–36), $\beta 4$ (residues: 41–47), $\beta 5$ (residues: 53–59), $\beta 6$ (residues: 65–71), $\beta 7$ (residues: 86–90), and $\beta 8$ (residues: 101–110) (Figure 1). The total charge in WT_u , *Mutant*, and

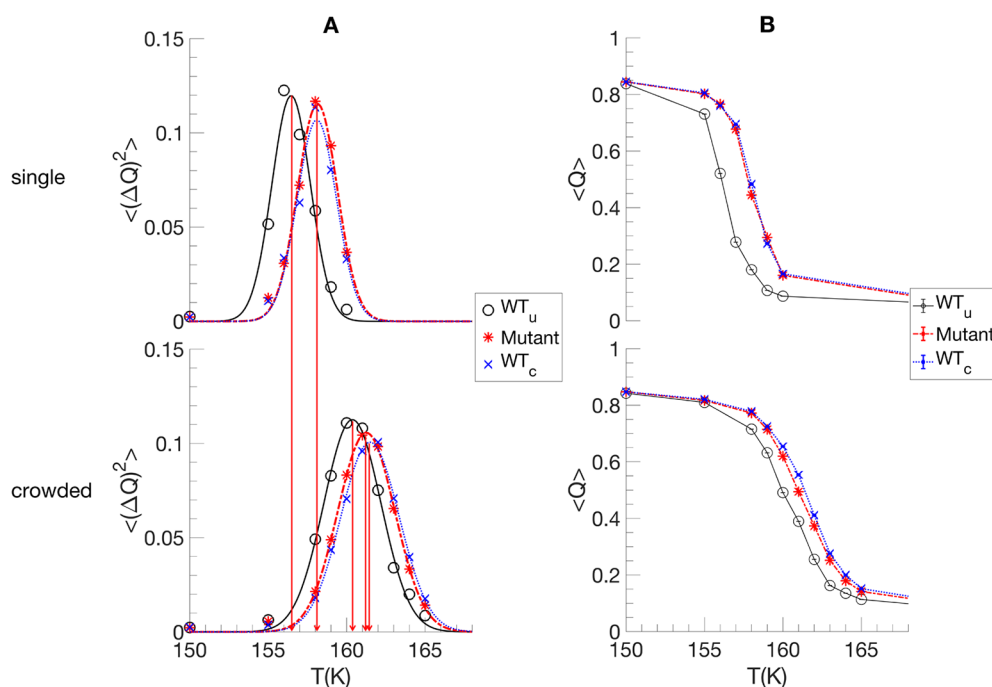


Figure 2. (A) Time average of the variance of Q vs temperature (T) in single protein systems (upper panel A) and crowded protein systems (lower panel A) for WT_u , *Mutant*, and WT_c . The peak values of $\langle \Delta Q^2 \rangle$ indicate a phase transition of the protein variants from folded to unfolded configurations and give the folding temperatures T_f . For both the single and crowded systems, the inclusion of electrostatic interactions increased T_f . Crowding also increased T_f for all the protein variants. (B) Q as a function of temperature for single (upper panel B) and crowded cases (lower panel B) for WT_u , *Mutant*, and WT_c . $\langle Q \rangle$ for all variants merge to a single value with a decrease in temperature in both the single and crowded ensembles, indicating that the variants have the same native states.

WT_c are 0, $-2e$, and $-1e$, respectively. Simulations were performed with unscreened Coulombic interactions between charged residues except for WT_u where all Coulombic interactions are neglected. These represent the extremes of ionic strength, corresponding to zero ionic strength for WT_c and infinite ionic strength for WT_u . The G41D mutation in *Mutant* adds a negative charge and is simulated with an unscreened Coulombic interaction, reducing the total charge of the protein and is “in between” the charged and uncharged cases when it comes to the total Coulombic interaction energy. We performed simulations in dilute conditions (single proteins) and self-crowded cases (64 proteins with the same Coulombic interactions: WT_c , *Mutant*, and WT_u). In crowded cases, the proteins of the same variant type act as self-crowders (for example, 64 WT_c proteins are used to understand the effect of self-crowding on WT_c). The ratio of unfolded to folded configurations after equilibration in the crowded systems was 1.3959, 0.5509, and 0.79077 in WT_u , WT_c , and *Mutant*, respectively.

2d. Order Parameters. To understand the dynamics of SOD1, we must be able to distinguish between different distinct ensembles of structures (e.g., the difference between “folded” and “unfolded”). In this paper, we will use multiple methods of quantifying the state of a protein at each sampled time step (many well-known and one novel variable). It is thus useful to describe these order parameters and indicate the differences between them.

1. In the main text, we compute the fraction of native contacts $Q = N_c^{-1} \sum_{(ij) \in \text{native}} \theta(1.2r_{ij}^N - r_{ij})$, widely used as a reaction coordinate in protein folding analysis^{59–61} to determine the state of a protein. In the definition of Q , “native” is the set of residue pairs that are in contact

in the native state (with a total of $N_c = 360$ native contacts), r_{ij} is the distance between residue pairs i and j of a protein, r_{ij}^N is the native distance between the residue pairs, and $\theta(x)$ is the Heaviside function. The set of native pairs in the definition of Q is identical to those in the Hamiltonian in eq 1. Note that charged interactions are nonspecific and will not be included in the list of native interactions (despite their potential importance to the folding of the protein). Q has a maximum value of 1, indicating all native pairs are in contact and hence a folded state, whereas $Q = 0$ indicates a state where no native contacts are formed. Q can be used to distinguish between folded (f) and unfolded (u) ensembles in a two-state system by thresholding using the location of the free energy barrier between the two minima. We also identify the temperature of maximum variance in the fraction of native contacts, $\langle \Delta Q^2 \rangle = \langle Q^2 \rangle - \langle Q \rangle^2$, as the folding temperature of each variant (WT_u , WT_c , and *Mutant*) in both ensembles (dilute and crowded).

2. Another common order parameter is the radius of gyration, $R_g^2 = N^{-1} \sum_i (r_i - R_{com})^2$, where $R_{com} = N^{-1} \sum_i r_i$ and N is the number of residues and measures the mean squared distance between each residue and the center of mass of the chain. An equivalent definition is $R_g^2 = 1/2N^2 \sum_{ij} r_{ij}^2$, which involves a sum over all pairs of possible interactions (instead of only native interactions). In the SI we show that Q and R_g produce similar two-state statistics and the same categorizations of folded and unfolded states for SOD1, suggesting that either order parameter can capture the global behavior of the protein.
3. Q incorporates solely native interactions (ignoring potentially important electrostatic interactions), while

R_g incorporates all interactions (including potentially irrelevant noninteracting pairs). In this paper, we will define a subset of all pairs of interacting residues ($\mathbf{b}$) that includes both native and non-native interactions and define the order parameter $b = |\mathbf{b}|^{-1} \sum_{(ij) \in \mathbf{b}} \theta(c - r_{ij})$, with c as an arbitrary cutoff. b (much like Q) ranges between 0 and 1, with 0 meaning no residue pairs in the set are within the cutoff value from each other and 1 meaning all pairs are within the cutoff distance. For b to be a useful order parameter, the choice of residue pairs contained in set $\mathbf{b}$ must reflect the underlying physics of the protein; a randomly selected set of pairs would not give rise to a useful order parameter. The justification for which pairs of residues to include in the set of pairwise interactions included in $\mathbf{b}$ is described in Results, but will include pairs of residues that are involved in the formation of contact between specific β -strands, either due to specific interactions or to electrostatics.

- To better understand the structure of the folded or unfolded states of a protein, we compute the probability of contact between residues (with contact defined as a pair of residues found within a distance c between one another, with c the same cutoff distance used in the definition of b above). We denote p_{ij}^{cs} , p_{ij}^{ms} , and p_{ij}^{us} (with a lower-case “p”) as the probability of contact between residue i and j in the single (dilute) protein ensemble for the charged, mutant, and uncharged systems, respectively, and where s denotes either the folded (f) or unfolded (u) state ($s \in \{f, u\}$). We likewise denote the intracontact probabilities in the crowded ensembles as P_{ij}^{cs} , P_{ij}^{ms} , and P_{ij}^{us} (with an upper-case “P”) for the charged, mutant, and uncharged variants, with $s \in \{f, u\}$. These intraprotein contact probabilities will be useful to illustrate the structure of the folded and unfolded configurations that take electrostatic interactions into account.
- In the crowded case, proteins may interact with one another, and we can compute the probabilities $P_{ij}^{cs_1s_2}$, $P_{ij}^{ms_1s_2}$, and $P_{ij}^{us_1s_2}$ for the contact between residues i and j on different proteins, having $s_1 \in \{f, u\}$ as the state of one protein and $s_2 \in \{f, u\}$ as the state of the other protein. We choose the same cutoff distance c for the distance threshold for the interprotein interaction cutoff as we did for the intraprotein interaction cutoff above. We compute the probability of intercontacts between residues i and j from two different proteins by summing over the number of contacts between i and j without regard for order, resulting in the intercontact probabilities that are symmetric. These interprotein contact maps will be useful in understanding how self-crowders interact with one another and of the possible relevance for aggregation of SOD1.

3. RESULTS

3a. Addition of Charge and Self-Crowders Increases Folding Stability. We determine the folding temperature of the three protein variants using the time average of the variance of Q as a function of temperature (Figure 2A), for both single and crowded systems. We fit the data with Gaussian to identify the peak in the variance, indicating a continuous phase transition. From Figure 2B, the mean value of Q for each variant shows a transition from folded to

unfolded states with midpoints at the folding temperatures indicated in Figure 2A. For the single systems, the folding temperatures were $T_f \approx 157$, 158, and 158 K for the WT_u , *Mutant*, and WT_c systems, respectively. For the crowded systems, $T_f \approx 160$, 161, and 161 K for the WT_u , *Mutant*, and WT_c respectively. For both the single and crowded cases, we see a significantly lower folding temperature between WT_u and the other variants. This is due to the complete absence of charge in WT_u , whereas WT_c has unscreened electrostatic interactions between charged residues with an extra negative charge at residue 41 in *Mutant*. Crowding thus increases the folding temperature for all the protein variants due to the excluded volume of the crowders⁶² stabilizing the folded states in WT_u and due to the net effect of the charge of the crowders in WT_c and *Mutant*. We find a similar folding temperature of the variants using energy fluctuation as a function of time, shown in Figure S2.

3b. Free Energy as a Function of Q Indicates a Two-State System. From Figure 2A we see that the folding temperature of each variant differs approximately by 1 K for both the single and crowded ensembles. We thus performed simulations near their approximately equal folding temperatures: 157 K for the single protein ensembles and 161 K for the crowded protein ensembles. The free energies are shown in Figure 3. For each variant, there are two deep wells in the free

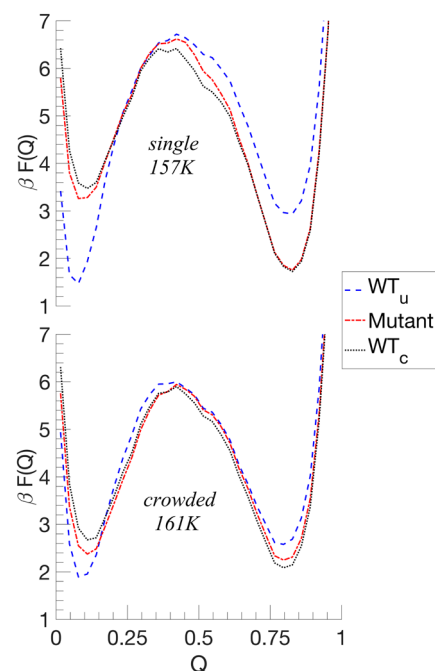


Figure 3. Free energy as a function of Q in single protein systems at 157 K (upper panel) and crowded protein systems at 161 K (lower panel) for WT_u (blue dashed line), *Mutant* (red dashed dot line), and WT_c (black dotted line). The energy barrier is maximum at about $Q = 0.42$, with two wells indicating distinct folded and unfolded states.

energy, corresponding to the folded (high Q values) and unfolded (low Q values) states using Q as a reaction coordinate. The absence of any additional local minima suggests that the folding and unfolding of SOD1 is a two-state system. Based on the locations of the peak of the free energy barrier, we define unfolded configurations to have $Q < 0.42$ and folded configurations to have $Q > 0.42$. The free energy as a function of the radius of gyration ($\beta F(R_g)$) is qualitatively

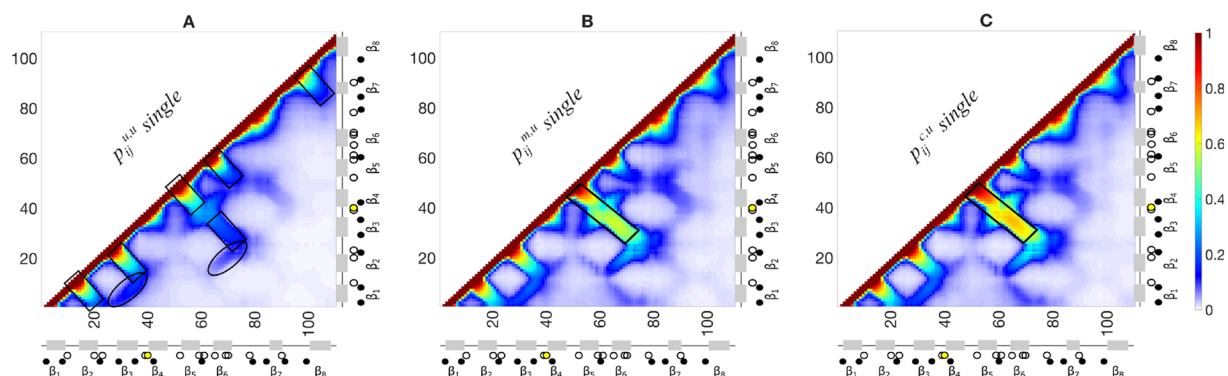


Figure 4. Intracontact probability values for the dilute (uncrowded) case. Panels A, B, and C show the probability of contact between unfolded proteins in the three ensembles for WT_u , *Mutant*, and WT_c . Axes show a schematic representation of β strands numbered 1–8. The filled black circles represent positively charged residues, the black empty circles represent negatively charged residues and the yellow filled circle represents residue 41, which is uncharged in WT_u and WT_c , but negatively charged in the *Mutant*. High values in the diagonal elements are due to the proximity of the residues in the protein chain structure. The greatest differences among the variants are in the interaction between residues 18–46 (β_3 , β_4) and 53–77 (β_5 , β_6 , and β bridge).

similar to Figure 3, validating our choice of Q to define folded and unfolded configurations (shown in Figure S1). From Figure 3 we see that there are differences between the free energy minima in folded and unfolded energy basins for all the variants in both the single and crowded ensembles because the folding temperatures of the variants are not exactly 157 and 161 K in the single and crowded cases, respectively.

3c. Probability of Intra- and Inter-Residue Contacts.

Comparison of the free energies for the uncharged, mutant, and charged variants do not indicate an obvious effect of the electrostatics beyond a change in the folding temperature. To better understand the effect of electrostatics, we compare the probability of contact between residues i and j for each variant in folded and unfolded configurations ($Q > 0.42$ and $Q < 0.42$, respectively) under single and crowded conditions. In this section, we consider a residue pair to be interacting if their distance is less than a cutoff value of $c = 1.5$ nm (the role of this cutoff is described in section 2d). The choice of 1.5 nm is arbitrary, with smaller values of the cutoff leading to noisier data (not shown). The average distance between native contacts is 6.4 Å, so this cutoff is a relatively large length-scale compared to the typical native interaction distance.

3c.1. Addition of Charge Increases Intracontact Probability between β Strands. Figure 4 shows the intracontact probability for the single protein variants at 157 K (close to the folding temperature for all three variants) for the unfolded configurations. Intracontact probabilities for the folded configurations do not show significant differences and are shown in Figure S5. In Figure 4, we see the unfolded intracontact probabilities which show numerous features indicating the organization of the unfolded state. We see a prominent dark red diagonal band with $P_{ij} = 1$ (which is the same for all the protein variants) due to the proximity of the residue pairs on the backbone. The other distinct regions in Figure 4 are the seven antidiagonal bands (perpendicular to the dark red bands, highlighted with boxes in Figure 4A), and two other small diagonal bands (highlighted by ovals in Figure 4A). The antidiagonal bands represent the interaction between β strands (from bottom left to top right: $\beta_1 - \beta_2$, $\beta_2 - \beta_3$, $\beta_3 - \beta_4$, $\beta_4 - \beta_5$, $\beta_5 - \beta_6$, $\beta_7 - \beta_8$) and loops in between the β strands. The small diagonal regions $\beta_1 - \beta_3$ and $\beta_2 - \beta_6$ and an antidiagonal one $\beta_3 - \beta_6$ represent the interactions between

the β strands not neighboring one another in the 1d chain structure. There is an additional interaction between $\beta_4 - \beta_7$ (not highlighted in Figure 4) with a lower probability of contact than the dominant $\beta_1 - \beta_3$, $\beta_2 - \beta_6$, and $\beta_3 - \beta_6$ contacts. These interactions are observed in all the protein variants, but a few are significantly reduced in WT_u . Those with a high probability of interaction ($\geq 20\%$) in WT_u must be formed due to nonelectrostatic effects (since charges are neglected in WT_u). We also see from Figure 4B,C that there is an additional antidiagonal band (not highlighted in Figure 4) representing interactions between $\beta_1 - \beta_6$ and $\beta_2 - \beta_5$, which are absent in WT_u . This increase in the probability of interaction must be due to the electrostatic effect in WT_c and *Mutant*. By comparing all the interactions highlighted by the intracontact probabilities, we see that there is a considerable increase in the contact probability between $\beta_4 - \beta_5$ and between $\beta_3 - \beta_6$ in *Mutant* (highlighted in Figure 4B) and WT_c (highlighted in Figure 4C) in comparison to WT_u .

Figure 4 shows a significant increase in the probability of interaction in the highlighted antidiagonal band. To better illustrate the effect of charge on the interactions between the residue pairs, we subtract $p_{ij}^{c,u}$ from $p_{ij}^{u,u}$, shown in Figure 5. Here, yellow and red represent positive values (residue pairs more likely to be found within 1.5 nm in WT_u than in WT_c) and green and blue signify negative values (residue pairs more likely to be found within 1.5 nm in WT_c than in WT_u). Overall, we see that the interaction probability between the β strands increases due to the addition of charge (highlighted in Figure 5 by the rectangular box). We found a similar pattern when computing the difference $p_{ij}^{m,u} - p_{ij}^{c,u}$ in the single ensemble as well as when we considered the differences in the crowded ensemble, shown in Figures S9 and S10.

The heightened contact probability in Figure 5 has only been demonstrated for the single protein case. To understand the impact of crowding we compared the differences in probabilities between WT_u and WT_c and *Mutant* and WT_c between the dilute and the crowded systems in Figure 6A and B, respectively. Here the yellow and red colors represent positive values, and hence, more interaction probability differences in crowded ensembles in comparison to single cases, whereas green and blue signify reduced interaction

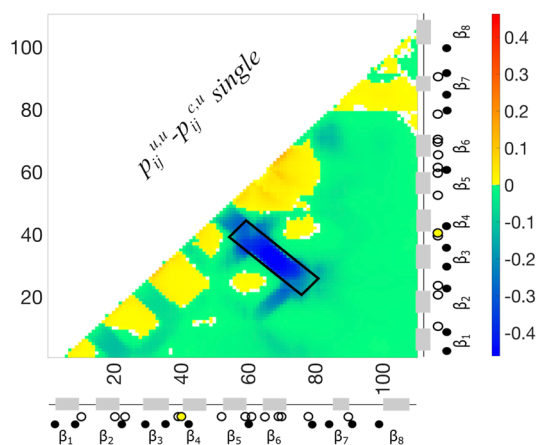


Figure 5. Comparison in contact probability for intracontacts in the uncrowded SOD1 system between residues from unfolded proteins between WT_u and WT_c . The major difference among the variants is in the interaction between residues 18–46 (β_3, β_4) and 53–77 (β_5, β_6 , and β bridge). The b region discussed in section 3b is defined by $\Delta p_{ij} = p_{ij}^{u,u} - p_{ij}^{c,u} < -0.2$. The interacting residue pairs are listed in Table S2.

probability differences. We see that crowding increased the differences in probabilities between WT_c and WT_u more in comparison to the differences in probabilities between WT_c and *Mutant* between $\beta_4 - \beta_5$ and $\beta_3 - \beta_6$. This is again due to the absence of charges in WT_u in contrast to *Mutant* and WT_c . The electrostatic effects from the crowders act alongside the steric effects to stabilize the unfolded configuration of *Mutant* and WT_c in comparison to WT_u .

3c.2. New Order Parameter, b , Indicates the Presence of Intermediate States in WT_c and *Mutant*. Figures 5 and 6 indicate a significant increase in the probability of contact between pairs of residues in $\beta_4 - \beta_5$ and $\beta_3 - \beta_6$ in the charged variants, which was further increased in the presence of crowding. To understand the effects of electrostatic interactions between these regions, we define a new order parameter, $b = |b|^{-1} \sum_{ij \in b} \theta(c - r_{ij})$ with b being a subset of interacting pairs as described in section 2d. The set b includes all pairs which are significantly more likely to be in contact in the presence of charge than the uncharged case, with $(p_{ij}^{u,u} -$

$p_{ij}^{c,u}) < -0.2$ (the blue region in Figure 5B). The order parameter b plays a role similar to the fraction of native contacts Q , but over a subset of both natively and non-natively interacting residue pairs (listed explicitly in Table S1; residues involved are also indicated in Figure S3). Figure 7 shows the

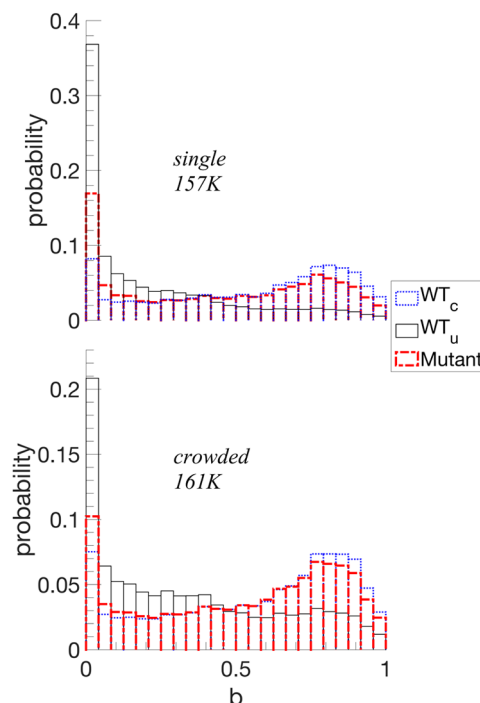


Figure 7. Normalized histogram of b for single systems (upper panel), crowded systems (lower panel). Blue represents WT_c , black represents WT_u , and red represents *Mutant*. There are two distinct peaks in *Mutant* and WT_c but not in WT_u for both single and crowded ensembles. This suggests the presence of a transition state even though it was not seen from the free energy as a function of Q plot (Figure 3 upper panel).

normalized histogram of b in single (top) and crowded (bottom) ensembles for WT_u (black solid line), *Mutant* (red dashed line), and WT_c (blue dotted line) for unfolded protein states ($Q < 0.42$). For both the single and crowded

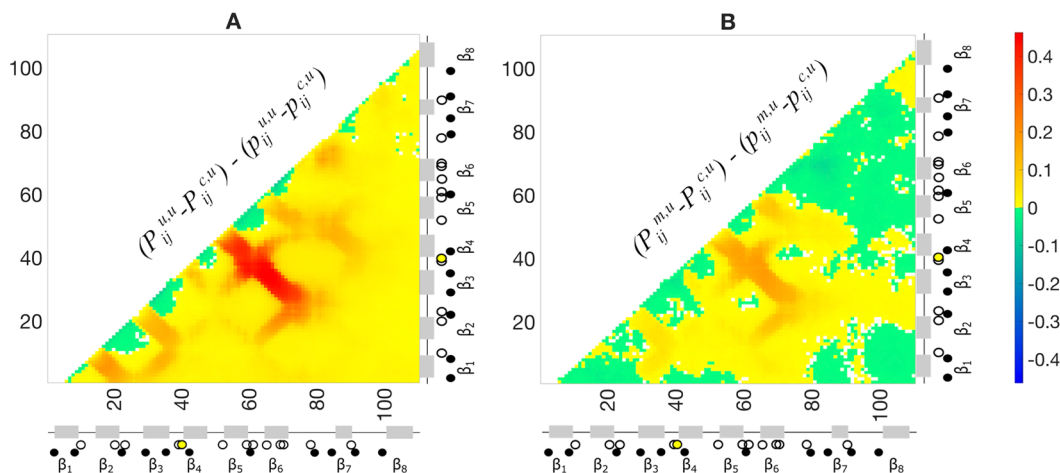


Figure 6. Comparison in probability difference values for intracontacts between residues from crowded and single unfolded proteins. Crowding increased $(P_{ij}^{u,u} - P_{ij}^{c,u})$ more in comparison to $(P_{ij}^{m,u} - P_{ij}^{c,u})$ between $\beta_4 - \beta_5$ and $\beta_3 - \beta_6$.

ensembles, we see that WT_u has a peak at $b = 0$ and the histogram is monotonically decreasing, while *Mutant* and WT_c have two distinct peaks at $b = 0$ and near $b \approx 0.8$. Thus, even though $\beta F(Q)$ does not indicate the presence of an intermediate state, this new order parameter b suggests *Mutant* and WT_c may have an intermediate state in the unfolded basin of attraction. The fact that $\beta F(Q)$ does not indicate an intermediate state might be because Q only depends on the interaction between the native state residue pairs, whereas b encompasses non-native interactions (including electrostatics) along with a few native interactions. In Figure S4, we also determine an effective radius of gyration $(R_g^*)^2 = \frac{1}{2|b|^2} \sum_{ij \in b} r_{ij}^2$ of the b -region as a function of b for the dilute simulations and find the typical distance between the residue pairs that define the b -region is ≈ 1 nm. When $b \approx 0.8$, most charged residues are found closest to an oppositely charged residue nearby in the sequence. The only exceptions to this pattern are the residue pairs 30–70, 36–66, and 43–53, despite their large separations in the sequence. Due to their large separation in the sequence, these residues may act as lynchpins to the formation of the intermediate configurations. This pattern is found in both the WT_c and *Mutant* variants, suggesting the mutation at 41 does not significantly alter the structure of the b region.

To illustrate the structure of the unfolded protein with and without the formation of the b -region, in Figure 8 we show a

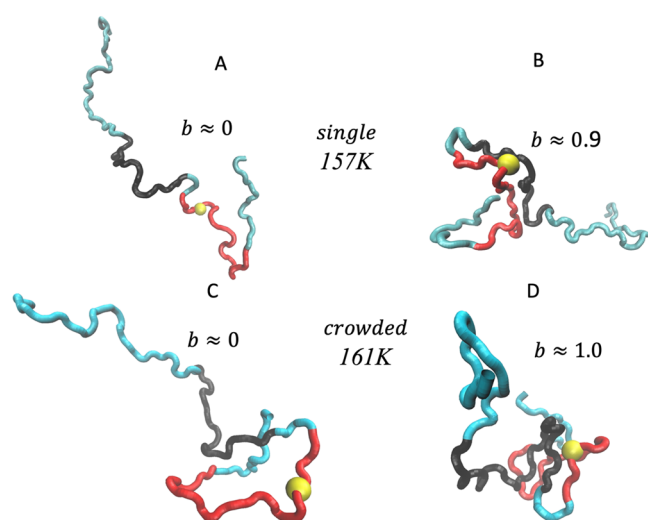


Figure 8. Representative snapshot of (*Mutant*) unfolded state for the single ensemble at 157 K (A, B) and crowded ensemble at 161 K (C, D). Red represents residues 18–46; black represents residues 53–77. The yellow sphere is residue 41. Cyan represents the rest of the protein. Panels A and C highlight the lack of interaction ($b \approx 0$) between the residues from the regions marked in red and black, whereas panels B and D highlight the interaction ($b > 0.8$) between the said regions.

representative snapshot of the unfolded *Mutant* with small and large values of b for single and crowded cases. The red color indicates the region containing β_3 and β_4 while the black indicates the part of the protein chain containing β_3 and β_6 . The yellow bead represents the site of the point mutation (Residue 41). In Figure 8A the protein resembles a random coil and has $Q = 0.0028$ and $b = 0$, and Figure 8B shows a representative snapshot of unfolded *Mutant* with $Q = 0.0972$

and $b = 0.8790$ (both in the single protein ensemble). Figures 8C and D show snapshots in the crowded ensemble with $Q = 0.0028$ and $b = 0.0019$, and $Q = 0.0972$ and $b = 0.9565$ respectively. The *Mutant* and WT_c variants have similar contact maps (Figures 4B,C and S9), and these snapshots from the *Mutant* simulations are also representative of the charged case.

Figure 7 suggests the presence of an intermediate defined by the order parameter b . Thus, we computed two-dimensional free energy as a function of b and Q ($F(b, Q)$), to get a more complete picture of the effect of charge in the folding of SOD1 variants. Figure 9A–C show the two-dimensional free energy for the single protein ensembles and Figure 9D–F show the free energy for the crowded ensembles. The insets show $\beta F(b, Q)$ for Q ranging from 0.14 to 0.16 and represent the free energy as a function of b for unfolded proteins. As seen in the insets of Figure 8, $\beta F(b, Q)$ for WT_u has a minimum at $b = 0$, as well as a weak well near $b \approx 0.8$ (with depth ≈ 0.5 kT), while two distinct minima are observed for the other protein variants, consistent with the distributions in Figure 7. For $Q > 0.42$, there is only one minimum in $\beta F(b, Q)$ at $b = 1$. We see that there are two distinct free energy minima at $b \approx 0$ and $b \approx 0.8$ in *Mutant* and WT_c in both single and crowded cases, while there is only one minimum in WT_u . This again suggests an intermediate defined by the reaction coordinate b , discussed further in section 4.

3c.3. Mutant's Interactions with Self-Crowders Differ from Wild Type. Figure S11 suggests that crowding decreases the intracontact probability between the β sheets in WT_u in comparison to the charged variants. To understand the effect of mutation under crowding, we next calculated the intercontact probabilities between residue pairs from unfolded–unfolded (U–U), unfolded–folded (U–F), and folded–folded (F–F) proteins (shown in Figures S12 and S13). The probabilities of residues being found within the cutoff distance of 1.5 nm are significantly lower between proteins, and thus to highlight the key differences in intercontact interactions due to the effect of the mutation, we implemented a threshold for the difference in the intercontact probabilities. In Figure 10, we show $(P_{ij}^{m;s1,s2} - P_{ij}^{c;s1,s2})/P_{ij}^{c;s1,s2}$ (that is, the difference between intercontacts in the *Mutant* and WT_c ensembles, normalized by the intercontact probability for WT_c). We only included 10% and 25% differences in these figures to highlight the regions where interprotein interactions differ most significantly. In Figure 10A, we see that residue 41 has three main interaction regions that are affected by the mutation (indicated by the red and the blue regions). There is a higher probability of interaction between residue 41 with residues 40 to 60 ($\beta_4 - \beta_5$) and a lower interaction with residues 1–18 (β_1) and 80–100 in WT_c in comparison to *Mutant*. This is also true for the U–F and the F–F interprobability contact, but with additional regions of elevated contact for the charged system (blue regions in Figure 10B,C, not discussed here). This difference in interaction is because residue 41 is negatively charged in *Mutant* but neutral in WT_c , so it prefers interaction with the positively charged residues 1 to 18 and residues 80 to 100 over residues 40 to 60, which have a nearly equal number of positive and negative charges in *Mutant*. However, residue 41 does not interact with all positive charges with higher probability in the *Mutant*, such as the residues 30 or 36 that are important in the formation of the b region. The point mutation thus has some interaction specificity between residues due to the constraints imposed by native interactions.

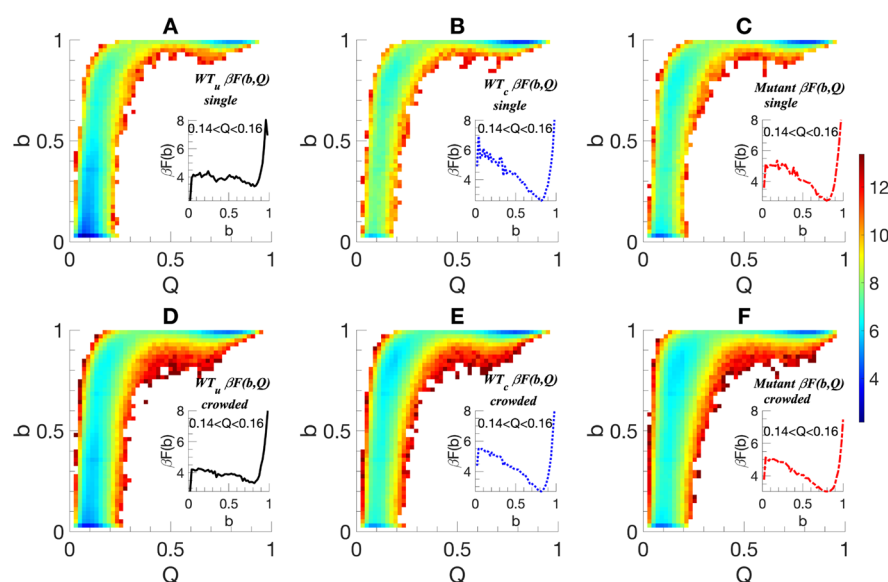


Figure 9. 2D projection of $\beta F(b, Q)$ for single and crowded systems at 157 and 161 K, respectively. Insets highlight the presence of two energy minima for WT_c and *Mutant* for Q values ranging between 0.14 and 0.16 corresponding to unfolded proteins. This folding pathway suggests that *Mutant* and WT_c may have an intermediate state.

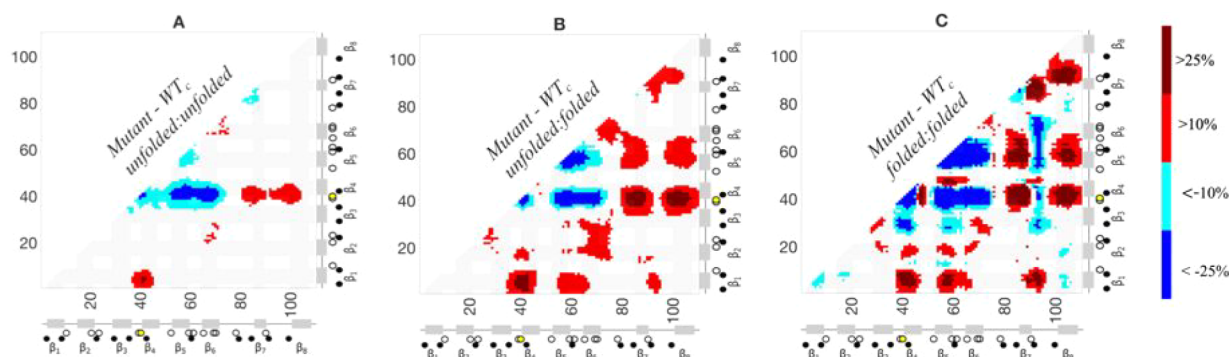


Figure 10. Thresholded relative differences in intercontact probability for the crowded SOD1 variants, WT_c and *Mutant* for (A) unfolded–unfolded, (B) unfolded–folded, and (C) folded–folded configurations. In comparison to WT_c dark red represents at least a 25% increase in contact probability in *Mutant*, light red at least a 10% increase, light blue at least a 10% decrease, and dark blue at least a 25% decrease. Axes show a schematic representation of β strands in gray numbered 1–8. The greyscale background in the lower triangle is a guide to identifying the β strands involved in the interaction shown by the probability of contact maps. The axes labels are the same as in Figure 4.

Figure 11 shows illustrative snapshots (generated using VMD) of these interprotein interactions. The top panels show the intercontact interactions in WT_c between residues 41 and 40–60 (the main blue region in Figure 10A). Residues 41–60 are highlighted in blue in one protein which is interacting with residue 41 (yellow sphere) on the other protein. The residues 80–100 are highlighted in red and are not interacting with residue 41 in panel 1. The middle panels show the interaction in *Mutant* between residues 41 and 80–100 (the two red regions in Figure 10A). The colors have the same meaning as in the top panel. The bottom panels highlight the interaction between residues 41 and 1–18 (in red) in *Mutant*. These snapshots are selected at random from the equilibrated protein trajectories with interactions between residues on distinct proteins that are in contact (within 1.5 nm) from the red and blue regions in Figure 10A. The cutoff of 1.5 nm is the same as was used for intraprotein contacts, with the justification for using this contact distance for interprotein interactions discussed in Supporting Information, S5.

4. DISCUSSION AND CONCLUSIONS

Formation of the b -Region in SOD1 is Sensitive to the Perturbation from Varying Ionic Strengths. Inside the highly packed cellular environment, macromolecular crowding gives rise to steric effects, which have been shown to have a stabilizing effect on the folded state of the proteins.⁶² Crowders may have additional interactions with proteins, including electrostatic effects, which may stabilize or destabilize the native state of the protein, depending on the distribution of charge.^{62,63} The net effect of the type of the crowders, ionic concentration, and surface charge distribution⁶⁴ will determine the protein stability. In our coarse-grained model, we investigated two extreme cases of ionic concentrations: an ionic strength of 0 in WT_c and infinity in WT_w ignoring the solvent charge screening to solely study the effects of charge on the proteins. Focusing on the extremes of the uncharged versus unscreened Coulombic interactions allowed us to identify a region of SOD1 (intermediate, defined by the reaction coordinate b) that is formed primarily due to the charged residues. However, it is important to acknowledge

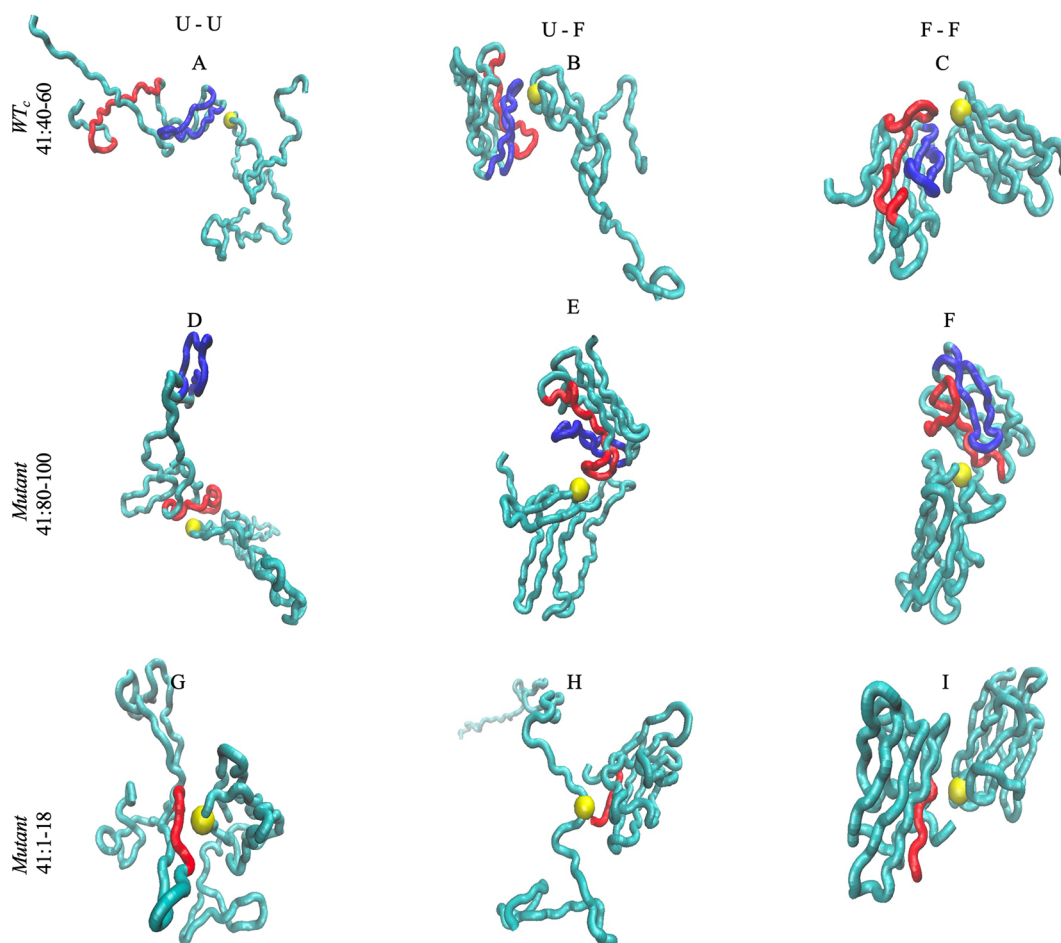


Figure 11. Representative snapshots of two interacting proteins from the crowded ensemble at 161 K. The top row shows WT_c and the bottom two rows represent *Mutant*. In panels A–F, the yellow sphere represents residue 41, residues 41–60 are colored red, and residues 60–100 are colored blue. In panels G–I, the yellow sphere represents residue 41m while residues 1–18 are represented by red. Panels A, D, and G show two unfolded protein configurations with residue 41 near the specified residues, panels B, E, and H show interactions between a folded and an unfolded configuration, and panels C, F, and I show interactions between two folded protein configurations. The first row emphasizes the preference of residue 41 to interact with residues 40–60 over residues 80–100 in WT_c , consistent with Figure 10A. The middle row emphasizes the preference of residue 41 to interact with residues 80–100 over 40–60 in *Mutant*, and the bottom row emphasizes the interaction of residue 41 with residues 1–18 in *Mutant* (consistent with Figure 10A).

that the effective solvent-charge screening length in a cellular environment is about 0.8 nm at room temperature,⁶⁵ which is neither of the extreme limits of vanishing nor infinite ionic strengths used in this paper. We found that once the intermediate was formed ($b \approx 0.8$) many of the charged residues in the b -region are found at an average distance of ≤ 9.5 nm, suggesting that screening effects will be moderate once the b -region is formed. We saw that for WT_c and *Mutant* in the single protein case the residue pairs 30–70, 36–66, and 43–53 are nearby for $b \approx 0.8$ despite their large separation in the sequence. Our simulations suggest that the b -region may be sensitive to the solvent ionic strength, but may play an important role in the folding pathway. In low ionic strength, the b -region may have an effectively longer screening length than under bulk conditions due to the volume excluded by the proteins (inaccessible to solvent) when they are in contact.

Adding Crowders Increases Native State Stability, While Mutation Destabilizes the Native State. In this paper, we studied the effect of self-crowding to gain a better insight into the self-assembly of proteins.⁶⁶ We analyzed the effect of crowding in two ways: by looking at how crowding affects the stability of one protein variant in the crowded

ensemble in comparison to the stability of other protein variants in the crowded case or by comparing the thermal stability of the protein variants between the single and the crowded ensembles. Comparing the wild type (WT_c) and *Mutant* within the single ensemble, we see from Figure 3 (top panel) that the free energy minima for the folded configurations (with high Q) are approximately the same at 157 K (which is very close to their folding temperature at approximately 158 K), while the unfolded basin is deeper for the *Mutant* than WT_c . In the crowded case (Figure 3 bottom panel), the depth of the folded well for WT_c is greater than for the *Mutant* at the folding temperature of 161 K, with the differences in depths in the unfolded basin unchanged by the addition of crowders. This suggests that the addition of self-crowders destabilizes the folded state of the *Mutant* in comparison to that of the WT_c in the crowded ensemble.

From Figure 2A we see that the folding temperatures of all the variants have increased by a few K with the addition of self-crowders in the crowded ensemble (bottom panel) in comparison to the single case (top panel). This suggests that even though the effects of crowding do not significantly change the folding temperature of the protein (consistent with the

previous studies^{15,24,25,58,62–64,67}), the variants were still stabilized in the crowded ensemble in comparison to the single case. Our finding is consistent with a previous study⁶² showing that crowding increases the thermal stability of proteins. In another recent paper,⁶⁸ using FRET imaging and molecular dynamics (MD) simulations Gnutt et al. studied the effect of quinary interactions on single point mutations of SOD1. They found the folding free energy of the G41D mutant in HeLa cells to be larger than for the G41D mutant in vitro, indicating that the intermolecular interactions present in the cellular environment destabilized the G41D mutation in SOD1. The wide range of quinary interactions that may be present in the complex cellular environment, including^{69,70} electrostatic interactions, hydrodynamic effects, and other nonspecific interactions, were not considered in this study and may explain why crowding is seen to stabilize the folded state for all of the variants. Destabilization of the folded state was observed using BSA as a crowding agent using experiments and molecular dynamics simulations in the Gnutt study, whereas we have found that self-crowding increases stability. A possible reason for this discrepancy arises from two differences between the crowders: (a) The folding temperature of BSA is well above the experimental conditions in the Gnutt study⁶⁸ (so that the crowders remain globular), whereas the self-crowders in our study are unfolded for a significant fraction of the time, and (b) BSA has a larger number of charged residues and greater magnitude of charge ($q = -13e$ at 10 mg/mL) at neutral pH⁷¹ than SOD1 ($q = -1e$), suggesting the increased strength of the destabilizing quinary interactions in BSA may explain the finding that self-crowding is stabilizing in this paper. An interesting future study would be to add additional nonspecific interactions (for example hydrodynamic or other solvent-mediated interactions) between SOD1 proteins to study the competition between folding and aggregation.⁷²

Formation of the *b*-Region in the Unfolded State Drives the Two-Dimensional Folding Pathway of SOD1.

In addition to the minor change in thermal stability of the protein variants, we found significant differences in intraresidue interaction probabilities for the unfolded configurations (with little change in the folded configurations; Figures 4 and S5–S7). This suggests that the effects of charge and crowding are more significant in the unfolded states compared to the folded states and led to the definition of the *b*-region in this paper. In another study,²³ Mojumdar et al. used force–extension experiments and coarse-grained simulation to show that the folding and unfolding of SOD1 are complex multistate processes with the presence of distinct intermediate states. In a more recent investigation,²⁴ Timr et al. used all-atom simulations in explicit solvent to identify the presence of intermediate states. There the authors showed that for SOD1 to fold, it should have β -sheet 1 (consisting of β -strands: β_1 , β_2 , β_3 , and β_6) and β -sheet 2 (consisting of β -strands: β_4 , β_5 , β_7 , and β_8). The two-dimensional free energy shown in Figure 9 is consistent with the two-dimensional folding pathway suggested there. In our paper, the intermediate defined by the *b* reaction coordinate primarily involves the interaction between $\beta_4 - \beta_5$ and $\beta_3 - \beta_6$. Starting from an unfolded configuration with $Q \approx 0$ and $b \approx 0$, Figure 9 suggests that the variants must pass through configurations with high *b* values, while they are still in the unfolded state (intermediate) and then fold (increase in *Q* with a high *b* value). This agrees with the 2D folded pathway suggested by Timr et al.

Mutation at 41 in the *b*-Region Alters the Interactions between Proteins Instead of within a Protein.

The negative charge in residue 41 in *Mutant* makes the net charge of the protein more negative than *WT*, and changes in the interprotein contacts involving residue 41 may give greater insights into the process by which dimerization and further aggregation begin. We found that the addition of the negative charge in the *Mutant* results in a decrease in the interactions of residue 41 with residues 80–100 compared to *WT*, with an increased probability of interaction of residue 41 with residues 1–18 (β_1) and 80–100 (β_7 , β_8). While it may be unsurprising that a negative charge interacts with regions containing positive charge, it is important to note that even though β_3 has two positive charges, residue 41 in *Mutant* does not have a high probability of contact with this region. One reason behind this might be the steric effects between β_6 and β_3 (Figure 4), suggesting the highlighted interaction of residue 41 with residues 80–100 is not solely due to electrostatics. β_3 is part of the *b*-region, and this suggests that the primary effect of the mutation is to alter the interaction between proteins, rather than change the folding landscape of individual proteins. Previous investigations have shown that the point mutation G41D in SOD1 is prone to prion-like aggregation and fibril formation,⁷³ associated with fALS. β_7 and β_8 are known to be the least stable and most flexible strands in SOD1 monomers.^{4,15,23,74,75} Thus, this switching of β -strand partners by residue 41 by the G41D mutant might shed new light on the mechanism behind the formation of fALS-associated aggregates.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.2c00819>.

Additional information about the structure of the folded configurations of SOD1, the specific definition of the *b*-region and additional details regarding contact probabilities in it, and the free energy as a function of R_g (rather than Q) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Greg Morrison – Department of Physics, University of Houston, Houston, Texas 77204, United States; Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States; orcid.org/0000-0002-1400-8092; Email: gcmorrison@uh.edu

Authors

Atsayee Sarkar – Department of Physics, University of Houston, Houston, Texas 77204, United States; Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States

Andrei G. Gasic – Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States; orcid.org/0000-0003-4054-6545

Margaret S. Cheung – Department of Physics, University of Houston, Houston, Texas 77204, United States; Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States; Pacific Northwest National Laboratory, Seattle Research Center, Seattle, Washington 98109, United States; orcid.org/0000-0001-9235-7661

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jpcb.2c00819>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (A.S., A.G., M.C., and G.M.) under NSF-PHY-2019745. A.G. is also supported by the Mathematical and Physical Sciences Ascending Postdoctoral Research Fellowship (MPS-Ascend) from the National Science Foundation Grant DMR-2137680. We thank Tim Burt, Lucas Babel, and Arya Datta for their participation in the early stage of the project. We thank Prof. Ebbinghaus and Dr. David Gnutz for the discussion on SOD1.

REFERENCES

- (1) Crapo, J. D.; Oury, T.; Rabouille, C.; Slot, J. W.; Chang, L. Y. Copper/Zinc Superoxide Dismutase Is Primarily a Cytosolic Protein in Human Cells. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 10405–10409.
- (2) Arnesano, F.; Banci, L.; Bertini, I.; Martinelli, M.; Furukawa, Y.; O'Halloran, T. v. The Unusually Stable Quaternary Structure of Human Cu,Zn-Superoxide 1 Is Controlled by Both Metal Occupancy and Disulfide Status. *J. Biol. Chem.* **2004**, *279*, 47998–48003.
- (3) Culik, R. M.; Sekhar, A.; Nagesh, J.; Deol, H.; Rumfeldt, J. A. O.; Meiering, E. M.; Kay, L. E. Effects of Maturation on the Conformational Free-Energy Landscape of SOD1. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E2546–E2555.
- (4) Teilum, K.; Smith, M. H.; Schulz, E.; Christensen, L. C.; Solomentssev, G.; Oliveberg, M.; Akke, M. Transient Structural Distortion of Metal-Free Cu/Zn Superoxide Dismutase Aberrant Oligomerization. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 18273–18278.
- (5) Sirangelo, I.; Iannuzzi, C. The Role of Metal Binding in the Amyotrophic Lateral Sclerosis-Related of Copper-Zinc Superoxide Dismutase. *Molecules* **2017**, *22*, 1429.
- (6) Khan, M. A. I.; Respondek, M.; Kjellstrom, S.; Deep, S.; Linse, S.; Akke, M. Cu/Zn Superoxide Dismutase Forms Amyloid Fibrils under Near Physiological Quiescent Conditions: The Roles of Disulfide Bonds and Effects of Denaturant. *ACS Chem. Neurosci.* **2017**, *8*, 2019–2026.
- (7) Zhu, C.; Beck, M. v.; Griffith, J. D.; Deshmukh, M.; Dokholyan, N. v. Large SOD1 Aggregates, Unlike Trimeric SOD1, Do Not Impact Cell in a Model of Amyotrophic Lateral Sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 4661–4665.
- (8) Sau, D.; de Biasi, S.; Vitellaro-Zuccarello, L.; Riso, P.; Guarnieri, S.; Porri, M.; Simeoni, S.; Crippa, V.; Onesto, E.; Palazzolo, I.; Rusmini, P.; Bolzoni, E.; Bendotti, C.; Poletti, A. Mutation of SOD1 in ALS: A Gain of a Loss of Function. *Hum. Mol. Genet.* **2007**, *16*, 1604–1618.
- (9) Pansarasa, O.; Bordoni, M.; Diamanti, L.; Sproviero, D.; Gagliardi, S.; Cereda, C. SOD1 in Amyotrophic Lateral Sclerosis: “Ambivalent” Behavior Connected to the Disease. *Int. J. Mol. Sci.* **2018**, *19*, 1345.
- (10) Huai, J.; Zhang, Z. Structural Properties and Interaction Partners of Familial ALS-Associated SOD1 Mutants. *Front. Neurol.* **2019**, *10*, 527.
- (11) Lindberg, M. J.; Bystrom, R.; Boknas, N.; Andersen, P. M.; Oliveberg, M. Systematically Perturbed Folding Patterns of Amyotrophic Lateral (ALS)-Associated SOD1 Mutants. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9754–9759.
- (12) Strange, R. W.; Yong, C. W.; Smith, W.; Hasnain, S. S. Molecular Dynamics Using Atomic-Resolution Structure Reveal Structural That May Lead to Polymerization of Human Cu-Zn Superoxide. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 10040–10044.
- (13) Schmidlin, T.; Kennedy, B. K.; Daggett, V. Structural Changes to Monomeric CuZn Superoxide Dismutase Caused by the Familial Amyotrophic Lateral Sclerosis-Associated Mutation A4V. *Biophys. J.* **2009**, *97*, 1709–1718.
- (14) Kumar, V. v.; Prakash, A.; Lynn, A. M. Alterations in Local Stability and Dynamics of A4V SOD1 in the presence of Trifluoroethanol. *Biopolymers* **2018**, *109*, e23102.
- (15) Ding, F.; Dokholyan, N. v. Dynamical Roles of Metal Ions and the Disulfide Bond in Cu, Zn Dismutase Folding and Aggregation. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 19696–19701.
- (16) Proctor, E. A.; Ding, F.; Dokholyan, N. v. Structural and Thermodynamic Effects of Post-Translational Modifications in Mutant and Wild Type Cu, Zn Superoxide Dismutase. *J. Mol. Biol.* **2011**, *408*, 555–567.
- (17) Ding, F.; Furukawa, Y.; Nukina, N.; Dokholyan, N. v. Local Unfolding of Cu, Zn Superoxide Dismutase Monomer Determines the Morphology of Fibrillar Aggregates. *J. Mol. Biol.* **2012**, *421*, 548–560.
- (18) Bille, A.; Jonsson, S. A. E.; Akke, M.; Irback, A. Local Unfolding and Aggregation Mechanisms of SOD1: A Monte Carlo. *J. Phys. Chem. B* **2013**, *117*, 9194–9202.
- (19) Das, A.; Plotkin, S. S. Mechanical Probes of SOD1 Predict Systematic Trends in Metal and Dimer of ALS-Associated Mutants. *J. Mol. Biol.* **2013**, *425*, 850–874.
- (20) Das, A.; Plotkin, S. S. SOD1 Exhibits Allosteric Frustration to Facilitate Metal Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 3871–3876.
- (21) Habibi, M.; Plotkin, S. S.; Rottler, J. Soft Vibrational Modes Predict Breaking Events during Force-Induced Unfolding. *Biophys. J.* **2018**, *114*, 562–569.
- (22) Harder, T.; Borg, M.; Bottaro, S.; Boomsma, W.; Olsson, S.; Ferkinghoff-Borg, J.; Hamelryck, T. An Efficient Null Model for Conformational Fluctuations in Proteins. *Structure* **2012**, *20*, 1028–1039.
- (23) Sen Mojumdar, S.; Scholl, Z. N.; Dee, D. R.; Rouleau, L.; Anand, U.; Garen, C.; Woodside, M. T. Partially Native Intermediates Mediate Misfolding of SOD1 in Single-Molecule Folding Trajectories. *Nat. Commun.* **2017**, *8*, 1881.
- (24) Timr, S.; Gnutz, D.; Ebbinghaus, S.; Sterpone, F. The Unfolding Journey of Superoxide Dismutase 1 Barrels under Crowding: Atomistic Simulations Shed Light on Intermediate States and Their Interactions with Crowders. *J. Phys. Chem. Lett.* **2020**, *11*, 4206–4212.
- (25) Gnutz, D.; Timr, S.; Ahlers, J.; Koenig, B.; Manderfeld, E.; Heyden, M.; Sterpone, F.; Ebbinghaus, S. Stability Effect of Quinary Interactions Reversed by Single Point Mutation. *J. Am. Chem. Soc.* **2019**, *141*, 4660–4669.
- (26) Solanki, A.; Neupane, K.; Woodside, M. T. Single-Molecule Force Spectroscopy of Rapidly Fluctuating, Marginally Structures in the Intrinsically Disordered Protein-Synuclein. *Phys. Rev. Lett.* **2014**, *112*, 158103.
- (27) Bokvist, M.; Gröbner, G. Misfolding of Amyloidogenic Proteins at Membrane Surfaces: The Impact of Macromolecular Crowding. *J. Am. Chem. Soc.* **2007**, *129*, 14848–14849.
- (28) Homouz, D.; Perham, M.; Samiotakis, A.; Cheung, M. S.; Wittung-Stafshede, P. Crowded, Cell-like Environment Induces Shape Changes in Aspherical Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 11754.
- (29) Zhou, Z.; Fan, J. B.; Zhu, H. L.; Shewmaker, F.; Yan, X.; Chen, X.; Chen, J.; Xiao, G. F.; Guo, L.; Liang, Y. Crowded Cell-like Environment Accelerates the Nucleation Step of Amyloidogenic Protein Misfolding. *J. Biol. Chem.* **2009**, *284*, 30148–30158.
- (30) Ma, Q.; Hu, J.-Y.; Chen, J.; Liang, Y. The Role of Crowded Physiological Environments in Prion and Prion-like Protein Aggregation. *Int. J. Mol. Sci.* **2013**, *14*, 21339–21352.
- (31) White, D. A.; Buell, A. K.; Knowles, T. P. J.; Welland, M. E.; Dobson, C. M. Protein Aggregation in Crowded Environments. *J. Am. Chem. Soc.* **2010**, *132*, 5170–5175.
- (32) Zimmerman, S. B.; Trach, S. O. Estimation of Macromolecule Concentrations and Excluded Volume Effects for the Cytoplasm of *Escherichia Coli*. *J. Mol. Biol.* **1991**, *222*, 599–620.

- (33) Asakura, S.; Oosawa, F. On Interaction between Two Bodies Immersed in a Solution of Macromolecules. *J. Chem. Phys.* **1954**, *22*, 1255–1256.
- (34) Minton, A. P. Excluded Volume as a Determinant of Macromolecular Structure and Reactivity. *Biopolymers* **1981**, *20*, 2093–2120.
- (35) Dhar, A.; Samiotakis, A.; Ebbinghaus, S.; Nienhaus, L.; Homouz, D.; Gruebele, M.; Cheung, M. S. Structure, Function, and Folding of Phosphoglycerate Kinase Are Strongly Perturbed by Macromolecular Crowding. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 17586–17591.
- (36) Gasic, A. G.; Boob, M. M.; Prigozhin, M. B.; Homouz, D.; Daugherty, C. M.; Gruebele, M.; Cheung, M. S. Critical Phenomena in the Temperature-Pressure-Crowding Phase Diagram of a Protein. *Phys. Rev. X* **2019**, *9*, 41035.
- (37) Zosel, F.; Soranno, A.; Buholzer, K. J.; Nettels, D.; Schuler, B. Depletion Interactions Modulate the Binding between Disordered Proteins in Crowded Environments. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 13480–13489.
- (38) Vancraenenbroeck, R.; Harel, Y. S.; Zheng, W.; Hofmann, H. Polymer Effects Modulate Binding Affinities in Disordered Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 19506.
- (39) Lin, Y.-H.; Forman-Kay, J. D.; Chan, H. S. Sequence-Specific Polyampholyte Phase Separation in Membraneless Organelles. *Phys. Rev. Lett.* **2016**, *117* (17), 178101.
- (40) Higgs, P. G.; Joanny, J. Theory of Polyampholyte Solutions. *J. Chem. Phys.* **1991**, *94*, 1543–1554.
- (41) Kantor, Y.; Kardar, M.; Li, H. Statistical Mechanics of Polyampholytes. *Phys. Rev. E* **1994**, *49*, 1383–1392.
- (42) Hu, J.; Chen, T.; Wang, M.; Chan, H. S.; Zhang, Z. A Critical Comparison of Coarse-Grained Structure-Based Approaches and Atomic Models of Protein Folding. *Phys. Chem. Chem. Phys.* **2017**, *19*, 13629–13639.
- (43) Huai, J.; Zhang, Z. Structural Properties and Interaction Partners of Familial ALS-Associated SOD1 Mutants. *Front. Neuro.* **2019**, *10*, 527.
- (44) Okado-Matsumoto, A.; Fridovich, I. Amyotrophic Lateral Sclerosis: A Proposed Mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9010–9014.
- (45) Crown, A. M.; Roberts, B. L.; Crosby, K.; Brown, H.; Ayers, J. I.; Hart, P. J.; Borchelt, D. R. Experimental Mutations in Superoxide Dismutase 1 Provide Insight into Potential Mechanisms Involved in Aberrant Aggregation in Familial Amyotrophic Lateral Sclerosis. *G3: Genes, Genomes, Genet.* **2019**, *9*, 719–728.
- (46) Prudencio, M.; Hart, P. J.; Borchelt, D. R.; Andersen, P. M. Variation in Aggregation Propensities among ALS-Associated Variants of SOD1: Correlation to Human Disease. *Hum. Mol. Genet.* **2009**, *18*, 3217–3226.
- (47) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (48) Clementi, C.; Nymeyer, H.; Onuchic, J. N. Topological and Energetic Factors: What Determines the Structural Details of the Transition State Ensemble and “En-Route” Intermediates for Protein Folding? An Investigation for Small Globular Proteins. *J. Mol. Biol.* **2000**, *298*, 937–953.
- (49) Bryngelson, J. D.; Wolynes, P. G. Spin Glasses and the Statistical Mechanics of Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 7524–7528.
- (50) Shoemaker, B. A.; Wang, J.; Wolynes, P. G. Structural Correlations in Protein Folding Funnels. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 777–782.
- (51) Clementi, C. Coarse-Grained Models of Protein Folding: Toy Models or Predictive Tools? *Curr. Opin. Struct. Biol.* **2008**, *18*, 10–15.
- (52) Kohn, J. E.; Millett, I. S.; Jacob, J.; Zagrovic, B.; Dillon, T. M.; Cingel, N.; Dothager, R. S.; Seifert, S.; Thiagarajan, P.; Sosnick, T. R.; Hasan, M. Z.; Pande, V. S.; Ruczinski, I.; Doniach, S.; Plaxco, K. W. Random-Coil Behavior and the Dimensions of Chemically Unfolded Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 12491–12496.
- (53) Piana, S.; Klepeis, J. L.; Shaw, D. E. Assessing the Accuracy of Physical Models Used in Protein-Folding Simulations: Quantitative Evidence from Long Molecular Dynamics Simulations. *Curr. Opin. Struct. Biol.* **2014**, *24*, 98–105.
- (54) Habibi, M.; Rottler, J.; Plotkin, S. S. As Simple As Possible, but Not Simpler: Exploring the Fidelity of Coarse-Grained Protein Models for Simulated Force Spectroscopy. *PLoS Comput. Biol.* **2016**, *12*, e1005211–e1005211.
- (55) Noel, J. K.; Levi, M.; Raghunathan, M.; Lammert, H.; Hayes, R. L.; Onuchic, J. N.; Whitford, P. C. SMOG 2: A Versatile Software Package for Generating Structure-Based Models. *PLoS Comput. Biol.* **2016**, *12*, e1004794.
- (56) Noel, J. K.; Onuchic, J. N. The Many Faces of Structure-Based Potentials: From Protein Folding Landscapes to Structural Characterization of Complex Biomolecules. In *Computational Modeling of Biological Systems: From Molecules to Pathways*; Dokholyan, N. v, Ed.; Springer US: Boston, MA, 2012; pp 31–54.
- (57) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
- (58) Anandakrishnan, R.; Aguilera, B.; Onufriev, A. v. H++ 3.0: Automating PK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res.* **2012**, *40*, W537–41.
- (59) Zegarra, F. C.; Homouz, D.; Eliaz, Y.; Gasic, A. G.; Cheung, M. S. Impact of Hydrodynamic Interactions on Protein Folding Rates Depends on Temperature. *Phys. Rev. E* **2018**, *97*, 32402.
- (60) Oliveira, R. J.; Whitford, P. C.; Chahine, J.; Wang, J.; Onuchic, J. N.; Leite, V. B. P. The Origin of Nonmonotonic Complex Behavior and the Effects of Nonnative Interactions on the Diffusive Properties of Protein Folding. *Biophys. J.* **2010**, *99*, 600–608.
- (61) Cho, S. S.; Levy, Y.; Wolynes, P. G. P versus Q: Structural Reaction Coordinates Capture Protein Folding on smooth Landscapes. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 586–591.
- (62) Cheung, M. S.; Klimov, D.; Thirumalai, D. Molecular Crowding Enhances Native State Stability and Refolding Rates of Globular Proteins. *Biophys. Comput. Biol.* **2005**, *102*, 4753–4758.
- (63) Sarkar, M.; Lu, J.; Pielak, G. J. Protein Crowder Charge and Protein Stability. *Biochemistry* **2014**, *53*, 1601–1606.
- (64) Ghosh, K.; de Graff, A. M. R.; Sawle, L.; Dill, K. A. Role of Proteome Physical Chemistry in Cell Behavior. *J. Phys. Chem. B* **2016**, *120*, 9549–9563.
- (65) Lin, S.-P.; Pan, C.-Y.; Tseng, K.-C.; Lin, M.-C.; Chen, C.-D.; Tsai, C.-C.; Yu, S.-H.; Sun, Y.-C.; Lin, T.-W.; Chen, Y.-T. A Reversible Surface Functionalized Nanowire Transistor to Study Protein–Protein Interactions. *Nano Today* **2009**, *4*, 235–243.
- (66) Fujiwara, N.; Wagatsuma, M.; Oba, N.; Yoshihara, D.; Tokuda, E.; Sakiyama, H.; Eguchi, H.; Ichihashi, M.; Furukawa, Y.; Inoue, T.; Suzuki, K. Cu/Zn-Superoxide Dismutase Forms Fibrillar Hydrogels in a pH-Dependent via a Water-Rich Extended Intermediate State. *PLoS One* **2018**, *13*, e0205090.
- (67) Bille, A.; Jensen, K. S.; Mohanty, S.; Akke, M.; Irbäck, A. Stability and Local Unfolding of SOD1 in the Presence of Protein Crowders. *J. Phys. Chem. B* **2019**, *123*, 1920–1930.
- (68) Gnutt, D.; Timr, S.; Ahlers, J.; König, B.; Manderfeld, E.; Heyden, M.; Sterpone, F.; Ebbinghaus, S. Stability Effect of Quinary Interactions Reversed by Single Point Mutations. *J. Am. Chem. Soc.* **2019**, *141*, 4660–4669.
- (69) McConkey, E. H. Molecular Evolution, Intracellular Organization, and the Quinary Structure of Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 3236–3240.
- (70) Guin, D.; Gruebele, M. Weak Chemical Interactions That Drive Protein Evolution: Crowding, Sticking, and Quinary Structure in Folding and Function. *Chem. Rev.* **2019**, *119*, 10691–10717.
- (71) Barbosa, L. R. S.; Ortore, M. G.; Spinozzi, F.; Mariani, P.; Bernstorff, S.; Itri, R. The Importance of Protein-Protein Interactions on the PH-Induced Conformational Changes of Bovine Serum Albumin: A Small-Angle X-Ray Scattering Study. *Biophys. J.* **2010**, *98*, 147–157.

(72) Dave, K.; Gasic, A. G.; Cheung, M. S.; Gruebele, M. Competition of Individual Domain Folding with Inter-Domain Interaction in WW Domain Engineered Repeat Proteins. *Phys. Chem. Chem. Phys.* **2019**, *21*, 24393–24405.

(73) Chan, P. K.; Chattopadhyay, M.; Sharma, S.; Souda, P.; Gralla, E. B.; Borchelt, D. R.; Whitelegge, J. P.; Valentine, J. S. Structural Similarity of Wild-Type and ALS-Mutant Superoxide Dismutase-1 Fibrils Using Limited Proteolysis and Atomic Force Microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 10934–10939.

(74) Nordlund, A.; Oliveberg, M. Folding of Cu/Zn Superoxide Dismutase Suggests Structural Hotspots for gain of Neurotoxic Function in ALS: Parallels to Precursors in Amyloid. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10218–10223.

(75) Yu, H.; Liu, X.; Neupane, K.; Gupta, A. N.; Brigley, A. M.; Solanki, A.; Sosova, I.; Woodside, M. T. Direct Observation of Multiple Misfolding Pathways in a Single Prion Molecule. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 5283–5288.

Recommended by ACS

Exploring the Interplay between Disordered and Ordered Oligomer Channels on the Aggregation Energy Landscapes of α -Synuclein

Xun Chen, Peter G. Wolynes, *et al.*

JULY 10, 2022

THE JOURNAL OF PHYSICAL CHEMISTRY B

[READ !\[\]\(d3b4f22af99c507f55d7924c8d6d7349_img.jpg\)](#)

Delineating the Aggregation-Prone Hotspot Regions (Peptides) in the Human Cu/Zn Superoxide Dismutase 1

, Md. Imtaiyaz Hassan, *et al.*

DECEMBER 03, 2021

ACS OMEGA

[READ !\[\]\(845d2e325a1bb70ecff27a3a373459f0_img.jpg\)](#)

Analytical Approaches for Deriving Friction Coefficients for Selected α -Helical Peptides Based Entirely on Molecular Dynamics Simulations

Aleksandra Wosztal, Robert Szoszkiewicz, *et al.*

OCTOBER 27, 2022

THE JOURNAL OF PHYSICAL CHEMISTRY B

[READ !\[\]\(7eab21365973e986874cef4081771357_img.jpg\)](#)

Binding Models of A β 42 Peptide with Membranes Explored by Molecular Simulations

Ke Wang, Wensheng Cai, *et al.*

AUGUST 19, 2022

JOURNAL OF CHEMICAL INFORMATION AND MODELING

[READ !\[\]\(d1f70f1feee1d35b8823b4642f89723e_img.jpg\)](#)

[Get More Suggestions >](#)