THE JOURNAL OF PHYSICAL CHEMISTRY B

pubs.acs.org/JPCB Article

Exploring Folding Aspects of Monomeric Superoxide Dismutase

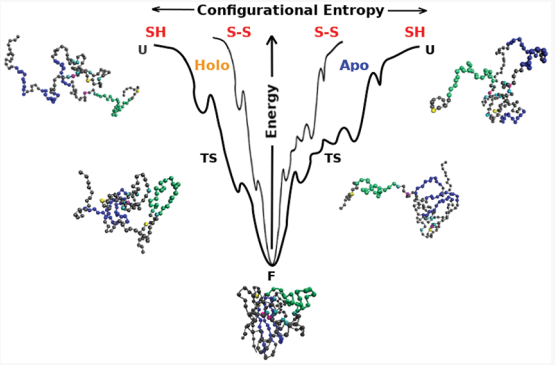
Published as part of The Journal of Physical Chemistry virtual special issue "David N. Beratan Festschrift".

Paulo R. Mouro, Ana P. R. Povinelli, Vitor B. P. Leite,* and Jorge Chahine*

Cite This: *J. Phys. Chem. B* 2020, 124, 650–661 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Recent studies have associated the absence of bound metals (Apo protein) and mutations in Cu–Zn Human Superoxide Dismutase (SOD1) with amyotrophic lateral sclerosis (ALS) disease, suggesting mechanisms of SOD1 aggregation. Using a structure-based model and modifying the energy of interaction between amino acids in the metal-binding site, we detected differences between the folding of the apo and holo proteins. The presence of metal ions decreases the free-energy barrier and also suggests that the folding pathway may change to reach the native state. The kinetics of folding of the apo and holo forms also correlates with the amount of free-energy barrier in the folding process. Also, the stability of the native state is significantly affected by the absence of metal ions. Our results, obtained from a very simplified model, correlate with more detailed studies, which also have shown that the transition and the native states are affected by the absence of the metal ions, hindering the folding of SOD1 and decreasing the stability of the native state. Regarding the disulfide bond, the results show that its absence decreases the stability of the native structure but affects the transition state less, suggesting that it is possibly made late in the folding process.



INTRODUCTION

Zn–Cu Human Superoxide Dismutase (SOD1) is a homodimeric enzyme in which each monomer contains 153 amino acids, 1 copper ion, and 1 zinc ion.[1−3] The dimeric SOD1 is the form present in most organisms. A rare exception is the monomeric SOD1 from the periplasm of *Escherichia coli*, which was extracted and purified ∼20 years ago,[4] when it was shown that the monomeric form is retained even at a large concentration of protein. This enzyme, in its native form, protects cells against oxidative damage and promotes the catalysis of the dismutation of superoxide radicals into molecular oxygen and hydrogen peroxide.[5,6] Zn and Cu ions, in addition to playing an important role in SOD1 enzymatic activity, also play an essential part in the stability of the native enzyme, as has been shown by experiments involving chemical or temperature denaturation.[6−10] Moreover, in the absence of metals, apoSOD1 is more prone to monomerization.[11,12] The presence of metals is relevant for the dimerization process and also for keeping two key loops properly structured for the enzyme function.[13] Hydrophobic interactions stabilize the dimeric form, burying ∼550 Å^2 of hydrophobic surface area at the interface of the two monomers,[2] a value roughly similar to the surface area of the active site, which is ∼10% of the total surface area of the molecule.[14] Loop IV, the metal-binding loop, and loop VII, the electrostatic loop, held together by a disulfide bond, form a channel that links the active site to the surface of the enzyme and creates a suitable electrostatic gradient that guides the superoxide anion to the metal site, where the catalysis takes place; see Figure 1. In addition, a study by Ciriolo et al.[15] suggests that the conserved residues in the electrostatic loop are crucial for an optimal copper uptake by the enzyme rather than the need for fast dismutation by acceleration of the substrate O_2^- into the active site. The conserved residue arginine 143 would be the main element responsible for attracting the superoxide to the active site. Neutral substitutions of this residue by Ala or Ile reduced the rate 10-fold, but replacements by residues of opposite charge, such as Asp and Glu, made the rate decrease ∼100-fold.[16]

As with other metalloenzymes or oligomeric proteins, SOD1 has a complex folding mechanism due to a long post-translational maturation process.[21] Fully mature, disulfide-oxidized SOD1 is highly stable with folding temperatures >90 °C.[22] The absence of metal ions in the apo form decreases the melting temperature to ∼52 °C, and the reduced disulfide bond decreases it to ∼42 °C.[23] All of these maturation steps related to the structure and folding of SOD1 have been extensively studied in the last several decades, and a large number of publications have shown evidence that SOD1

Received: October 13, 2019
Revised: January 2, 2020
Published: January 3, 2020

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650–661

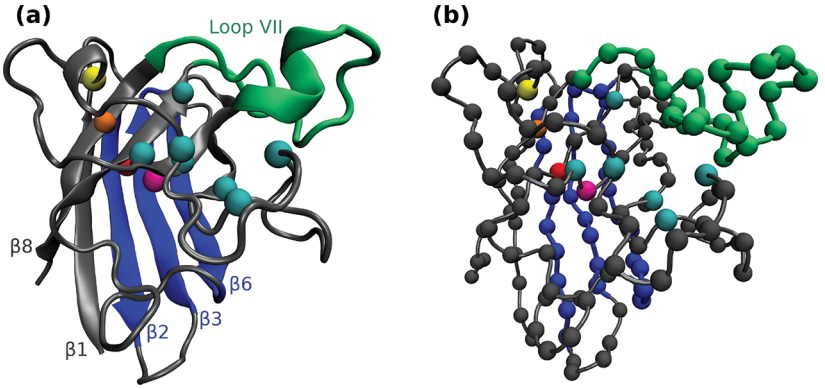


Figure 1. Schematic representation of native SOD1 monomer (PDB code 1HL5, chain A). (a) Representation of secondary structures according to the SCOP criterion.[17] Results have shown that β-sheets 2, 3, and 6 have a higher propensity to form aggregates,[18] and these regions are highlighted in blue in this figure. The electrostatic loop is shown in green (loop VII, residues 121−142). The cyan spheres represent the C_α carbon of the residues that bind to copper and zinc. The yellow and orange beads represent the C_α carbon of cysteines 57 and 146, respectively. The pink and red beads indicate the location of the residues 47 and 117, respectively, which have high ϕ values according to refs 19 and 20. (b) Monomeric SOD1 representation at the C_α level of simplification. The colors follow the same pattern described in panel a.

mutants can aggregate and lead to amyotrophic lateral sclerosis (ALS) disease. More than 100 mutations in the SOD1 gene have been associated with this pathological condition.[24,25] Mutations, along with the absence of Cu and Zn ions, have been regarded as possible agents for the SOD1 pathogenic behavior.[26]

Computational simulations have helped in the understanding of how proteins fold.[27−31] The use of simplified models that take into account the energy landscape theory, such as the structure-based model (SBM),[32,33] has been able to capture some essential characteristics of the protein-folding problem[32,34−37] and is in agreement with other analytical and experimental methods.[33,38−45]

The objective of this study is to obtain a coarse-grained description of monomeric Human Superoxide Dismutase folding. Kinetics and thermodynamics will be addressed, focusing mainly on the steps related to the maturation process, which include metal binding (Zn and Cu ions) and the disulfide bond (C57−C146). The simplified model used in this work is a standard C_α SBM (Cα-SBM).[32] One of the motivations for using such a simple model is that the apo and disulfide reduced forms are similar to the holo structure in terms of the backbone, as shown by the alignment of the holo (PDB code 1HL5), apo (1HL4), and reduced disulfide (2AF2) structures. By inserting or removing contacts energies between specified C_α atoms, which represent the residues in the simplified model, we get a coarse-grained insight into features of these forms, which can be correlated, at least qualitatively, with results from more detailed theoretical and experimental studies.

COMPUTATIONAL DETAILS

The folding process of SOD1 was studied by computational simulations using the Cα-SBM.[32] In this model, each residue is represented by a bead at the Cα position. For a protein with native conformation Γ^0, the potential energy of any configuration Γ is given by

$$V(\Gamma, \Gamma^0) = \sum_{\text{bonds}} \epsilon_r(r - r_0)^2 + \sum_{\text{angles}} \epsilon_\theta(\theta - \theta_0)^2 + \sum_{\text{backbone}} \epsilon_\phi \left\{ [1 - \cos(\phi - \phi_0)] + \frac{1}{2}[1 - \cos(3(\phi - \phi_0))] \right\} + \sum_{\text{contacts}} \epsilon_c \left[5\left(\frac{d_{ij}}{r_{ij}}\right)^{12} - 6\left(\frac{d_{ij}}{r_{ij}}\right)^{10} \right] + \sum_{\text{non-cont}} \epsilon_{NC}\left(\frac{\sigma_{NC}}{r_{ij}}\right)^{12} \quad (1)$$

where r_0 is the distance between two α carbons in the native structure, θ_0 is the angle formed by three residues (in sequence) in the native conformation, and ϕ_0 is the dihedral angle formed by four consecutive α carbons. The fourth term (10−12 Lennard-Jones potential) describes the long-range interaction between $C_\alpha(i)$ and $C_\alpha(j)$, where d_{ij} is the distance between carbons i and j in the native structure. Native contacts were obtained by the Shadow Map algorithm.[46] To verify if a native contact is being made at a certain step of the simulation, we have used as a criterion the distance between the amino acids of the respective contact in this conformation. If it is $<1.2d_{ij}$, then it is considered that the contact is made. The last term of the potential refers to all α-carbons that do not make a native contact, using σ_{NC} = 4 Å as the minimum approach distance. The other constants, ϵ_r, ϵ_θ, ϵ_ϕ, and ϵ_{NC}, are all in ϵ_c units, and their values are, respectively, 100, 20, 1, and 1.[32,37] The value of ϵ_c, which is 1.0 kJ/mol, scales the energy and also the temperature, which is measured as a reduced unit, $T = \epsilon_c/k$, where k is the Boltzmann constant, taken as 1.

We used GROMACS-4.5.5[47] to perform the simulations and the SMOG online tool[48] to generate the input files. The thermodynamics of folding was obtained using the replica-exchange method (REMD), making use of 24 "replicas", whereas each replica is at a different temperature. More details of the temperature and replica-exchange sampling of each system can be found in the Supporting Information. Each temperature was simulated at 5×10^8 steps, using an integration step of 0.5×10^{-15} s. Conformations were stored every 2500 steps, and exchanges between neighboring replicas were allowed at every 10 000 simulation steps. To study the kinetics aspects (folding time), we performed 200 simulations without REMD at low temperature (around room temperature), and the kinetics analysis (fraction unfolded versus time) was described through the average of the 200 simulations. The folding temperature was considered the temperature at which

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661

the free energies of the folded and unfolded states have the same value.

Using the native holo structure from chain A of the PDB code 1HL5 to build our monomeric SOD1 model, we try to describe the folding process of five monomeric systems, three apo and two holo forms, modeling each one just by changing some contact energies around the residues next to the metal site, or the C57−C146 disulfide bond. More details of the native contact map for each system can be found in Figure S1 in the Native Contact Map section of the Supporting Information. The monomeric systems are labeled and characterized as follows: (i) $mApo_{SH}$: for the apo reduced form, in which cysteines are protonated and do not interact, so there is no favorable energy associated with their contact; (ii) $mApo_{conf}$: for the apo form with a conformational reversible disulfide bond; that is, the cysteines have favorable contact energy associated with their contact but with strength such that it can be undone throughout the simulation; (iii) $mApo_{SS}$: for the apo disulfide oxidized form, which means a covalent (permanent) bond is formed between the cysteines; (iv) $mHolo_{SS}$: for the holo disulfide oxidized form, with the same covalent bond as in the previous SS case; and (v) $mHolo_{conf}$: for the holo with a conformational reversible disulfide bond (analogous to the previous conf case). Although metals are not explicitly present in these models, the simulations designated as "Holo" are associated with existent native interaction energies between amino acids in the Zn and Cu binding sites. On the contrary, in all Apo cases, the contacts among residues around the metal site were removed. The changes in the disulfide bridge between C57 and C146 are indicated by the indices conf, SS, and SH. In the conf case ($mApo_{conf}$ and $mHolo_{conf}$), the bond was modeled like any other native contact, being governed by the potential "10−12" of the Cα-SBM force field. In the SS cases ($mApo_{SS}$ and $mHolo_{SS}$), the disulfide bond was always present, and it was modeled by a harmonic potential.

The reason and motivation to perform the study in different ways to form the disulfide bond is due to the significant amount of experimental data, mainly for the apo monomer, in the reduced and oxidized forms. Moreover, the disulfide bond mechanism is still a matter of discussion in the SOD1 maturation process. Table 1 summarizes the five systems studied.

Table 1. Studied Systems

system	MS[a]	C57−C146[b]
$mApo_{SH}$	no	reduced
$mApo_{conf}$	no	conformational
$mApo_{SS}$	no	oxidized
$mHolo_{conf}$	yes	conformational
$mHolo_{SS}$	yes	oxidized

[a]MS indicates the presence of the metal-site contact energies. [b]Indicates the type of interaction between C57 and C146.

In addition to the native contact fraction (Q), commonly used as an order parameter in protein-folding simulations,[36,37] we also used β_T[49] as a reaction coordinate, which is given by the difference between the radius of gyration of the unfolded R_g^U, the native R_g^N, and any other state (R_g)

$$\beta_T = \frac{a_r + b_r(R_g^U - R_g)}{a_r + b_r(R_g^U - R_g^N)} \quad (2)$$

where a_r and b_r are obtained experimentally depending on the denaturant that is used to follow the folding/refolding process. In our case, we follow the values determined by Geierhaas et al.[49] for the urea as a denaturant. Like the native contact fraction (Q), β_T is also normalized, with values between 0 and 1. Q or β_T equal to 0 means that the protein is completely unfolded, and closer to 1, the protein approaches its native structure. From the radius of gyration, calculated using the GROMACS 4.5.5 plugin (g_gyrate), the value of β_T is obtained, allowing us to estimate the m values, which describe the dependence of the unfolding free energy on the denaturant concentration. m values are interesting quantities because they are strongly correlated with the amount of buried hydrophobic surface in the folding process. These values also quantify the cooperativity of folding and unfolding transitions. The relation between β_T and m values is given by[49]

$$\beta_T = \frac{m_{D\text{-}TS}}{m_{D\text{-}N}} \quad (3)$$

where D-TS is the conformational change between the denatured state and the transition state, D-N is the transition between the denatured state and the native state, and the m values describe the amount of buried hydrophobic surface and also the cooperativity related to these transitions. The m values, obtained from experiments, inserted in eq 3 give β_T, which will be compared with the theoretical values from simulations. This procedure will be used to probe how well the transition-state ensemble from simulations correlates with experimental data. The fact that a structure-based potential relies on the native structure is crucial to check if conformations, other than those belonging to the native ensemble, are being reliably probed.

RESULTS AND DISCUSSION

Thermodynamics of Monomeric SOD1. This section describes the thermodynamics of the SOD1 monomer folding. To make comparisons between experimental and simulation data, it is necessary to find a correspondence between the real temperatures (in Kelvin) and the reduced temperatures used in the simulations. An obvious correspondence is obtained from the folding temperature, T_f, at which the simulated system shares one thermodynamic feature with the real system: zero stability. Then, it would be simple to find a reduced temperature that could correspond to a real temperature by associating the T_f value from simulations with the T_f from experimental thermodynamic data, usually known as the melting temperature, T_m. However, the correspondence between reduced and real temperatures to find the room temperature in reduced units is not that simple.

The linear relation used to find the room temperature in reduced units, which takes only the melting temperature as a reference, yields very low reduced temperatures. As a consequence, it leads to very high unfolding barriers. This is not reasonable for moderate values of the stability. In the present study, the reduced temperatures that correspond to real room temperatures are chosen according to two reference temperatures: one at which the stability is zero (T_f) and another where the unfolding barrier is zero (T_u). This is not always possible, but for the case of the oxidized apo

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661

monomer,[50] the rates for folding (unfolding) are obtained as a function of temperature, and the value T_u = 370 K completely removes the unfolding barrier. Along with the folding temperature T_f = 324 K, for which the stability is zero, it is possible to propose a linear scale $T^* = a^*T + b$ between the reduced temperature T^* and T in Kelvin. For the other two forms, the reduced apo monomer and the oxidized holo monomer, we could not find enough data to calculate the constants a and b; we estimated their "room temperature" by taking their known values of T_f and using the room temperature of the oxidized apo monomer as a reference.

By adopting the above criteria to find the corresponding simulation room temperature, it is possible to compare our computational thermodynamics results with experimental ones. Figure 2 shows the free-energy profile at room temperature for the three cases, from which we can draw the comparison: $mApo_{SH}$, $mApo_{SS}$, and $mHolo_{SS}$. In the simulations of all of the monomeric forms, the native structure used to build the model is the same, as the crystallographic structures are very similar. Then, in our modeling, the difference in each system is just some contact energies around the residues next to the disulfide bond, or in the metal site, as explained in the Computational Details section. In this case, the free-energy profile in Figure 2 indicates that under the same disulfide bond condition, the holo form is always more stable than the apo one. This fact shows the importance of metals, represented, in the present coarse-grained model, by the contact energies among the residues that coordinate with the Zn and Cu ions in the native structure.

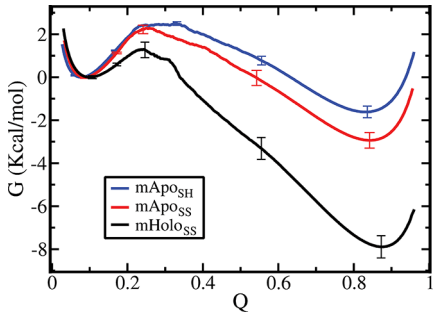


Figure 2. Thermodynamics of monomeric systems. Free-energy profile according to the fraction of native contacts Q at "room temperature". Each color represents one monomeric system and follows the pattern: $mApo_{SH}$ (blue curve), $mApo_{SS}$ (red curve), and $mHolo_{SS}$ (black curve). The curves were obtained by applying the WHAM[51,52] method on the 24 replicas of the temperature of each system. Error bars were calculated by splitting the simulation data into thirds and calculating the variation between the three splits. The color of each error bar follows the same pattern of each free-energy curve.

All of the free-energy profiles in Figure 2 show a typical two-state protein-folding process without the presence of metastable intermediate states. The peak of the barrier located at a small Q value ($Q \approx 0.25$) means relatively high energy at the transition state because only 25% of the contacts are made. The predominance of enthalpy in relation to entropy for the formation of the free-energy barrier of SOD was detected by Kayatekin et al.[50] using experimental data. In that study, the authors showed a favorable entropy of activation due to interaction with the solvent. However, the high enthalpic cost surmounted the change in the entropy, so that the barrier has an enthalpic character. In the simulations of the present study, the activation entropy is unfavorable because only configurational entropy is involved, which is reduced at the transition state. Then, the barrier clearly has an entropic character, but its relatively high value is due to the small decrease in the enthalpy caused by the reduced number of contacts (only 25%) made at the transition state.

The apo monomer in the presence of zinc shows a significantly higher rate of folding.[53,54] The experimental folding rates of the apo and metalled (zinc) forms in the mentioned reference are 0.056 and 0.75 s^{-1}, respectively, which means that zinc binding lowers the barrier by ~1.6 kcal/mol. Regarding the stability, it is increased by ~8 kcal/mol. By comparing the values from the red (apo) and black (holo) curves of Figure 2, the values regarding the difference in barrier and stability are 1.3 and 5 kcal/mol, respectively. The indicates a reasonable agreement for the barrier, whereas for the stability, the agreement is not as good.

The blue curve in Figure 2 represents the free-energy profile of the reduced apo monomer, $mApo_{SH}$, whose stability decreases 1.3 kcal/mol with respect to the oxidized apo form, represented by the red curve. Besides a lower stability, the reduced form has a slightly higher free-energy barrier, ~0.3 kcal/mol, than the oxidized $mApo_{SS}$. Data from Kayatekin et al.[55] show that the presence of a disulfide bond accelerates folding by a factor of two, which means that the free-energy barrier of the reduced form is higher by a factor of 0.4 to 0.5 kcal/mol. Their study showed a lower stability of the reduced form, ~1.5 kcal/mol less than the oxidized form. The simulations also show a reduction of the stability on the order of 1.3 kcal/mol of the reduced form relative to oxidized one. It must be noted that absolute values of the barriers and stabilities from simulations do not show a very good agreement with experimental results. For instance, the experimental value for the folding barrier of the oxidized apo monomer is ~13 kcal/mol, with a stability of 4.8 kcal/mol,[55] but the red curve of Figure 2 for the oxidized apo monomer shows corresponding values of 2.2 and 2.9 kcal/mol for the barrier and stability, respectively. Only the difference between quantities $\Delta\Delta G$ has shown a reasonable agreement between the simulations and the experimental data, possibly due to the cancellation of solvent effects.

The 1.5 kcal/mol[55] lower stability of the reduced apo form is less than the value suggested by a polymer model that estimates the contribution of a disulfide bond to the change in conformational entropy of a protein,[56] given by

$$\Delta S = -2.1 - (3/2)R \ln(n) \quad (4)$$

where n is the number of residues in the loop formed by the disulfide bond and R is the universal gas constant. For SOD1, n is 89, and $T\Delta S$ at room temperature is ~4.6 kcal/mol, suggesting that the effect of the disulfide bond is more complicated than just to decrease the entropy of the unfolded state when compared with the reduced form. Reference 50 shows the enthalpy and entropy of folding at 298 K of the disulfide intact form $\Delta H(\text{oxi})$ = 28.9 kcal/mol and $T\Delta S(\text{oxi})$ = 24.1 kcal/mol, equivalent to a stability value of 4.8 kcal/mol. From ref 57, these numbers are calculated for the reduced form but at 310 K. By using the Kirchoff approximation,[58] the values at 298 K are $\Delta H(\text{red})$ = 37.8 and $T\Delta S(\text{red})$ = 34.4 equiv to a stability value of 3.4 kcal/mol, consistent with values from other studies.[57,59] From these numbers, it is possible to calculate the differences between the reduced and oxidized forms $\Delta\Delta H$ = 8.9 and $T\Delta\Delta S$ = 10.3, whose difference is −1.4

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661

kcal/mol, the negative value meaning that the reduced form is less stable. A higher amount of entropy must be removed from the unfolded state of the reduced form to reach the native structure. By comparing the values of $T\Delta S$(red) = 34.4 kcal/mol and $T\Delta S$(oxi) = 24.1 kcal/mol of both forms, it is possible to note that much less entropy must be removed from the unfolded state of the oxidized form. Also, a lower amount of enthalpy must be lost by the oxidized form to fold, as indicated by comparing ΔH(red) = 37.9 kcal/mol and ΔH(oxi) = 28.9 kcal/mol.

The motivation to discuss these experimental data, which are not part of the present study, is that, from the experimental point of view, it is difficult to decompose the conformational free energy from the solvent free energy. In the previously mentioned reference,[50] the refolding of the disulfide intact form of SOD1 shows a favorable entropy of activation, which is a composition of conformational entropy and enthalpy decrease and solvent effects, resulting in a positive contribution to overcoming the free-energy barrier of folding. It is necessary to highlight this aspect to make comparisons with data from simulations on simplified models, which do not explicitly take into account the solvent effects. Only conformational changes are probed by the present simulations, which makes the character of the barrier be entropic, as both enthalpy and entropy decrease. The enthalpy and entropy decrease monotonically with the reaction coordinate, making it impossible to have a favorable activation entropy or an unfavorable activation enthalpy. This is the case of one of the mentioned experimental studies,[50] where combined conformational and solvent effects lend an enthalpic character to the free-energy barrier and a favorable entropy of activation. Then, more than trying to compare the absolute numbers related to the free-energy profiles from simulations and experimental studies, the data obtained from simulations may serve as complementary data by estimating the free energy due to conformational changes. This is the purpose of the next paragraph: to add data from simulations to experimental data to try to decouple changes in the chain from those occurring in the solvent.

This attempt is not free of errors not only due to experimental errors but also mainly due to the simplified model used, which takes into account only the backbone of the chain. More than to show exact numbers, the analysis must be seen from a qualitative perspective and starts by considering that it is possible to decompose the difference in the experimental values $\Delta\Delta G_{exp} = \Delta G_{red} - \Delta G_{oxi}$ of the free energy of the two forms as a sum of the configurational (conf) and hydration (hyd) contributions $\Delta\Delta G_{exp} = \Delta\Delta G_{conf} + \Delta\Delta G_{hyd}$. From simulations, we have already mentioned that $\Delta\Delta G_{conf}$ = 1.3 kcal/mol, which is not much different than the experimental value around 1.5 kcal/mol; this would imply that $\Delta\Delta G_{hyd}$ is nearly zero. The difference in the stabilities of reduced and oxidized forms is mainly due to conformational aspects. The present simulations point to similar values for the changes in the enthalpy and entropy of folding for both forms (data not shown). Then, the experimentally higher values for these changes for the reduced form could possibly come from the solvent. As for the higher variation of the entropy and enthalpy for the folding process of the reduced form, one possible explanation is that water is more organized in the unfolded state due to the more exposed apolar residues when compared with the oxidized form. The present simulations show a significant reduction of the radius of gyration of the oxidized form relative to the reduced structure in its unfolded states, which can be seen in Figure 3. This could explain the

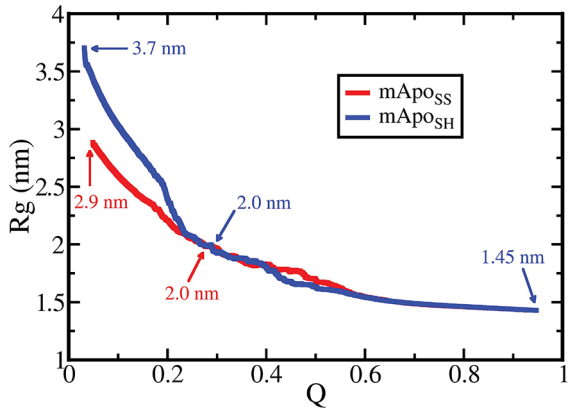


Figure 3. Radius of gyration, R_g, as a function of the native contact fraction, Q, for the mApo$_{SH}$ (in blue) and mApo$_{SS}$ (in red) cases. The arrows indicate the positions and the value of R_g in the unfolded, transition (2.0 nm), and native (1.45 nm) states.

reason for the higher variation of entropy in the folding process of the reduced form, which would be nearly compensated by the also higher variation of enthalpy compared with the oxidized form. Finally, the data from the entropy and enthalpy from the present simulations (data not shown) point to an enthalpic compensation for the entropic reduction of the oxidized form (due to disulfide) in the unfolded state, thus reducing the expected value of 4.6 kcal/mol predicted by the polymer model.

Figure 3 shows the behavior of the radius of gyration, R_g, as a function of the fraction of native contacts, Q, in the case mApo$_{SH}$ (in blue) and mApo$_{SS}$ (in red). It can be noted that the value of R_g in the transition state is ∼2.0 nm for both. The folded state also shows the same behavior for both systems, independent of the oxidation of the disulfide bond, with R_g = 1.45 nm. This latter value is nearly equal to the radius of gyration of the crystallographic structure, which is a consequence of using SBMs. The unfolded states show R_g equal to 3.7 and 2.9 nm for the reduced and oxidized cases, respectively. Using the values of R_g obtained for the reduced case, and applying them in the eq 2, we find the theoretical value of β_T, which gives us a value of 0.71. Experimental data[50,59] show that for the apo and reduced monomers, the m values are $m_{D\text{-}TS}$ = 1.05, whereas $m_{D\text{-}N}$ = 1.53. When we compare these data with eq 3, we find an experimental value of β_T equal to 0.69, which is very close to the theoretical value. For the apo and oxidized cases, using the values of R_g, we can find the value of m_{TS} = 0.65, which is quite reasonable when compared with the value of 0.62 determined by Kayatekin et al.[50] This is a good indication that the ensembles are being reasonable probed, at least for the comparison with this experimental data.

Folding: Difference between Apo and Holo Cases. Studies have shown that in the absence of metals, in addition to losing its function, SOD1 also has its stability reduced, which correlates with data shown in Figure 2. This means that holo SOD1 remains folded at higher temperatures than the apo forms. However, there are still discussions in the literature about the exact moment the metals bind to the enzyme. It has been proposed that in vivo SOD1 needs the assistance of the

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661

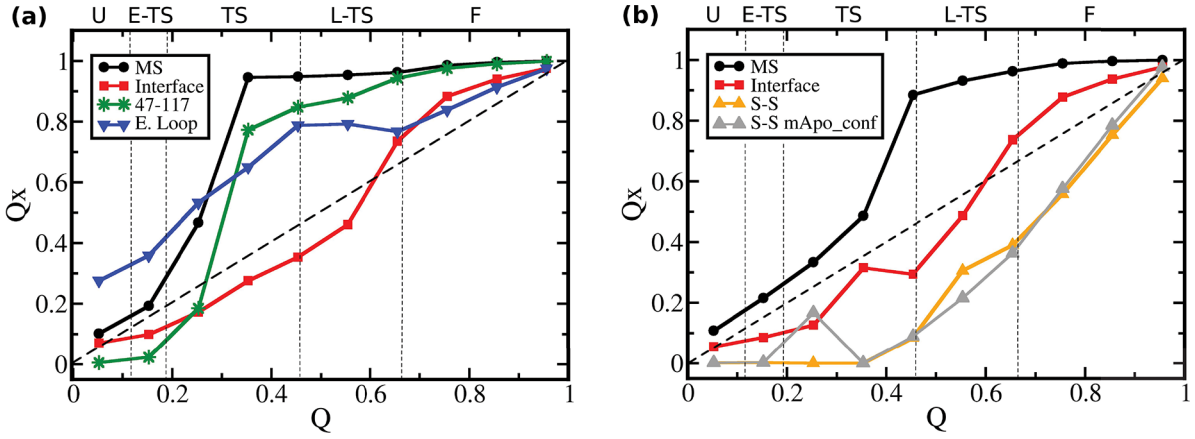


Figure 4. Metal-binding site, CA57−CA146 disulfide bond (S−S), interface, and electrostatic loop formation in monomeric holo SOD1. Q_x is the average native contact fraction in each region: metal binding site (MS, black ●), interface (red ■), CA57−CA146 disulfide bond (S−S, yellow ▲ for $mHolo_{conf}$ and gray ▲ for $mApo_{conf}$ in panel b), CA47−CA117 (47−117, green *), and electrostatic loop contacts (E. Loop, blue ▼). The "interface" is considered the region in the monomer that makes contact with the other monomer when the dimer is formed. Each graph was made at the respective folding temperature of each system: (a) $mHolo_{SS}$ and (b) $mHolo_{conf}$. The gray line and ▲ (SS $mApo_{conf}$ in panel b) indicate the formation of the disulfide bond (57−146) in the Apo case. This curve is just to compare the holo and the apo conformational cases. The abbreviations, U, E-TS, TS, L-TS, and F, refer to unfolded, early transition state, transition state, late transition state, and folded, respectively. All of these regions were defined by dividing the free-energy barrier in three (equals) portions according to is height. The dashed diagonal line indicates the formation of the total monomer.

copper chaperone Ccs1, which would bind an immature form of SOD1 in its apo and disulfide reduced forms, inducing structural changes and increasing the affinity for zinc.[60] Then, later on, close to the native structure, Ccs1 assists the folding completion by binding the copper, oxidizing the intramolecular disulfide bond, and joining the two monomers to form the dimeric and native structure of SOD1.[61−64] The in vitro scenario is different according to several studies. The disulfide bond is probably made early, before the transition state, which causes a modest acceleration of the folding process. Moreover, a higher affinity for zinc demands the formation of the disulfide bond.[55]

Results from the present study show that the metal-binding site formation occurs in the early stages of the SOD1 folding, starting before the transition state, as shown by Figure 4. Q_x is the fraction of native contacts, where x represents different regions of the chain, such as the metal site (MS), disulfide bond (S−S), contact between residues 47 and 117, contacts between residues at the interface region, and contacts between residues in the electrostatic loop (E. loop). Our model is too simple to distinguish between zinc and copper neighborhoods. That is the reason to mention the metal site (Cu + Zn). Residues 47 and 117 have relatively high ϕ values[19,20] and they make contact in the native structure, which is the motivation to monitor its formation. Figure 4 shows these order parameters in two different mHolo scenarios: (a) when the disulfide bond is always made and (b) when it is made along the folding process. In both cases, the formation of the MS region (represented by circles) steeply increases in the transition state, compatible with higher folding rates upon zinc binding described in some studies.[53,54] The contact between residues 47 and 117 (represented by *) also shows a fast increase in its frequency of formation in the transition state ensemble, compatible with a high ϕ values of both residues. As already mentioned, the comparison between the experimental and simulated values of β_T showed good agreement, an indication that the transition-state ensemble is being well probed, as it is clearly shown that the contact between residues 47 and 117 is also compatible with their ϕ values.

The residues of the monomer present in the dimer interface are represented by squares. Their order parameter, in Figure 4a,b, shows that the formation of this region is gradual, following the chain folding. On the contrary, the disulfide bond between C57 and C146, represented by ▲ in Figure 4b, in both conformational cases, $mHolo_{conf}$ (black triangle) and $mApo_{conf}$ (gray ▲), is formed late, having its growth after the L-TS state. Finally, the region surrounding the electrostatic loop (▼ in Figure 4a) was shown to be one of the first regions of the protein to begin to form, always ahead of the "chain folding" (dashed line).

Another way of verifying the effect that the modifications in the contact map can have on the SOD1 folding is by calculating the route measure, $R(Q)$,[42] as seen in Figure 5. The measure of $R(Q)$ shows how specific the path is, followed by the protein in its folding process. For the SOD1 monomer, high $R(Q)$ values around the TS region were observed for the five simulated cases, suggesting a specific TS state, which means it is necessary to make specific contacts to overcome the

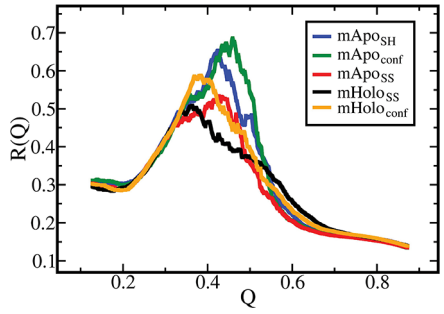


Figure 5. Folding route measure, $R(Q)$, for five monomeric systems: $mHolo_{SS}$ (black curve), $mHolo_{conf}$ (orange curve), $mApo_{SS}$ (red curve), $mApo_{conf}$ (green curve), and $mApo_{SH}$ (blue curve). The five curves were obtained at the respective folding temperature of each system, where Q is the native contact fraction of each one.

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661

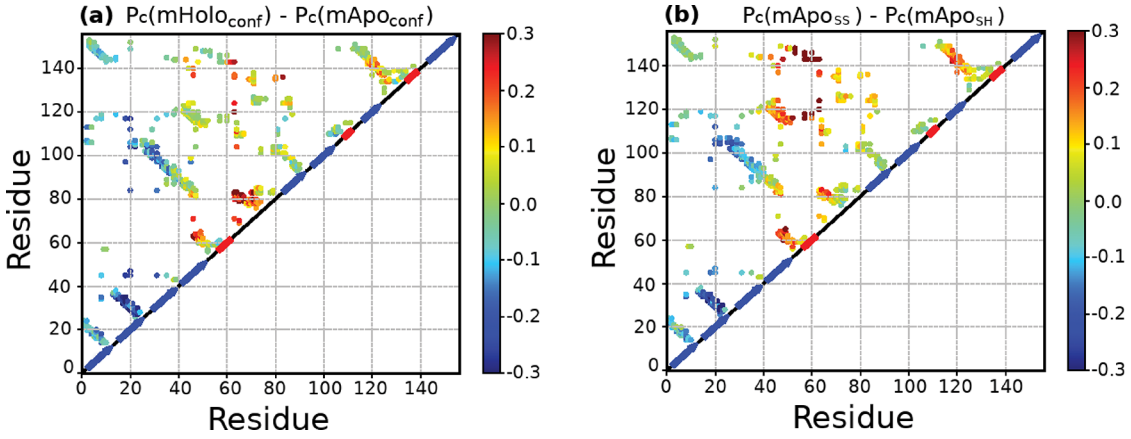


Figure 6. Native contact probability at TS state. (a) Difference in the probabilities of native contact formation (P_c) in the TS region between $mHolo_{conf}$ and $mApo_{conf}$, $P_c(mHolo_{conf}) - P_c(mApo_{conf})$. (b) Difference in the probabilities of native contact formation in the TS region between $mApo_{SS}$ and $mApo_{SH}$, $P_c(mApo_{SS}) - P_c(mApo_{SH})$. The values of the probabilities were obtained at the respective folding temperature of each system. Positive probability values, given by the color bar in the graph, show that there was a higher propensity of the contact to be formed in $mHolo_{conf}$ than in $mApo_{conf}$ in panel a and in $mApo_{SS}$ than in $mApo_{SH}$ in panel b. The blue arrows, red rectangles, and black lines in the diagonal of the figure represent the secondary structures of the native protein: β-sheets, helices, and coils, respectively.

free-energy barrier to reach the native state. When comparing our results with those obtained by Clementi et al.,[42] the $R(Q)$ profile of monomeric SOD1 resembles the profile of SH3, which has a high $R(Q)$ value in the transition state. Similarly, the SH3 protein is also a "β-barrel" protein, which may suggest a typical behavior of proteins with this motif.

The high value of $R(Q)$ for the reduced ($mApo_{SH}$) and conformational ($mApo_{conf}$) apo cases shows that in these simulations, the folding route has made the transition state even more specific. High values of $R(Q)$ also may indicate a relatively rugged free-energy landscape, and the search for this specific path may hinder the folding process. In most cases, high values of $R(Q)$ in the transition state indicate proteins with low folding rates, and this seems to happen in the absence of the disulfide bond for the SOD1 apo case. This possible "delay", associated with the presence of local minima on the energy landscape, can lead the protein to the formation of pathogenic structures, such as amyloid aggregates. By forcing the disulfide bond oxidation from the beginning of the folding for both the $mHolo_{SS}$ and the $mApo_{SS}$ cases, one can see that this causes a decrease in the specificity of contacts. The holo case was the one that showed the lowest value of $R(Q)$, suggesting that the "presence" of the metals can make the transition state less specific, paving the way to overcome the barrier more efficiently, which is compatible with an increase in the folding rate.

The $R(Q)$ calculation quantifies the specificity of the pathways, but it does give information about the contacts that the amino acid residues are forming. Figure 6 shows the difference in the probability of native contact formation in the transition state in panel a between $mHolo_{conf}$ and $mApo_{conf}$ and in panel b between $mApo_{SS}$ and $mApo_{SH}$. The region near the zinc loop (region between residues 60 and 80) is ~30% more likely to form in the holo than in the apo case. Figure 6a also shows that in addition to affecting the formation of the metal site, the apo structure is more likely to form other secondary structures than the holo case. More specifically, the amino acid residues that appear to most likely form contacts in the TS of the $mApo_{conf}$ structure are in β-sheets 2, 3, and 6. In this case, the differences in the behavior of $R(Q)$ may not only be related to increased specificity but also be due to a change in the protein folding pathway.

Figure 6b reflects the perturbation caused by the disulfide bond in the folding of the apo monomer. It shows the difference that the early formation of the disulfide bond between C57 and C146 can make in the folding as compared with the case in which it does not form (the reduced case). In Figure 6b, we see that the result of panel a is maintained in the sense that now the $mApo_{SH}$ forms contacts between β-sheets 2, 3, and 6 more often than the apo oxidized form. However, the oxidation of the disulfide bond caused an increase in the formation of the region between the metal site and the electrostatic loop, suggesting an allosterism between these regions in the transition state. By simulating a force spectroscopy experiment on an apo monomer from SOD and using an all-atom model and explicit solvent, Habibi et al.[65] were able to see that the last sheets to be unfolded are exactly β-sheets 2, 3, and 6. Ivanova et al.[18] studied the propensity to form aggregates of amyloid fibers, which indicated that some of the β-sheets are formed with a higher probability. In our simulations, the apo reduced system is more prone to form aggregates. Thus our results corroborate the idea that the absence of metals, together with the reduction of the disulfide bond, may favor a misfolding process.

The SOD1 folding process has also been characterized by having some features that differentiate it from Ig-like β-sandwich proteins, as shown by Nordlund and Oliveberg by the ϕ-value analysis.[19] According to the experimental ϕ-value analysis, the apoSOD1 folding nucleus may be composed by the structuring of consecutive β-sheets (β1, β2, β3, β4) involving distant β-sheets, as in the case of the interaction between β4 and β7. One way to note regions that are part of the folding nucleus is by calculating ϕ values for the native contact. From simulations, it is possible to monitor the frequency of each native contact and define ϕ values for contact between amino acids i and j according to the contact frequency of formation, f, in each ensemble: unfolded (U), transition (TS), and folded (F), by the expression $\phi_{ij} = (f_{ij}^{TS} - f_{ij}^{U})/(f_{ij}^{F} - f_{ij}^{U})$. This defined ϕ_{ij} seems natural because the main term of the potential is based on the contact map of the native structure. If the frequency of contact between residues i and j

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661

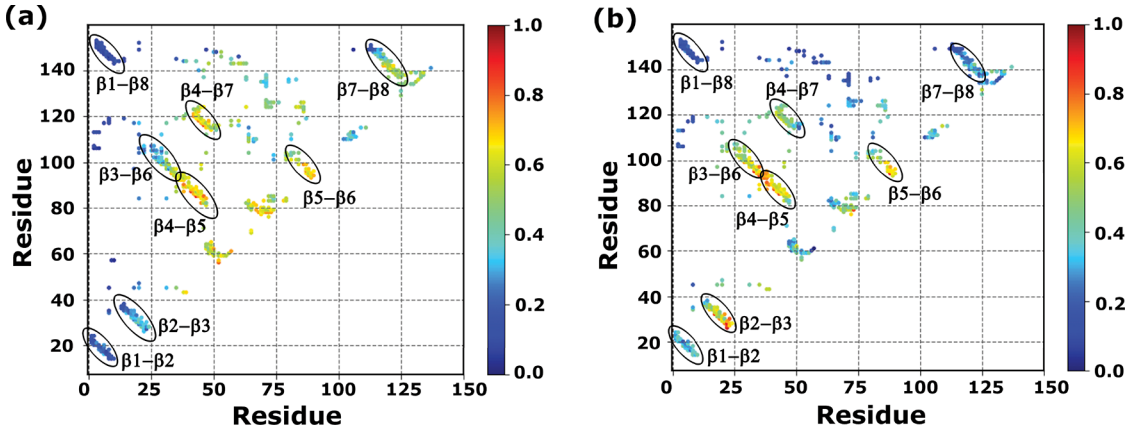


Figure 7. (a) ϕ values for native contacts of the $mApo_{SS}$. (b) ϕ values for native contacts of the $mApo_{SH}$. The ϕ values were obtained at the respective folding temperature of each system and are quantified by the color bar in the graph. Black ellipses are used to indicate contacts between residues that join two β-sheets.

in the transition state, f_{ij}^{TS}, is nearly equal to corresponding frequency in the folded state, f_{ij}^{F}, then the value of ϕ_{ij} is close to 1, and if $f_{ij}^{TS} \approx f_{ij}^{U}$, then ϕ_{ij} will tend to 0. Then the quantity ϕ_{ij} can capture the aspect of the neighborhood of the pair (i, j) in the transition state. The exact value of ϕ_{ij} is not to be compared with the experimental ϕ_i, which measures the ϕ value of residue i, but there must be an expected correlation between the quantities (ϕ_i, ϕ_j) and ϕ_{ij}. For instance, residues 47 and 117 have ϕ values 0.4 and 0.8, according to Nordlund et al.[19] One of the ellipses in Figure 7a for the apo oxidized form represents the contacts between β-sheets β4 and β7, which includes residues 47 and 117. The figure shows that some ϕ_{ij} values are close to 0.8. The contact between residues 47 and 117 is described in detail by the green curve of Figure 4a, which shows that its formation really mainly occurs in the transition state. The corresponding ellipse in Figure 7b, for the reduced form, shows lower ϕ_{ij} values, suggesting a less cooperative contact formation of β-sheets 4 and 7. Our results show that the apoSOD1 folding nucleus is dependent on the disulfide bond 57−146 condition, as seen in Figure 7. The interaction between residues of β4 and β7 were shown to be more relevant in the $mApo_{SS}$ (Figure 7a) than in the $mApo_{SH}$ case (Figure 7b).

According to Figure 7, it is possible to note that for the oxidized form (Figure 7a), β4 and β7 have some contacts with high ϕ values (>0.7), indicating that these are a fundamental part of the folding nucleus, consistent with the results of Nordlund.[19] As for the reduced form, the contacts joining β4 and β7 have intermediate ϕ values (<0.5), as seen in Figure 7b, and the folding nucleus can be drawn mainly by the structuring of consecutive β-sheets: β2, β3, β4, and β5. However, regardless of the disulfide bond condition, β-terminal contacts (β1−β8) have low ϕ values, indicating that these strands are largely unstructured in the transition state and, consequently, are not part of the apoSOD1 monomer folding nucleus.

Kinetics of SOD1 Monomer. Until now, all data from the simulations have referred to thermodynamics. In this section, the kinetics of folding related to the order parameter is described by performing 200 events of folding, starting from uncorrelated unfolded structures. Figure 8 shows the mean unfolded fraction $1 - Q$ as a function of the simulation time for the three apo cases. The chain is considered folded when the unfolded fraction $1 - Q$ reaches 0.17, equivalent to the

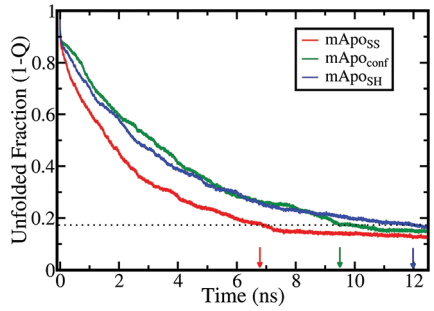


Figure 8. Monomeric apo SOD1 folding kinetics. Formation of the total native contacts in the monomeric apo cases: oxidized (red curve, $mApo_{SS}$), with the disulfide bond made along the folding process (green curve, $mApo_{conf}$), and the reduced (blue curve, $mApo_{SH}$) against time. The dashed line was set to the value of the unfolded fraction that corresponds to the native minimum ($Q = 0.83$) obtained by the free-energy curve (Figure 2). The arrows indicate the time "τ" in which each curve reaches the dashed line: $mApo_{SS}$, $\tau = 6.80$ ns; $mApo_{conf}$, $\tau = 9.5$ ns; and $mApo_{SH}$, $\tau = 12$ ns. All three curves were obtained at the same temperature ($T = 1.211$, in reduced units) using an average of 200 simulations.

folded fraction $Q = 0.83$, which corresponds to the folded minimum in Figure 2.

The simulations are performed at low reduced temperatures to fold the chain in a reasonable amount of computational time. The simulations are performed for three different cases: $mApo_{SS}$ (red), $mApo_{conf}$ (green), and $mApo_{SH}$ (blue). The arrows point to the value 0.17 (unfolded fraction) of the order parameter. The time units, in nanoseconds, do not correspond to real folding times, which are on the orders of seconds for SOD1. This is the reason the comparisons are made just as relative to one another. The unfolded fraction $1 - Q$ reaches 0.17 in 12.0 ns for the reduced case $mApo_{SH}$ and 6.8 ns for the oxidized form $mApo_{SS}$, a factor of 1.8 times faster. This is consistent with the results of Figure 2, which show a slightly lower free-energy barrier of 0.3 kcal/mol for the oxidized form, which, according to the transition-state theory, makes the rate increase by a factor ∼1.7. Moreover, this factor of 1.8 from kinetics is also consistent with the experimental result,[55] which finds that upon oxidizing the disulfide bond, the folding reaction accelerates two-fold.

Correlating Mutations to Temperature Changes in SOD1. This section is devoted to discussing correlations, in

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661

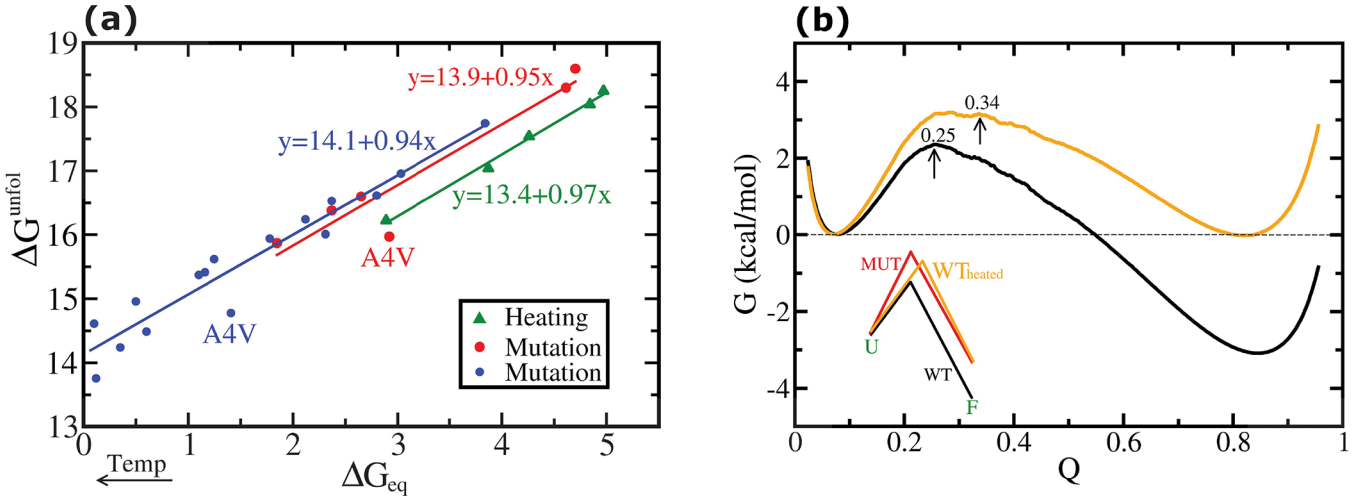


Figure 9. (a) Unfolding barrier (ΔG^{unf}) against stability (ΔG_{eq}). The green triangles are results obtained in the supplementary data of ref 50, which show the rates for folding and unfolding as a function of temperature. The unfolding barrier (in kcal/mol) is calculated by taking a prefactor found in the same reference. The solid red circles represent the variation of the unfolding barrier (in kcal/mol) as a function of mutations using experimental data from ref 55. The blue symbols are mutation data from Lindberg et al.[67] The black arrow indicates the direction of increasing temperature. The equations show the linear fit for each case, labeled by color. (b) Curves represent the free energy as a function of the reaction coordinate from our simulations and show the shift of the transition state as the temperature is increased. These results motivate the diagram shown at the bottom, where, for the same thermodynamic stability, mutations produce a higher change of the unfolding barrier when compared with the change in temperature.

regards to thermodynamics and kinetics, between mutations and temperature changes for the apo SOD1. Figure 9a shows three plots of unfolding barrier against stability. The triangles are results taken from the supplementary data of Kayatekin et al.,[50] which show the rates for folding and unfolding as a function of temperature for the wild-type apo. The unfolding barrier is calculated by taking a prefactor found in the same reference. The solid red and blue circles represent the unfolding barrier for mutants whose thermodynamic and kinetic parameters are found in table 1 of ref 55 and table 1 of ref 67, respectively. Interestingly, the triangles and circles have similar linear regressions, as shown by the slopes and also the y intercept, which measure the unfolding barrier at zero thermodynamic stability. For the green symbols, which represent temperature changes, the unfolding barrier at zero stability is ∼13.4 kcal/mol, a value very close to 13.9 (red symbols) and 14.1 kcal/mol (blue symbols) obtained from mutation experiments. Additionally, the slopes of the three curves are also very similar (0.97 for heating experiments and 0.94 and 0.95 (data in blue and red, respectively) for mutation experiments). The slopes measure the loss in the kinetic stability when the thermodynamic stability decreases. The similarity of these three slopes is significant. A study by Tzul et al.,[66] measured the slopes for eight different proteins (in mutation experiments), finding values between 0.75 and 1.33, which is a rather broad range. The fact that the slopes from heating experiments are close to the corresponding quantities from mutation experiments is not likely to be a coincidence due to the mentioned broad range of values for the slopes.

The mutation A4V stays significantly away from the line. To be placed in the straight line, A4V should have ΔG_{unfol} equal to 16.8 kcal/mol, but the actual value is only 15.8 kcal/mol. The unfolding rate is increased by a factor of 6 ($\exp(16.8 - 15.8)/RT$) if compared with the trend of the other mutations. Data from Kayatekin et al.[50] show that A4V has the highest unfolding rate ($k_u = 1 \times 10^{-3}$ s^{-1}) among all five mutants. In the folding/unfolding experiments conducted in Kayatekin et al.,[55] the authors mention that at 37 °C, A4V is too destabilized to allow an accurate fitting of the data. The A4V mutation is also known to be the most lethal mutant of SOD1, and it is often associated with aggregation.

The mutation data represented in blue (in Figure 9a), taken from Lindberg et al.,[67] are almost a continuation of the red line, which is a regression of the red symbols representing mutation data from Kayatekin et al.[50] The differences are due to the thermodynamic stabilities measured by the two groups. For instance, the first research team obtains 3 kcal/mol for the native apo monomer, whereas the same quantity is ∼4.6 kcal/mol, as measured by the second team. This difference slightly shifts the blue curve to the left. Otherwise, the two lines would coincide, as the slopes are nearly the same. Even so, the red and blue lines are nearly segments of the same straight line, as shown by Figure 9a. Again, the data in blue show that mutation A4V is the farthest point to the straight line, as was already shown in the data from Kayatekin et al.[50] (in red). Mutation A4V, in both data sets (blue and red), should have the unfolding barrier 0.8 to 1 kcal/mol ($>kT$ = 0.6 kcal/mol) higher to approximate the straight lines. This mutation seriously affects the kinetic stability.

In regards to the three regressions shown in Figure 9a, the lines representing mutations and the temperature increase are not only parallel but also very close to each other. Despite the fact that all of the data in this section were taken from other studies, the comparison we have made (blue, green, and red curves) led us to convey a new proposition. Because the green line (from heating experiments) is very close to the red and blue lines (mutations), also with similar slopes, it may be possible to obtain information on one data set (heating) by using another data set (mutation). This may be a particular feature of SOD1 or may occur for other proteins as well. Additionally, the proximity of the lines representing mutations and heating experiments may indicate that mutations, along with evolution, would occur as a search for thermodynamic and kinetic stabilities.

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661

It is opportune to discuss a possible reason that the green curve is slightly below the red line. Our understanding is that it is related to the shift of the transition state to the right when the temperature is increased, as shown by our data, represented by Figure 9b. At low temperature, the transition state is at the value 0.25 of the reaction coordinate. At a higher temperature, the position changes to 0.34, as shown. As for mutations, those shown in Kayatekin et al.,[50] do not significantly change the position of the transition state. These data show the m values of the wild-type apo monomer for different temperatures and also for the four mutations. From those values, it is possible to calculate the parameter $\alpha = m^{\ddagger}_{U\to N}/(m^{\ddagger}_{U\to N} - m^{\ddagger}_{N\to U})$, which measures the position of the transition state. For the four mutations, these values are 0.64 (L38V), 0.65 (G93A), 0.63 (L106A), and 0.64 (S134). By comparing with the α of the wild type (at T = 298 K), which is 0.65,[50] it is possible to notice that the position of the transition state is nearly the same. However, when looking at the α of the wild type at T = 313 K, the value increases to 0.69. Our data also show this shift to the right, as seen in Figure 9b. Studies show the shift of the transition state to the native region as temperature is increased (Hammond's postulate), but what about the vertical movement (the barrier height)? By looking at the measurements by Kayatekin et al. related to the two data sets (mutations and heating, red and green lines of Figure 9a, respectively) and choosing the same thermodynamic stability in two data sets, it is possible to see that the unfolding barrier is lower for the heating process when compared with mutations. For instance, the data show that for mutation L106V, the thermodynamic stability is 2.6 kcal/mol, and the corresponding unfolding rate is $k_u = 5 \times 10^{-4}$ s^{-1}, whereas the same quantities for the wild type at T = 313 K are 2.8 kcal/mol and $k_u = 3 \times 10^{-3}$ s^{-1}. Then, for similar values of the stability, the heating process shows an unfolding rate one order of magnitude higher, corresponding to a lower unfolding barrier. This is the reason the green curve is below the red curve in Figure 9a. There are combined effects in temperature experiments, that is, a shift of the transition state to the right and a smaller increase in the barrier (when compared with the mutation in the same thermodynamic stability). They suggest the diagram shown at the bottom of the Figure 9b, where the black curve represents the wild type, the red curve represents a mutant, and the orange curve represents the wild type at high temperature, which shows a shift of the transition state to the right and places the barrier below the position of the mutation curve. This diagram may justify the aspect of Figure 9a, where the green line (heating) is placed below the red and blue lines (mutations). As for mutations that similarly affect the native and transition states (high ϕ values), this behavior may change, as in the case of mutation A4V, highlighted in Figure 9a. This mutation has a relatively low unfolding barrier, which places it well below the red line. From a kinetic and thermodynamic point of view, A4V behaves as the wild-type molecule at a temperature 313 K, which is represented by the last triangle at the bottom of the green line, close to the red circle representing A4V, or at an even higher temperature, according to data in blue.

CONCLUSIONS

There is good agreement between the values of the reaction coordinate β_T related to the transition state of the SOD1 monomer, studied by experimental techniques and simulations. It suggests that simulations may have probed the folding transition reasonably well, thus producing acceptable values of changes of entropy and enthalpy related to conformational changes. Thus the simulations helped us to distinguish between a combined effect of solvent and conformational effects from pure conformational changes. The proposed way to model the monomer forms, reduced, oxidized, apo, and holo, showed results that correlate with experimental data from a thermodynamic and also a kinetic point of view. In particular, the kinetics of folding of the simplified model showed a relation between the relaxation times of the different forms, which agrees with the same relation obtained from thermodynamics. Additionally, important contacts necessary to overcome the free-energy barrier were detected by simulations, which showed a fast increase in their frequency in the transition state. Experimentally these contacts are made between residues with relatively high ϕ values. This is another factor that favors the reliability of the conformational changes produced by simulations. Also, the electrostatic loop and zinc site, experimentally suggested as preorganized, showed a relatively organized pattern before the transition state, as revealed by the fraction of formation as a function of the reaction coordinate that describes the folding transition. Regarding the disulfide bond, the folding of this simplified model suggests that it is formed late, after the transition state, possibly qualitatively agreeing with the in vivo folding process. In this regard, a ϕ-value analysis[59] also suggests that the disulfide bond may not be formed at the transition state, but for the oxidized form, the disulfide bond between cysteines 57 and 146 is able to decrease the specificity of the folding pathway and marginally affect the folding nucleus, facilitating the formation of β-sheets 4 and 7. As for the reduced form, where the cysteines come close to each other late in the folding process, the folding nucleus concentrates on β-sheets 2, 3, and 4. The simulations also associated a more specific route with a higher free-energy barrier, suggesting one possible factor for the longer conformational search near the transition state to overcome the barrier. Finally, in the last section, we observed the remarkably similar behavior in the linear regressions of the unfolding barrier, as function stability, related to mutations and heating, which can be seen as a quantitative way to correlate mutations to the search for thermal stability.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.9b09640.

Native contact map and information about contacts that were removed to model and simulate each SOD1 monomer system presented in this work, information about the computational details of the folding kinetic simulations, and the replica-exchange sampling in the folding simulations for all five systems studied using REMD (PDF)

AUTHOR INFORMATION

Corresponding Authors

Vitor B. P. Leite − *São Paulo State University (UNESP), IBILCE, São José do Rio Preto, Brazil, and Rice University, Houston, Texas;* orcid.org/0000-0003-0008-9079; Email: vitor.leite@unesp.br

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661

Jorge Chahine − *São Paulo State University (UNESP), IBILCE, São José do Rio Preto, Brazil*;
Email: jorge.chahine@unesp.br

Other Authors

Paulo R. Mouro − *São Paulo State University (UNESP), IBILCE, São José do Rio Preto, Brazil*

Ana P. R. Povinelli − *São Paulo State University (UNESP), IBILCE, São José do Rio Preto, Brazil*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.9b09640

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Ronaldo Junio de Oliveira and Fernando Bruno da Silva for the help with ϕ-value scripts. P.R.M. and A.P.R.P. were supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES - Brazil. V.B.P.L. was supported by FAPESP grants 2018/18668-1 and 2016/19766-1, and the Center for Theoretical Biological Physics was sponsored by the National Science Foundation NSF Grant PHY-1427654.

■ REFERENCES

(1) Das, A.; Plotkin, S. S. SOD1 exhibits allosteric frustration to facilitate metal binding affinity. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 3871−3876.
(2) Tainer, J. A.; Getzoff, E. D.; Beem, K. M.; Richardson, J. S.; Richardson, D. C. Determination and analysis of the 2 structure of copper, zinc superoxide dismutase. *J. Mol. Biol.* **1982**, *160*, 181−217.
(3) Valentine, J. S.; Doucette, P. A.; Zittin Potter, S. Z. Copper-zinc superoxide dismutase and amyotrophic lateral sclerosis. *Annu. Rev. Biochem.* **2005**, *74*, 563−593.
(4) Battistoni, A.; Folcarelli, S.; Gabbianelli, R.; Capo, C.; Rotilio, G. The Cu,Zn superoxide dismutase from Escherichia coli retains monomeric structure at high protein concentration. Evidence for altered subunit interaction in all the bacteriocupreins. *Biochem. J.* **1996**, *320* (Pt 3), 713−716.
(5) Sturtz, L. A.; Diekert, K.; Jensen, L. T.; Lill, R.; Culotta, V. C. A fraction of yeast Cu/Zn superoxide dismutase and its metallochaperone, CCS, localize to the intermembrane space of mitochondria: A physiological role for SOD1 in guarding against mitochondrial oxidative damage. *J. Biol. Chem.* **2001**, *276*, 38084−38089.
(6) Hassan, H. M.; Fridovich, I. Paraquat and Escherichia coli. Mechanism of production of extracellular superoxide radical. *J. Biol. Chem.* **1979**, *254*, 10846−10852.
(7) Biliaderis, C. G.; Weselake, R. J.; Petkau, A.; Friesen, A. D. A calorimetric study of human CuZn superoxide dismutase. *Biochem. J.* **1987**, *248*, 981−984.
(8) Parge, H. E.; Getzoff, E. D.; Scandella, C. S.; Hallewell, R. A.; Tainer, J. A. Crystallographic characterization of recombinant human CuZn superoxide dismutase. *J. Biol. Chem.* **1986**, *261*, 16215−16218.
(9) Mei, G.; Rosato, N.; Silva, N.; Rusch, R.; Gratton, E.; Savini, I.; Finazzi-Agro, A. Denaturation of human copper-zinc superoxide dismutase by guanidine hydrochloride: a dynamic fluorescence study. *Biochemistry* **1992**, *31*, 7224−7230.
(10) Ding, F.; Dokholyan, N. V. Dynamical roles of metal ions and the disulfide bond in Cu, Zn superoxide dismutase folding and aggregation. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 19696−19701.
(11) Arnesano, F.; Banci, L.; Bertini, I.; Martinelli, M.; Furukawa, Y.; O'Halloran, T. V. The unusually stable quaternary structure of human Cu,Zn-Superoxide Dismutase 1 is controlled by both metal occupancy and disulfide status. *J. Biol. Chem.* **2004**, *279*, 47998−48003.
(12) Strange, R. W.; Yong, C. W.; Smith, W.; Hasnain, S. S. Molecular dynamics using atomic-resolution structure reveal structural fluctuations that may lead to polymerization of human CuZn superoxide dismutase. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 10040−10044.
(13) Getzoff, E.; Cabelli, D.; Fisher, C.; Parge, H.; Viezzoli, M.; Banci, L.; Hallewell, R. Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. *Nature* **1992**, *358*, 347−351.
(14) Tainer, J.; Getzoff, E.; Richardson, J.; Richardson, D. Structure and mechanism of copper, zinc superoxide dismutase. *Nature* **1983**, *306*, 284−287.
(15) Ciriolo, M. R.; Battistoni, A.; Falconi, M.; Filomeni, G.; Rotilio, G. Role of the electrostatic loop of Cu,Zn superoxide dismutase in the copper uptake process. *Eur. J. Biochem.* **2001**, *268*, 737−742.
(16) Fisher, C. L.; Cabelli, D. E.; Tainer, J. A.; Hallewell, R. A.; Getzoff, E. D. The role of arginine 143 in the electrostatics and mechanism of Cu, Zn superoxide dismutase: Computational and experimental evaluation by mutational analysis. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 24−34.
(17) Murzin, A. G.; Brenner, S. E.; Hubbard, T.; Chothia, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **1995**, *247*, 536−540.
(18) Ivanova, M. I.; Sievers, S. A.; Guenther, E. L.; Johnson, L. M.; Winkler, D. D.; Galaleldeen, A.; Sawaya, M. R.; Hart, P. J.; Eisenberg, D. S. Aggregation-triggering segments of SOD1 fibril formation support a common pathway for familial and sporadic ALS. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 197−201.
(19) 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 disease. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10218−10223.
(20) Schmidlin, T.; Ploeger, K.; Jonsson, A. L.; Daggett, V. Early steps in thermal unfolding of superoxide dismutase 1 are similar to the conformational changes associated with the ALS-associated A4V mutation. *Protein Eng., Des. Sel.* **2013**, *26*, 503−513.
(21) Rakhit, R.; Chakrabartty, A. Structure, folding, and misfolding of Cu,Zn superoxide dismutase in amyotrophic lateral sclerosis. *Biochim. Biophys. Acta, Mol. Basis Dis.* **2006**, *1762*, 1025−1037.
(22) Roe, J. A.; Butler, A.; Scholler, D. M.; Valentine, J. S.; Marky, L.; Breslauer, K. J. Differential scanning calorimetry of copper-zinc-superoxide dismutase, the apoprotein, and its zinc-substituted derivatives. *Biochemistry* **1988**, *27*, 950−958.
(23) Rodriguez, J. A.; Shaw, B. F.; Durazo, A.; Sohn, S. H.; Doucette, P. A.; Nersissian, A. M.; Faull, K. F.; Eggers, D. K.; Tiwari, A.; Hayward, L. J.; et al. Destabilization of apoprotein is insufficient to explain Cu,Zn-superoxide dismutase-linked ALS pathogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10516−10521.
(24) Bruijn, L. I.; Miller, T. M.; Cleveland, D. W. Unraveling the mechanisms involved in motor neuron degeneration in ALS. *Annu. Rev. Neurosci.* **2004**, *27*, 723−749.
(25) Andersen, P. M. Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. *Curr. Neurol. Neurosci. Rep.* **2006**, *6*, 37−46.
(26) Banci, L.; Bertini, I.; Boca, M.; Girotto, S.; Martinelli, M.; Valentine, J. S.; Vieru, M. SOD1 and amyotrophic lateral sclerosis: mutations and oligomerization. *PLoS One* **2008**, *3*, e1677.
(27) Dobson, C. M. Protein folding and misfolding. *Nature* **2003**, *426*, 884−890.
(28) Baldwin, R. The nature of protein folding pathways: The classical versus the new view. *J. Biomol. NMR* **1995**, *5*, 103−109.
(29) Dill, K. A.; Chan, H. S. From Levinthal to pathways to funnels. *Nat. Struct. Mol. Biol.* **1997**, *4*, 10−19.
(30) Pande, V. S.; Grosberg, A. Y.; Tanaka, T. On the theory of folding kinetics for short proteins. *Folding Des.* **1997**, *2*, 109−114.
(31) Nelson Onuchi, J.; Nymeyer, H.; Garcia, A.; Chahine, J.; Socci, N. D. The energy landscape theory of protein folding: Insights into folding mechanisms and scenarios. *Adv. Protein Chem.* **2000**, *53*, 87−152.

(32) 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.
(33) Whitford, P. C.; Noel, J. K.; Gosavi, S.; Schug, A.; Sanbonmatsu, K.; Onuchic, J. N. An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields. *Proteins: Struct., Funct., Genet.* **2009**, *75*, 430−441.
(34) Clementi, C.; Garcia, A. E.; Onuchic, J. N. Interplay among tertiary contacts, secondary structure formation and side-chain packing in the protein folding mechanism: All-atom representation study of protein L. *J. Mol. Biol.* **2003**, *326*, 933−954.
(35) Wang, J.; Oliveira, R. J.; Chu, X.; Whitford, P. C.; Chahine, J.; Han, W.; Wang, E.; Onuchic, J. N.; Leite, V. B. P. The topography of funneled landscapes determines the thermodynamics and kinetics of protein folding. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 15763−15768.
(36) Mouro, P. R.; de Godoi Contessoto, V.; Chahine, J.; Junio de Oliveira, R.; Pereira Leite, V. B. Quantifying nonnative interactions in the protein-folding free-energy landscape. *Biophys. J.* **2016**, *111*, 287−293.
(37) Contessoto, V. G.; Lima, D. T.; Oliveira, R. J.; Bruni, A. T.; Chahine, J.; Leite, V. B. P. Analyzing the effect of homogeneous frustration in protein folding. *Proteins: Struct., Funct., Genet.* **2013**, *81*, 1727−1737.
(38) Nymeyer, H.; Garcia, A. E.; Onuchic, J. N. Folding funnels and frustration in off-lattice minimalist protein landscapes. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 5921−5928.
(39) 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.
(40) Oliveira, R. J.; Whitford, P. C.; Chahine, J.; Leite, V.; Wang, J. Coordinate and timedependent diffusion dynamics in protein folding. *Methods* **2010**, *52*, 91−98.
(41) Koga, N.; Takada, S. Roles of native topology and chain-length scaling in protein folding: A simulation study with a Gō-like model. *J. Mol. Biol.* **2001**, *313*, 171−180.
(42) Chavez, L. L.; Onuchic, J. N.; Clementi, C. Quantifying the roughness on the free energy landscape: Entropic bottlenecks and protein folding rates. *J. Am. Chem. Soc.* **2004**, *126*, 8426−8432.
(43) Snow, C. D.; Sorin, E. J.; Rhee, Y. M.; Pande, V. S. How well can simulation predict protein folding kinetics and thermodynamics? *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 43−69.
(44) Gosavi, S.; Chavez, L. L.; Jennings, P. A.; Onuchic, J. N. Topological frustration and the folding of interleukin-1 beta. *J. Mol. Biol.* **2006**, *357*, 986−996.
(45) Chu, X.; Gan, L.; Wang, E.; Wang, J. Quantifying the topography of the intrinsic energy landscape of flexible biomolecular recognition. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E2342−E2351.
(46) Noel, J. K.; Whitford, P. C.; Onuchic, J. N. The shadow map: A general contact definition for capturing the dynamics of biomolecular folding and function. *J. Phys. Chem. B* **2012**, *116*, 8692−8702.
(47) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(48) Noel, J. K.; Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. SMOG@ctbp: simplified deployment of structure-based models in GROMACS. *Nucleic Acids Res.* **2010**, *38*, W657−W661.
(49) Geierhaas, C. D.; Nickson, A. A.; Lindorff-Larsen, K.; Clarke, J.; Vendruscolo, M. BPPred: A Web-based computational tool for predicting biophysical parameters of proteins. *Protein Sci.* **2006**, *16*, 125−134.
(50) Kayatekin, C.; Cohen, N. R.; Matthews, C. R. Enthalpic barriers dominate the folding and unfolding of the human Cu, Zn superoxide dismutase monomer. *J. Mol. Biol.* **2012**, *424*, 192−202.
(51) Ferrenberg, A. M.; Swendsen, R. H. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.* **1988**, *61*, 2635−2638.
(52) Ferrenberg, A. M.; Swendsen, R. H. Optimized Monte-Carlo data analysis. *Phys. Rev. Lett.* **1989**, *63*, 1195−1198.
(53) Kayatekin, C.; Zitzewitz, J. A.; Matthews, C. R. Zinc binding modulates the entire folding free energy surface of human Cu,Zn superoxide dismutase. *J. Mol. Biol.* **2008**, *384*, 540−555.
(54) Li, H.-T.; Jiao, M.; Chen, J.; Liang, Y. Roles of zinc and copper in modulating the oxidative refolding of bovine copper, zinc superoxide dismutase. *Acta Biochim. Biophys. Sin.* **2010**, *42*, 183−194.
(55) Kayatekin, C.; Zitzewitz, J. A.; Matthews, C. R. Disulfide-reduced ALS variants of Cu, Zn superoxide dismutase exhibit increased populations of unfolded species. *J. Mol. Biol.* **2010**, *398*, 320−331.
(56) Pace, C. N.; Grimsley, G. R.; Thomson, J. A.; Barnett, B. J. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem.* **1988**, *263*, 11820−11825.
(57) Vassall, K. A.; Stubbs, H. R.; Primmer, H. A.; Tong, M. S.; Sullivan, S. M.; Sobering, R.; Srinivasan, S.; Briere, L.-A. K.; Dunn, S. D.; Colón, W.; et al. Decreased stability and increased formation of soluble aggregates by immature superoxide dismutase do not account for disease severity in ALS. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 2210−2215.
(58) Privalov, P. In *Stability of Proteins Small Globular Proteins*; Anfinsen, C., Edsall, J. T., Richards, F. M., Eds.; Advances in Protein Chemistry 33; Academic Press, 1979; pp 167−241.
(59) Lindberg, M. J.; Normark, J.; Holmgren, A.; Oliveberg, M. Folding of human superoxide dismutase: Disulfide reduction prevents dimerization and produces marginally stable monomers. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 15893−15898.
(60) Boyd, S. D; Liu, L.; Bulla, L.; Winkler, D. D Quantifying the interaction between copper-zinc superoxide dismutase (Sod1) and its copper chaperone (Ccs1). *J. Proteomics Bioinf.* **2018**, *11*, 1−5.
(61) Culotta, V. C.; Klomp, L. W. J.; Strain, J.; Casareno, R. L. B.; Krems, B.; Gitlin, J. D. The copper chaperone for superoxide dismutase. *J. Biol. Chem.* **1997**, *272*, 23469−23472.
(62) Brown, N. M.; Torres, A. S.; Doan, P. E.; O'Halloran, T. V. Oxygen and the copper chaperone CCS regulate posttranslational activation of Cu,Zn superoxide dismutase. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 5518−5523.
(63) Carroll, M. C.; Girouard, J. B.; Ulloa, J. L.; Subramaniam, J. R.; Wong, P. C.; Valentine, J. S.; Culotta, V. C. Mechanisms for activating Cu- and Zn-containing superoxide dismutase in the absence of the CCS Cu chaperone. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 5964−5969.
(64) Furukawa, Y.; O'halloran, T. V. Posttranslational modifications in Cu,Zn-superoxide dismutase and mutations associated with amyotrophic lateral sclerosis. *Antioxid. Redox Signaling* **2006**, *8*, 847−867.
(65) Habibi, M.; Rottler, J.; Plotkin, S. S. The unfolding mechanism of monomeric mutant SOD1 by simulated force spectroscopy. *Biochim. Biophys. Acta, Proteins Proteomics* **2017**, *1865*, 1631−1642.
(66) Tzul, F. O.; Vasilchuk, D.; Makhatadze, G. I. Evidence for the principle of minimal frustration in the evolution of protein folding landscapes. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E1627−E1632.
(67) Lindberg, M. J.; Byström, R.; Boknäs, N.; Andersen, P. M.; Oliveberg, M. Systematically perturbed folding patterns of amyotrophic lateral sclerosis (ALS)-associated SOD1 mutants. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9754−9759.

https://dx.doi.org/10.1021/acs.jpcb.9b09640
J. Phys. Chem. B 2020, 124, 650−661