# Protein folding from heterogeneous unfolded state revealed by time-resolved X-ray solution scattering

Tae Wu Kim[a,b,c], Sang Jin Lee[a,b,c], Junbeom Jo[a,b,c], Jong Goo Kim[a,b,c], Hosung Ki[a,b,c], Chang Woo Kim[a], Kwang Hyun Cho[a], Jungkweon Choi[c], Jae Hyuk Lee[d], Michael Wulff[e], Young Min Rhee[a,1], and Hyotcherl Ihee[a,b,c,1]

[a]Department of Chemistry, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea; [b]KI for the BioCentury, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea; [c]Center for Nanomaterials and Chemical Reactions, Institute for Basic Science, Daejeon 34141, Republic of Korea; [d]XFEL Beamline Department, Pohang Accelerator Laboratory, Pohang 37673, Republic of Korea; and [e]Experiments Division, European Synchrotron Radiation Facility, F-38043 Grenoble Cedex, France

Edited by D. Thirumalai, University of Texas at Austin, Austin, TX, and accepted by Editorial Board Member Michael L. Klein May 5, 2020 (received for review August 3, 2019)

**One of the most challenging tasks in biological science is to understand how a protein folds. In theoretical studies, the hypothesis adopting a funnel-like free-energy landscape has been recognized as a prominent scheme for explaining protein folding in views of both internal energy and conformational heterogeneity of a protein. Despite numerous experimental efforts, however, comprehensively studying protein folding with respect to its global conformational changes in conjunction with the heterogeneity has been elusive. Here we investigate the redox-coupled folding dynamics of equine heart cytochrome *c* (cyt-c) induced by external electron injection by using time-resolved X-ray solution scattering. A systematic kinetic analysis unveils a kinetic model for its folding with a stretched exponential behavior during the transition toward the folded state. With the aid of the ensemble optimization method combined with molecular dynamics simulations, we found that during the folding the heterogeneously populated ensemble of the unfolded state is converted to a narrowly populated ensemble of folded conformations. These observations obtained from the kinetic and the structural analyses of X-ray scattering data reveal that the folding dynamics of cyt-c accompanies many parallel pathways associated with the heterogeneously populated ensemble of unfolded conformations, resulting in the stretched exponential kinetics at room temperature. This finding provides direct evidence with a view to microscopic protein conformations that the cyt-c folding initiates from a highly heterogeneous unfolded state, passes through still diverse intermediate structures, and reaches structural homogeneity by arriving at the folded state.**

protein folding | cytochrome c | time-resolved X-ray scattering | ensemble | molecular dynamics simulation

Proteins should retain unique three-dimensional (3D) structures to properly express its biological functions in living cells. The linear sequence of amino acids in a protein released from the ribosome encodes not only its native structure but also the folding mechanism by which this structure is attained (1). According to Levinthal's argument (2), an ensemble of unfolded proteins heterogeneously populated with various conformations will not likely find its native state through a random search in the vast conformational space but rather recovers its native form through well-defined folding pathways.

Numerous theoretical studies have been conducted to describe the complex folding dynamics in view of free-energy landscapes (3–10). Among various conceptual frameworks, it has been suggested that the funnel-like free-energy landscape with many microstates is a scheme pertinent to describing the protein folding (3, 4, 7, 10–12). In this picture, a key factor determining dynamics of protein folding is the native bias. This refers to the dominance of interactions stabilizing the specific native conformation over nonnative interactions that may guide to topologically distinct traps or frustrated conformations, so-called misfolded states (3–5, 13). Another factor is a conformational heterogeneity in the unfolded state and on-route intermediates, which can potentially alter the kinetic behavior in the protein folding due to the possibility of broadly distributed pathways. Thus, this folding model has an advantage for the explanation of conformational heterogeneity and its energetics along the reaction pathway of protein folding. To verify the scenario of the funnel-like free-energy landscape in protein folding, numerous experiments with diverse techniques such as laser-induced temperature jump, rapid mixing, and single-molecule spectroscopy have been conducted (14–24). Nevertheless, comprehensively understanding protein folding, especially in regard to its global conformational changes paired with the structural heterogeneity, has been elusive. This is simply due to the complex nature of any protein. Observing temporal evolutions of protein molecules with such intrinsic diversity can be ideally achieved with computational means such as molecular dynamics (MD) simulations. At present, however, reliably accomplishing the goal is not a trivial task at all. This is partly because simulating a long enough time scale is still a large burden. Perhaps more importantly, commonly available force field parameters are usually optimized for folded structures and may not equally behave well during the entire course of folding events from diverse unfolded configurations.

## Significance

**Protein undergoes a folding pathway to find a three-dimensional structure that expresses its biological function in a living cell. From vast theoretical studies, the funnel-like free-energy landscape has been recognized as a prominent scheme to describe protein folding, but it has been unclear how the dynamic behavior of protein folding is linked to conformational heterogeneity. Here we employ the time-resolved X-ray scattering method combined with systematic structural analysis to track the folding of cytochrome c. The kinetics and structural changes obtained from the data reveal that various protein conformations in the unfolded state take the complex folding pathways, resulting in the stretched exponential kinetics. This study may provide structural insights into fundamental principles governing protein folding.**

Author contributions: T.W.K., Y.M.R., and H.I. designed research; T.W.K., S.J.L., J.J., J.G.K., H.K., J.C., J.H.L., M.W., and H.I. performed research; T.W.K., S.J.L., J.G.K., C.W.K., K.H.C., and Y.M.R. analyzed data; T.W.K., Y.M.R., and H.I. wrote the paper; and T.W.K., S.J.L., J.J., J.G.K., H.K., C.W.K., K.H.C., Y.M.R., and H.I. contributed interpretation of results.

The authors declare no competing interest.

This article is a PNAS Direct Submission. D.T. is a guest editor invited by the Editorial Board.



See online for related content such as Commentaries.

[1]To whom correspondence may be addressed. Email: ymrhee@kaist.ac.kr or hyotcherl.ihee@kaist.ac.kr.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1913442117/-/DCSupplemental.

First published June 15, 2020.

The elusivity can be well exemplified with cytochrome *c* (cyt-c). It is a small metalloprotein with 104 amino acids and a heme group that is covalently bound to the protein backbone. Since the pioneering study of cyt-c folding initiated by the dissociation of carbon monoxide from the heme group (25), cyt-c has been recognized as an appropriate model system to probe the folding of single-domain protein in a time-resolved manner. There have been many experimental studies that tracked the cyt-c folding in the wide time region from submicroseconds to late milliseconds (18, 26–36). These experiments revealed that the self-assembly of protein backbones concomitant with the collapse of extended conformation begins on submillisecond regime. Further, a maximum rate of protein folding was carefully scrutinized by using the optical triggering to refold cyt-c, and the experiment showed the diffusion rate of $(\sim 35\ \mu s)^{-1}$ for the local contact of two sites separated by ~50 residues (37). If we assume that the local contacts in the early process of protein folding correctly reflect the native interactions, we can expect a rapid transition to the native state through the nucleation-collapse mechanism (38). However, the time-resolved experiments showed that the actual folding time in cyt-c takes several tens to hundreds of milliseconds (18, 26–34), which are much slower than the diffusion time of protein backbone reported from an earlier study (37). To explain the kinetic difference between the fast intrachain motion of the protein backbone and the relatively slow folding, the slow phase event on the millisecond time region has been obscurely presumed as either the sequential transitions through successive intermediates (26–29, 31–33) or the detours via frustrated conformations on the route to the native conformation (18, 38). This kinetic complexity during cyt-c folding has been presumed to originate from the complication of folding pathways associated with the ligations by natively or nonnatively contacting residues to the heme group. By employing imidazole binding to the heme group or by protonating misligating histidine residues at low pH toward suppressing the nonnative contact formation, some studies showed that the two-state folding mechanism was operational (39, 40), whereas the absence of imidazole accompanied the additional ligand exchange by methionine (26, 28, 29, 33). Meanwhile, other studies showed deviation from the two-state kinetics irrespective of the suppression of nonnative coordination (41, 42). In the end, even for this rather simple cyt-c with a single domain, there is no consensus on the interpretation of the folding dynamics in view of its conformational heterogeneity.

Time-resolved X-ray solution scattering (TRXSS) may serve as an effective way of obtaining a comprehensive view, mainly owing to the superb structural sensitivity from X-ray scattering. Because the redox state of heme acting as a cofactor is coupled with the folding free energy, it is possible to find a denaturing condition where the protein is fully unfolded in one oxidation state while it is fully folded in the other state. In a certain denaturing condition, the change in the redox state of the cofactor resculpts the free-energy landscape, and this brings about a strong native bias toward the folded state as shown in Fig. 1. In this work, we combined TRXSS, also known as time-resolved X-ray liquidography (43–51), with the external electron injection method in order to investigate the redox-coupled folding dynamics of cyt-c. We focus on examining 1) how the structurally heterogeneous nature of the unfolded state changes during the course of folding and 2) what type of kinetic behaviors are accompanied and linked with the inherent heterogeneous nature. The ensemble optimization analysis for the X-ray scattering data based on MD simulations shows that the initial oxidized state which contributes to the folding dynamics has a heterogeneously populated ensemble with diverse unfolded conformations. During folding, this converts to a narrowly populated ensemble around a well-organized folded conformation through an intermediate state, which still contains conformationally heterogeneous characteristics. Surprisingly, the slow folding phase of cyt-c accompanying the overall structural change manifests the stretched-exponential behavior even at room temperature. The finding from the analysis of X-ray scattering data provides direct evidence for the protein folding from the heterogeneously populated unfolded state through the multiple kinetic pathways and insights into the fundamental principles governing protein folding in view of global protein conformations and the related kinetic behaviors.

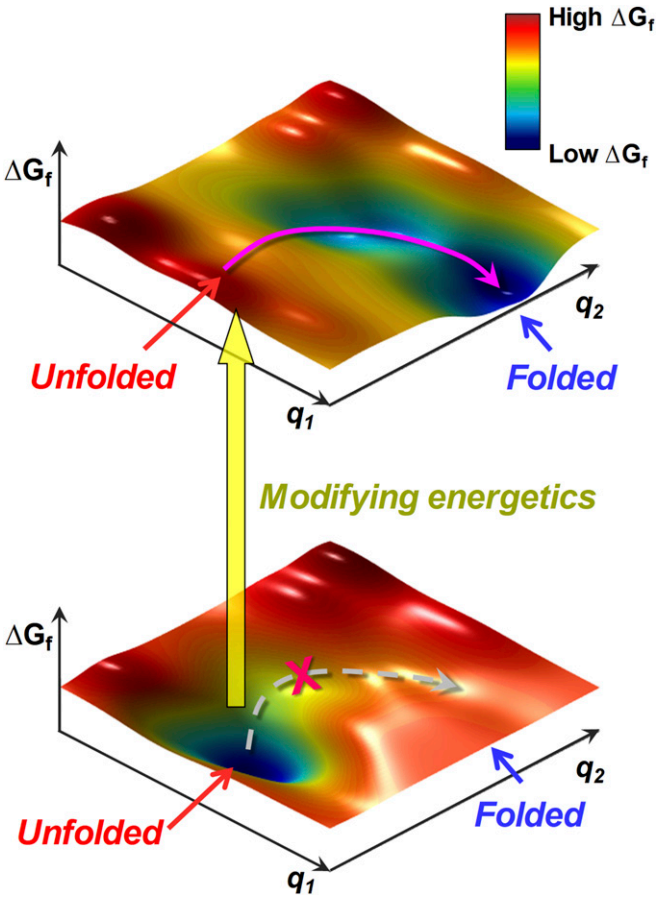


**Fig. 1.** Free-energy landscapes induced by external electron injection. Conventional free-energy surfaces for two different conditions as a function of representative reaction coordinates ($q_1$ and $q_2$). The reaction coordinates correspond to local structural changes in protein such as vibrational modes of specific residues or distances between residues. The free-energy surface can be resculpted by the change in protein environment, indicated as the yellow arrow, such as oxidation state of cofactor and solvent condition. From this, the basin of the unfolded state in the free-energy surface (*Bottom*) changes as a hillside in the resculpted energy surface (*Top*). The change of energetics activates the spontaneous folding pathways, indicated as the magenta arrow, along the reaction coordinates. As indicated with the gray dotted arrow, the pathway is inhibited in the initial condition.

## Results

**Structural Analysis of X-Ray Scattering Curves.** Prior to analyzing the TRXSS data, we first investigate the static X-ray scattering data for both the ferric state (oxidized) and the ferrous state (reduced) in order to confirm that ferric cyt-c is unfolded at the denaturation condition of 3.5 M GdnHCl and to investigate the heterogeneity of unfolded state. To overcome the limitation of adopting a single conformation as commonly employed in analyzing the structure of a well-structured protein (45–47, 52), we applied the ensemble optimization method (EOM) combined with MD simulations (53–55). Briefly, protein structures describing the conformational landscape are first generated from MD simulation snapshots and

are used to calculate theoretical X-ray scattering curves corresponding to the individual conformational states. In this analysis, the experimental X-ray scattering curve is assumed to arise from an ensemble including a certain number of coexisting conformational states. This is well suited for describing structural features of a potentially flexible system (53, 55). The schematic of this analysis and its details are described in *SI Appendix* (see *SI Appendix*, Fig. S1 and Note 1, for more information). By using a genetic algorithm (56), the ensemble that best fits the experimental data is finally chosen for extracting the optimal structural parameters such as the radius of gyration ($R_g$) and the 3D protein conformation. To guarantee wide enough conformational sampling in MD simulations, we employed three different simulation times of 100 ns, 500 ns, and 1 μs and checked the dependence of the final EOM results on the time. Specifically, we repeated the entire EOM analyses using the pools generated from six independent MD trajectories with each simulation time.

Fig. 2 *A*–*C* shows the distributions of $R_g$ derived from the EOM analysis for the static X-ray scattering curves of the oxidized and the reduced states based on different simulation times. The primitive $R_g$ distributions directly from MD simulations before EOM are also shown in Fig. 2 (*SI Appendix*, Fig. S2). Even though the maximum $R_g$ values from different pools are rather different, all of the pools cover the $R_g$ values (24–29 Å) of unfolded cyt-c reported from previous experiments (30, 57). The results from the EOM analyses show that while the $R_g$ distributions of the oxidized state optimized with six independent MD simulations quite disagree with each other with 100-ns simulation time, they eventually well overlap after 500 ns. This simply demonstrates that the conformational distribution in view of $R_g$ from the MD simulation is converging with the increase in the simulation time (*SI Appendix*, Fig. S2). Even though there are some marginal differences in the optimized ensembles from the MD simulations of 1-μs duration, it is obvious that the various unfolded conformations are heterogeneously populated in the oxidized state under 3.5 M GdnHCl. The average $R_g$ value from the EOM based on the 1-μs-long simulation changes from ∼26 Å for the oxidized state to ∼13 Å for the reduced state (*SI Appendix*, Fig. S3). Based on the representative protein conformations determined from the EOM analysis (*SI Appendix*, Figs. S4 and S5), we observe that at least four unfolded conformations with both extended and entangled structures were needed for describing the ensemble of oxidized cyt-c. According to a small-angle X-ray scattering (SAXS) study of oxidized cyt-c at differing GdnHCl concentrations (57), an equilibrium between native, partially unfolded, and fully unfolded states exists in the denaturing condition, and at 3.5 M GdnHCl the molar ratio of the three is ∼0:4:6. In addition, based on the existence of an additional singular vector from the singular value decomposition (SVD) analysis, it was vaguely presumed that one more denatured state may exist. Our model of oxidized cyt-c determined from EOM with at least four representative states has a higher heterogeneity than was proposed in this previous SAXS study (57). For the ensemble of reduced cyt-c, only one folded conformation was enough indeed. We note that the theoretical scattering curve from one individual representative structure in the ensemble of oxidized cyt-c cannot reproduce the experimental data (*SI Appendix*, Fig. S4). This aspect implies the inherent conformational heterogeneity in the oxidized cyt-c under 3.5 M GdnHCl.

**TRXSS.** The reaction scheme for the folding of equine heart cyt-c triggered by external electron transfer is shown in Fig. 3*A*. Upon 355 nm excitation of nicotinamide adenine dinucleotide (NADH) acting as the external electron source, a reactive species with solvated electron is generated that rapidly reduces the oxidized cyt-c with the unfolded protein conformation (26–29, 31, 33, 58). Based on the equilibrium unfolding curves for oxidized and reduced cyt-c (Fig. 3*B* and *SI Appendix*, Figs. S6 and S7), at neutral pH the concentration of guanidine hydrochloride (GdnHCl) in the range from 2.8 to 4.5 M induces the reduced cyt-c to be fully folded, whereas the same condition drives the oxidized form to the unfolded state with a fraction higher than 90%. This is in good agreement with the results reported from earlier studies (29, 33–35). Therefore, the concentration of 3.5 M GdnHCl is a proper condition to initiate spontaneous folding of unfolded cyt-c through the change in the redox state of the cofactor. Using the oxidized cyt-c dissolved in the phosphate buffer with 20 mM NADH and 3.5 M GdnHCl, we performed TRXSS experiments in order to investigate the folding dynamics triggered by the external electron transfer.

Time-resolved difference scattering curves, $q\Delta S(q, t)$, were measured in a wide time region covering from 31.6 μs to 316 ms (*SI Appendix*, Fig. S8). The difference scattering curves show not only a significant amplitude in the small-angle region ($q < 0.4$ Å$^{-1}$), which is related to the global conformational change, but also oscillatory features in the wide-angle region ($0.4$ Å$^{-1} < q < 2.2$ Å$^{-1}$), which can be attributed to the subtle structural change such as the rearrangement of the secondary structure (45, 47, 50) or the thermal response of solvent with the relaxation of the photoexcited molecule (45, 51). Also, from previous studies (27, 58), it is known that photoexcited NADH forms photoproducts such as $(NAD)_2$ as well as it recovers into its ground state. To discriminate the contributions by solvent heating and $(NAD)_2$ formation from the TRXSS data, we implemented an additional experiment using the buffer solution with the same composition but without the protein (see *SI Appendix*, Note 2, for details). In principle, the formation of $(NAD)_2$ and the accompanying increase in electron density may alter the scattering signal in the small-angle region ($q < 0.4$ Å$^{-1}$). The difference scattering signal of the protein-free buffer solution in the small-angle region is small and featureless, whereas the signal in the wide-angle region ($0.8$ Å$^{-1} < q < 2.2$ Å$^{-1}$) is dominant (*SI Appendix*, Fig. S8*B*). From the comparison of the difference scattering curves of the two solutions (*SI Appendix*, Fig. S9), we can verify that the effect of the formation of photoproduct on the entire $q$ region is negligible and that only the response to solvent heating is prominent in the wide-angle region. Therefore, to extract the photoinduced signal of pure protein from the time-resolved X-ray data, the contribution of solvent heating was removed by subtracting the signal that was obtained from a separate experiment on the protein-free solution by following the established protocols (45, 47, 51).

**Kinetics of Protein Folding.** Fig. 4*A* shows the time-resolved X-ray difference scattering curves, $q\Delta S(q, t)$, which are the heating-free experimental data. Quantitative kinetic analysis was conducted by using SVD to determine the rate constants that describe the time-dependent change of the heating-free data (45–47, 59). SVD decomposed the original data into time-independent components (left singular vectors), their associated time dependence (right singular vectors [RSVs]), and the relative contribution of each component (singular values). From this analysis (*SI Appendix*, Fig. S10), we identified two significant singular vectors and the corresponding autocorrelation factors, suggesting that two species discernible by X-ray scattering are formed in the measured time range (Fig. 4*B*). The simplest two-state kinetics involving two species will be a single-exponential conversion of the first species to the other one, and globally fitting with this assumption failed to give an acceptable fit (the case with $\beta = 1$ in Fig. 4*C*). This deviation points toward the existence of parallel pathways connecting the two species. We considered two candidates for two-state kinetics with parallel pathways: 1) bifurcated kinetics from the first to the second species with two exponentials, which can be considered as the simplest parallel kinetics, and 2) stretched exponential kinetics with $e^{-(t/\tau)^{\beta}}$, as an extreme case with parallel pathways, in which the β value stands

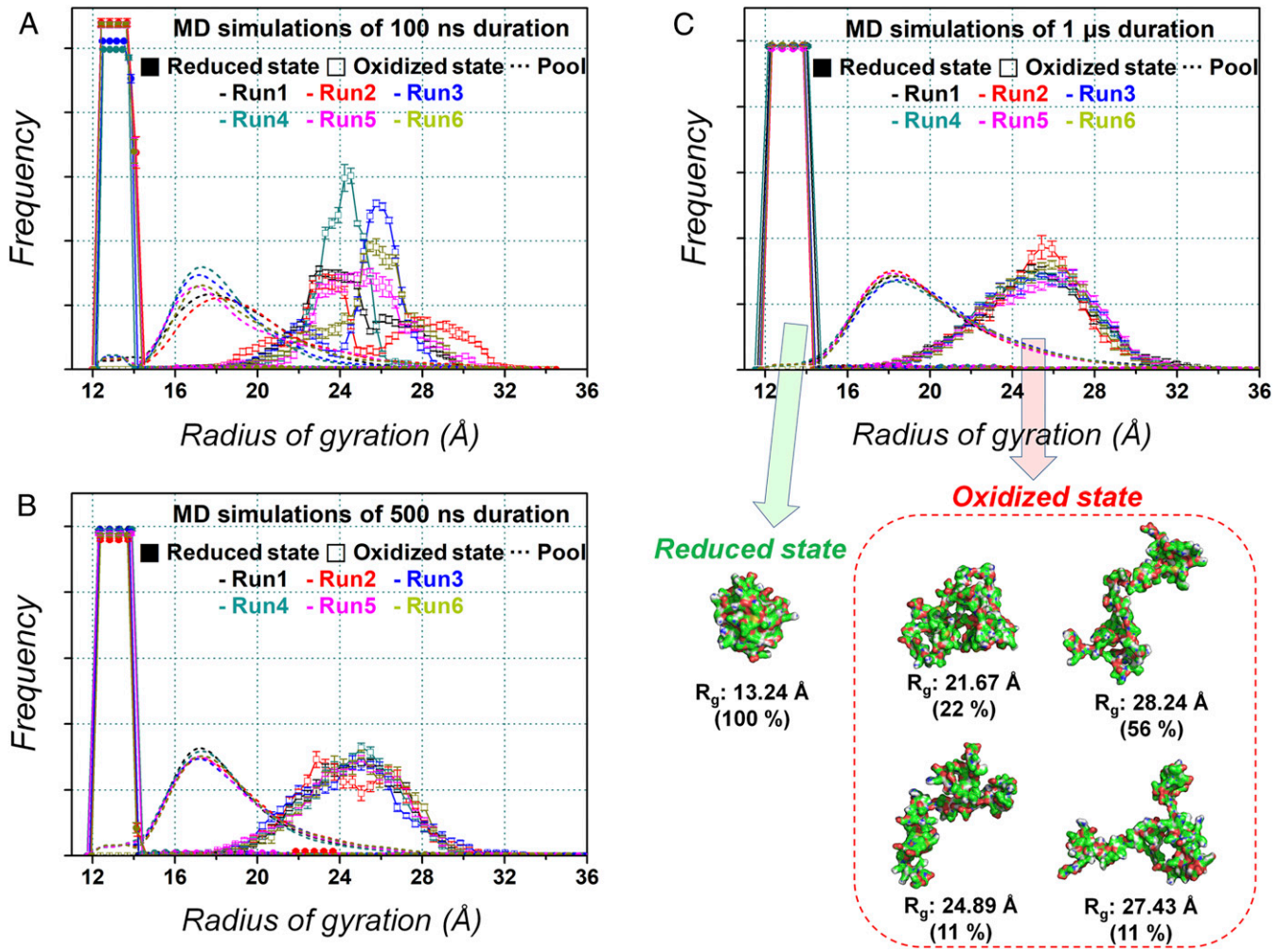


**Fig. 2.** Structural analysis of X-ray scattering curves based on EOM. (*A*–*C*) Distributions of $R_g$ determined from the EOM analysis for the static X-ray scattering curves of the oxidized and the reduced states based on unfolding MD simulation times of (*A*) 100 ns, (*B*) 500 ns, and (*C*) 1 μs. For each simulation time, a total of six independent trajectories were sampled by starting from one common structure. Each panel shows the $R_g$ distributions of finally optimized ensembles (squares) as well as the conformational space coverage by the candidate protein structures from the MD simulations (dotted lines). The theoretical scattering curves of the final ensembles well fit the experimental curves within the error bars (*SI Appendix*, Fig. S4). As the simulation time for sampling the candidate structures increases, the discrepancy among the six unfolded distributions optimized with the six trajectories becomes inconspicuous. (*C*, *Bottom*) The representative structures of folded and unfolded state with those relative populations. It is clear that various unfolded conformations are heterogeneously populated in the optimized unfolded ensemble (open squares) and single conformation is homogeneously populated in the optimized folded ensemble (closed squares).

for describing the extent of superposition of many exponential decays (60). Both the biexponential with timescales of 16.4 (±5.1) ms and 193 (±34) ms and the stretched exponential with $\tau = 185$ (±75) ms and $\beta = 0.70$ (±0.08) show good agreements with RSVs (*SI Appendix*, Fig. S10*B*). We note that the bifurcated kinetics is different from the consecutive sequential kinetics suggested in earlier optical spectroscopic studies (26, 28, 29), where three species were required to explain two time constants. Because the SVD analysis suggests that only two species formed in the measured time region, we ruled out the possibility of this sequential kinetics (*SI Appendix*, Fig. S11).

On the basis of both stretched exponential and bifurcated kinetics models, we performed principal component analysis using the time constants determined from the SVD analysis. The details on this kinetic analysis are described in *SI Appendix* (see *SI Appendix*, Note 3, for more information). From this analysis, we extracted the species-associated difference scattering curves (SADSs) and time-dependent contributions of SADSs. Those determined from the stretched exponential kinetics are displayed in Fig. 4. Each SADS contains information on the structure of a protein intermediate. The first SADS formed in the microsecond regime has the monotonous feature with a small negative amplitude around 0.1 $Å^{-1}$ in the difference scattering curve, whereas the second SADS formed in the late millisecond regime has a prominent oscillatory feature in the $q$ space from 0.05 to 0.8 $Å^{-1}$. Even though the first SADS has a much smaller contribution than the second one, the omission of the first species in the folding kinetics results in a worse agreement between the experimental and the theoretical data (*SI Appendix*, Fig. S12), indicating that the cyt-c folding should involve the first species. The small decrease of the scattering intensity in the small-angle region of the first SADS implies that the protein conformation of the first intermediate is slightly more unfolded relative to that of the initial state. To identify the origin of the conformational change in the transition from the first SADS to the second one, we compared the difference of the static scattering

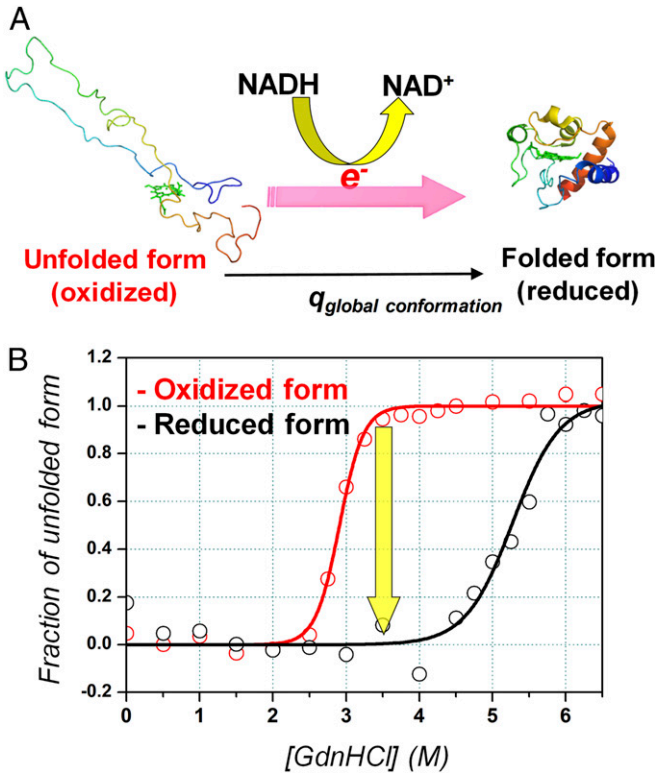


**Fig. 3.** Schematic of folding dynamics of cyt-c induced by external electron injection. (*A*) Unfolded cyt-c with the oxidized heme group undergoes folding dynamics triggered by the electron injection from NADH acting as an external electron donor to the heme group in cyt-c. (*B*) Equilibrium unfolding curves for oxidized ($Fe^{3+}$) and reduced ($Fe^{2+}$) cyt-c as a function of molar concentration of GdnHCl. These curves were reconstructed from the circular dichroism spectroscopic data. The TRXSS experiment was performed by employing the photoinduced electron transfer under the concentration of 3.5 M GdnHCl.

curves of the reduced and the oxidized forms [$\Delta S_{static}(q) = S_{reduced}(q) - S_{oxidized}(q)$] with the appropriately scaled second SADS as shown in Fig. 4*B*. The good agreement between them indicates the formation of the folded state through the transition from the first to the second species. When the scattering curves were reconstructed from the linear combination of SADSs obeying the stretched exponential kinetics, they exhibited excellent agreements with the experimental curves (Fig. 4*A*), confirming that cyt-c folding faithfully follows non-exponential kinetics. In fact, almost the same SADSs are also obtained from the bifurcated kinetics model because both cases are indistinguishable in the trace (*SI Appendix*, Fig. S11). Therefore, the choice between the stretched exponential and the bifurcated kinetics in explaining the cyt-c folding dynamics cannot be made solely based on the kinetic trace. For this reason, we implemented an additional analysis based on the maximum entropy method (MEM) whereby the time-dependent signal can be explained by a distribution of time constants (61). From this analysis, the optimal distribution of log-time constants was extracted from the experimental data (see *SI Appendix*, Note 3 and Fig. S13, for details). The broad distribution of log-time constants covering from the early to the late millisecond regime supports that the experimental trace likely follows the stretched exponential kinetics rather than the bifurcated one. The details about the folding kinetics in this regard are given in *Discussion*.

## Discussion

The kinetic analysis shows that the TRXSS data can be explained by the following scheme for the folding dynamics induced by the redox state change of the cofactor:

$$\mathbf{U}(\mathbf{Fe}^{3+}) \xrightarrow{<31.6\mu s} \mathbf{U}'(\mathbf{Fe}^{2+}) \rightarrow \mathbf{F}(\mathbf{Fe}^{2+}).$$

Here U is the initial unfolded state prior to the photoreduction of cyt-c, and U′ and F are the intermediate and the final photoproduct involved in the folding dynamics, respectively. The results from the EOM analysis for the static scattering curve of oxidized cyt-c shown in Fig. 2 suggest that the initial U state with an oxidized form of the iron-containing heme group has the conformational heterogeneity including various unfolded structures. The U′ intermediate is formed within 31.6 μs, which is the first time delay of our experiment. Considering the time of ∼33 μs taken to reduce the oxidized cyt-c with solvated electron and NAD radical produced from NADH photoexcitation (27, 58), the U′ state likely possesses a reduced form of the iron-containing heme group. In addition, the decrease in scattering amplitude in the small-angle region of the first SADS indicates that its volume and size slightly increases in comparison with the U state. From the comparison of the second SADS and the difference between static X-ray scattering curves of the oxidized and the reduced cyt-c, we find that the final F state has the completely folded conformation with well-organized secondary structures and reduced heme.

As mentioned already, based on the global kinetic analysis of TRXSS data, the main folding event in the transition from U′ to F can be explained by both stretched exponential and bifurcated kinetic models. This is from a typical uncertainty problem in the experimental data with a limited signal-to-noise ratio. While the MEM results favor the stretched exponential kinetics over the bifurcated one, having additional logical arguments other than the fitting quality in the kinetic analysis will be desirable. To address this, we consider the structural information available from the EOM analysis for the static scattering data. The ensemble of oxidized cyt-c linked to the initial U state shows the prominent structural heterogeneity including at least four representative conformations with both extended and entangled structures (Fig. 2 and *SI Appendix*, Fig. S4). If we apply the stretched exponential kinetics, the different unfolded conformations with the $R_g$ values of 20 to 30 Å in the U state individually undergo reduction to the structurally heterogeneous U′ state, which is kinetically expressed as one representative species. After this, various unfolded conformations in the U′ state will likely adopt different folding pathways with different folding times toward the homogeneous folded state F with $R_g$ of 13.2 Å. The superposition of many parallel pathways with the different timescales is compatible with the stretched exponential behavior. When we assume bifurcated kinetics, two parallel pathways may be attributed to the folding event from the heterogeneously populated U′ state to the folded state. As an example, some portion of the unfolded structures may possibly rapidly fold without a detour to a misfolded state, and the other remaining portion may progress through a misfolding pathway toward the native state. Namely, the two different timescales can originate from the presence or absence of the misfolded state in the folding pathway. Although this hypothetical situation may support the bifurcated kinetics, it is unclear how the extremely heterogeneous conformations in the U′ state will favor only two pathways toward the final folded state. It will be more persuading that the heterogeneous unfolded conformations adopt multiple pathways that potentially pass through various misfolded conformations on the free-energy surface. The superposition of such various folding paths actually merges to a single stretched exponential kinetics, and therefore, we suggest that the

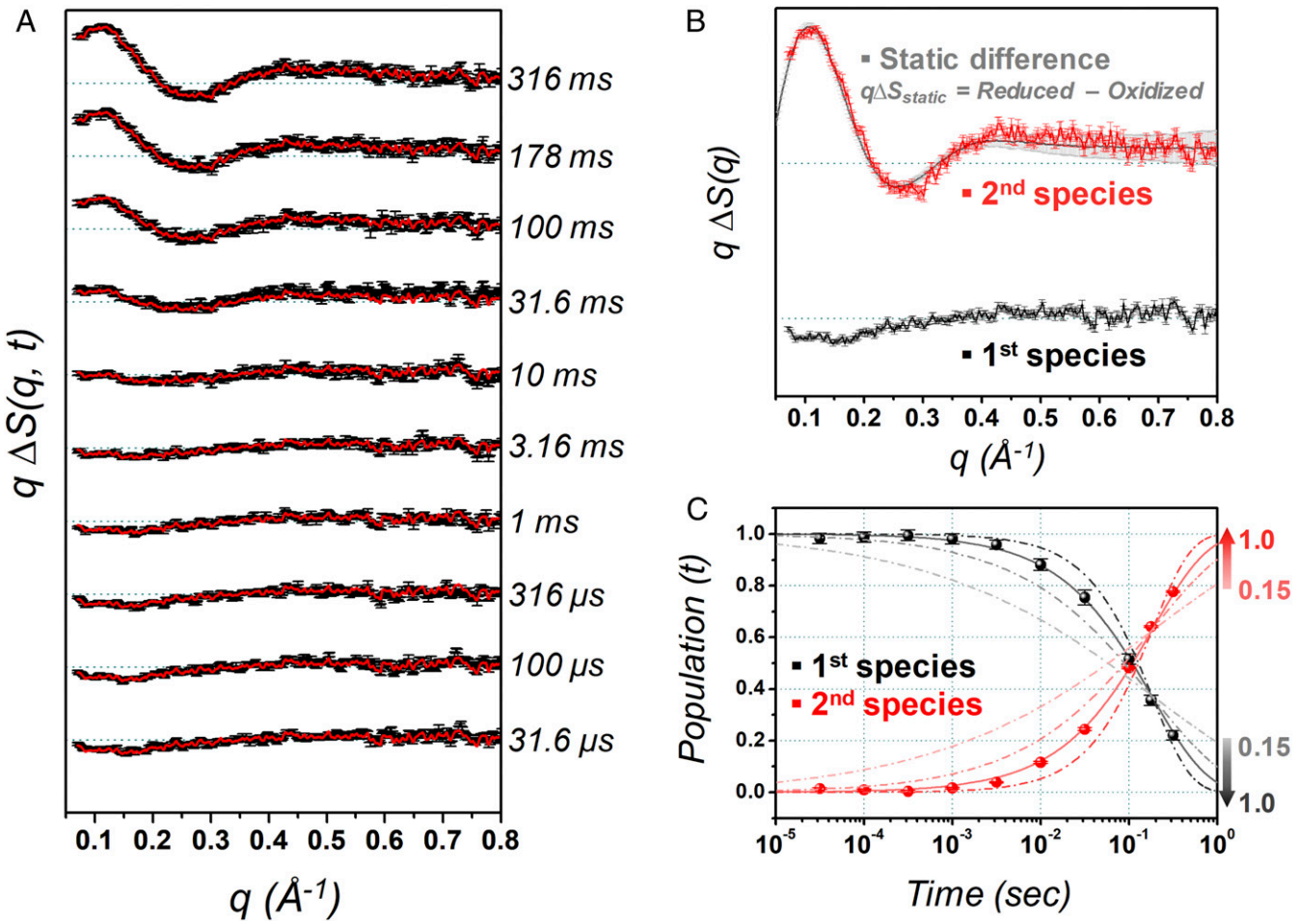


**Fig. 4.** TRXSS data for the folding dynamics of cyt-c. (*A*) Experimental (black) and theoretical (red) difference scattering curves. (*B*) SADSs extracted from the global kinetic analysis. The second SADS (red) is compared with the difference for the static X-ray scattering curves of reduced and oxidized forms (gray). The high similarity between the second SADS and the static difference, $q\Delta S_{static}$, confirms that the observed dynamics is linked with the protein folding. (*C*) Time-dependent population change of the intermediates (dots) is compared with the theoretical population dynamics (lines) based on the stretched exponential kinetics. The values shown on the right side of graph represent the β values in the stretched exponential, $e^{-(t/\tau)^{\beta}}$, ranging from 0.15 to 1.0. The experimental population dynamics is in good agreement with the stretched exponential kinetics with the time constant (τ) of 185 (±75) ms and the β value of 0.70 (±0.08).

stretched exponential will be the operational kinetics in the cyt-c folding.

Previous studies of cyt-c folding suggested the presence of kinetic phases in the microsecond time regime called as burst phase (*SI Appendix*, Fig. S14). Using time-resolved circular dichroism (TRCD) spectroscopy on electron-transferred cyt-c folding, Kliger and coworkers reported the time constant of 5 μs in the burst phase and attributed it to the rapid formation of partial secondary structures linked with the collapse of extended random coil (33). This earliest event was similarly observed with time-resolved optical rotatory dispersion (TRORD), where the fast phase process was assigned to the misligation of His33 (26, 28, 29). These studies suggested that the early-phase intermediate generated with the sudden collapse of random coil converts directly to the native state during the slow phase in hundreds of milliseconds. In terms of timescales, the U′ state observed in our study can be compared to the misligated state accompanying the partial formation of the secondary structure observed by TRCD and TRORD (26, 28, 29, 33). The first SADS, however, exhibits a decrease of the scattering amplitude in the small-angle region, which implies a slight increase of the protein volume and size in comparison with the U state. In other words, this slightly expanded U′ state, occurring before the rate-determining folding, is structurally distinguished from the molten globule state that possesses more partially folded conformations formed by the sudden collapse of the random coil (30, 41, 62). This difference is likely a consequence of adopting a strong denaturing condition with 3.5 M GdnHCl. According to an earlier fluorescence study (41), the relative amplitude of the early-phase kinetic trace of fluorescence significantly decreases with an increase in the concentration of GdnHCl. This is natural as GdnHCl tends to destabilize compact states that may form in the initial stage of folding (30, 57). In addition, destabilization of higher charged heme after oxidation in the low dielectric protein environment may be mitigated at high GdnHCl concentrations (63), and the reduction of the heme may not energetically prefer the early-phase protein motion necessary for forming the molten globule state such as the contraction of the backbone chain.

The kinetic analysis of TRXSS data shows that the U′ state undergoes transition to the final state with the time constant of 185 ms. In the TRCD study with multiexponential fitting, the kinetics in this time regime showed a decay with a time constant of 110 ms, which was attributed to the displacement of misligated residues at the heme cofactor (33), although the uncertainty is rather large due to the limited signal-to-noise ratio in the TRCD data. Using TRORD spectroscopy, a single decay of 160 ms in

the slow phase was observed at a slightly lower GdnHCl concentration of 3.3 M and was assigned to the formation of the native structure with the ligation process between Met-80 and the heme (29). A similar folding time of 185 ms was reported for tuna heart cyt-c, containing a tryptophan residue at the site occupied by His33 in equine heart cyt-c (28). With transient absorption spectroscopic experiments, it was shown that the slow phase folding of reduced cyt-c follows a single exponential decay with a lifetime of hundreds of milliseconds that coincides with the change in tryptophan fluorescence (31, 32). Our time constant of 185 ms in the stretched exponential kinetics identified by TRXSS relatively well matches those from the previous studies (*SI Appendix*, Fig. S14). However, we note that the heterogeneous transition of U′ → F expressed as the stretched exponential is in stark contrast to the normal exponential behavior interpreted as sequential transitions of intermediates with the conformational homogeneity in earlier studies (26, 28, 29, 31, 32). On the other hand, a stopped-flow mixing study for the reduced cyt-c combined with spectroscopic measurements observed two phases in the refolding kinetics when the GdnHCl concentration was smaller than ∼1.5 M, although at 3.5 M GdnHCl, which is the same concentration used in our work, only a single phase was observed (42). The two phases could be explained by either the conformational heterogeneity of unfolded cyt-c or the kinetic partitioning of the homogenous unfolded state (64). Our observation on the global conformational changes supports the former scenario even at the high GdnHCl concentration condition. This distinction is possible as TRXSS directly detects the overall structure unlike typical spectroscopic tools.

Indeed, the heterogeneous transition to the final folded state in slow phase folding of cyt-c may represent the experimental epitome of the funnel-like free-energy landscape in protein folding, which assumes multiple transitions from different microstates in the unfolded state to a unique global minimum of the folded state (12). A plausible scenario consistent with our observation is that the different unfolded structures in the initial oxidized state adopt conformation-specific pathways and stop by on-route intermediate configurations located at different positions of the free-energy landscape. Specifically, at an early stage of folding, some fractions of unfolded protein structures will be forming native interactions and are led to the correctly folded state, whereas others will face trap sites at local minima on the folding energy landscape. The differences in folding pathways originating from the differences in the unfolded conformations will likely bring about different folding times for each folding pathway. The stretched exponential kinetics observed with cyt-c in the slow phase can be interpreted as the sum of many single exponentials, and this aspect supports the kinetic partitioning of the heterogeneous unfolded state in folding, which may arise from the rugged features around the high-energy region of folding free-energy landscape. Our scenario for the folding dynamics of cyt-c can also be connected to the highly stretched exponential behavior in the early phase of folding, reported from a low-temperature stopped-flow study for the refolding of ferric cyt-c (18). As noted earlier, this behavior will correspond to the many parallel pathways initiating from as many different unfolded conformations. Thus, our results provide clear evidence for the heterogeneous folding of cyt-c involving structurally distinct multiple unfolded conformations and are in contrast to the view with unidirectional sequential pathways from a homogeneously defined unfolded state discussed in earlier studies (26–29, 31, 33, 34). A previous SAXS study reported that the high concentration of GdnHCl increased the population of fully unfolded state involving a high content of random coils (57), and another spectroscopic study showed that as the concentration of GdnHCl increased, folding events on the timescale of hundreds of milliseconds slowed down at room temperature (31). If these results are applied to our observation, we expect that the folding time may change rather sensitively with the GdnHCl concentration as it will affect the population ratio of microscopic unfolded states and their thermodynamic energetics. For example, at a higher concentration of GdnHCl the population of unfolded conformation with larger $R_g$ values and extended structure will increase, which may bring about the increase of folding time correlated with the unfolded conformation of cyt-c.

We note that the chain diffusion timescale in cyt-c is several tens of microseconds (65). Considering this relatively fast backbone diffusion, the slow interconversion among many conformers in the U′ state may seem incompatible. This is resolved when they cannot be interconverted by chain diffusion alone. There are at least two factors that contribute to increasing the barriers between the unfolded conformers in the U′ state: 1) heme–residue misligation and 2) nonnative configurations of proline residues. In the case of misligated conformers, the interconversion toward a native-like ligation requires the displacement of the misligated residues, which is not likely to occur by the chain diffusion. In addition, we cannot rule out the existence of nonnative configurations of proline residues in the unfolded cyt-c. According to previous studies (42, 66), the timescale for transcis isomerization of proline is much larger than the chain diffusion time. Overall, we speculate that the origin of the heterogeneous kinetics from U′ to F is the heterogeneity of unfolded conformers in the U′ state due to misligating residue–heme interactions and the proline isomerization which cannot be interconverted only by fast chain diffusion. This aspect eventually results in a broad distribution of folding times.

We also note that we have utilized MD simulations for generating decoy structures for the EOM process based on selections with the genetic algorithm. Thus, how well the structural ensemble from the MD trajectories represents the folded and the unfolded state is not of concern, as long as the trajectories can cover the thermally accessible configuration space well within the limit of accessible simulation times. Of course, it would be ideal if the simulated ensemble could directly match the scattering data without any ensemble adjustment through EOM. In fact, when we compare the $R_g$ distribution from any structure pool presented in Fig. 2 *A–C* against the distribution after its EOM treatment for the unfolded state, we find that the structure pool data display higher probabilities on the smaller $R_g$ side. Namely, the MD structures are generally more compressed than the actual experimental structures, which was also reported by the theoretical study of disordered proteins (67). Thus, we can infer that the adopted force field parameters unphysically prefer more compact structures even at the high temperature (800 K) employed for our simulations. It is likely that reproducing folded state characters were more importantly considered while optimizing the force field. Improving it such that the characters of less structured states are better reproduced will be important for studying proteins in general, especially with intrinsically disordered proteins. Our approach may provide a useful tool toward this end as it generates a straightforward target for benchmarking. Namely, by comparing the theoretical scattering data based on the simulated ensemble against the actually measured experimental figure, one can evaluate the fidelity of the adopted force field model for both folded and unfolded states. Although our EOM analysis scheme with MD simulations successfully extracted the structural information about the heterogeneous ensemble of unfolded cyt-c, there remains room for further improving on how to generate the candidate protein conformations. For example, the molecular form factor (MFF) analysis can be an alternative way for denatured and intrinsically disordered proteins as SAXS data can be analyzed by predetermined dimensionless MFFs without any additional MD simulations (68).

## Conclusions

In summary, using TRXSS, we were able to establish the folding dynamics of cyt-c with the nonexponential kinetics in the

transition between two states (U′ and F). The combination of the results obtained from the kinetic and structural analyses of X-ray scattering data revealed that the various protein conformations in the unfolded state take the multiple relaxation pathways toward the folded conformation, resulting in stretched exponential behaviors. Fig. 5 summarizes the detailed scheme for the spontaneous folding dynamics of cyt-c, especially highlighting the multiple relaxation pathways linked with the conformational heterogeneity in the unfolded state. This observation is direct evidence of protein folding from the heterogeneous unfolded states in view of microscopic protein conformations. The populations that can form nonnative ligation on the heme group are likely to induce various misfolded structures located in local minima in the folding free energy. Fast escaping from these minima will result in fast folding kinetics, while getting trapped in deep energy minima will require much longer time and detours toward the native state. The existence of all those complex pathways, associated with various states of heme ligations and the related conformers, will surely add a level of heterogeneity to its folding and unfolding dynamics. When such complicating factors combine together, they will contribute to rendering the kinetics to appear diverse with a stretched exponential behavior for cyt-c. Thus, how much our conclusion regarding the folding path diversity can be generalized should be considered as an open question, and answering that question in the future will definitely be interesting and important. We believe that our approach can surely contribute toward addressing this by further studying other folding systems. If the TRXSS experiment is implemented at various temperatures as employed in previous studies (14, 18), it may allow us to estimate the ruggedness of the folding free-energy landscape and to unravel the conformational heterogeneity in the early phase. This addition will ultimately facilitate topologically mapping the overall folding pathways in view of the microstate ensemble and its conformations. Such a capability will also be important for developing new computational models. In fact, constructing protein interaction models that can properly describe unstructured states is not a trivial task, primarily because experimental data that can validate the models are severely lacking. Unstructured protein states are more affected by interfacial behaviors such as solvent exposure and related polarization effects on protein surfaces, for which modern models are still being heavily developed. Our method can provide a convenient tool for evaluating their reliability. Thus, experimental developments in this area may contribute well toward advancing prediction powers of newly designed theoretical tools.

## Methods

**Sample Preparation.** Equine heart cyt-c purchased from Sigma Aldrich was dissolved in 100 mM of Na-phosphate 20 mM, 20 mM of NADH, and 3.5 M of GdnHCl at pH 7.0. A centrifuge with the cellulose acetate membrane filter was used to remove the aggregated particles in the protein solution. After that, the protein solution was bubbled using nitrogen gas for 30 min in order to make an oxygen-free buffer. The protein samples were prepared just before the X-ray scattering experiments.

**Steady-State Spectroscopic Measurements.** Prior to starting the time-resolved experiment, we performed the steady-state spectroscopic experiments using UV-visible absorption and UV circular dichroism (UV-CD) spectroscopies (JASCO J-815) to check the state of the protein sample.

**Measurement of X-Ray Scattering Curve.** Steady-state X-ray scattering data for the oxidized and the reduced form of cyt-c were collected at the 14-ID-B beamline at Advanced Photon Source. For the steady-state measurement for the reduced cyt-c, the oxidized cyt-c was reduced by the addition of sodium dithionite (4 mM) into the oxygen-free buffer. During the measurement, a



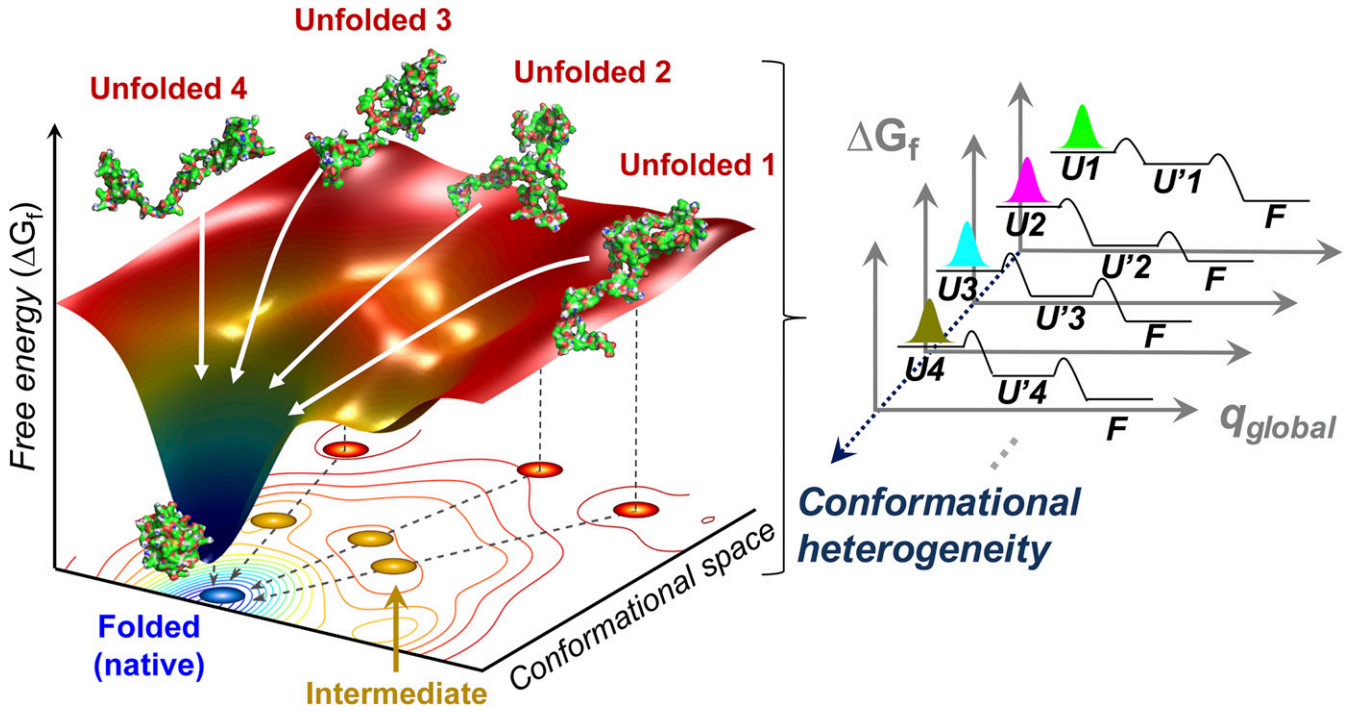


**Fig. 5.** Schematics of the folding dynamics in a funnel-like free-energy landscape. Upon the photoreduction of cyt-c, the heterogeneously populated initial unfolded state (red dots in the projected plane) undergoes spontaneous folding along the multiple parallel pathways (white arrows). Based on the EOM analysis with one MD trajectory over 1-μs duration, four conformations were obtained to represent the unfolded state. The number of representative unfolded conformations somewhat depends on the trajectory identity. The rugged features around the top of the free-energy funnel may be associated with the existence of intermediates (yellow dots in the projected plane) that are visited in the early phase of folding. The observed stretched exponential folding kinetics is the result of the superposition of the multiple pathways from the unfolded substates ($U_i$, $i = 1$ to $N$) to the folded state ($F$) via intermediate substates ($U'_i$, $i = 1$ to $N$), indicating the conformational heterogeneity in the cyt-c folding.

1-mM cyt-c solution dissolved in 100 mM Na-phosphate, 20 mM NADH, and 3.5 M GdnHCl at pH 7.0 was flowed to the capillary flow cell in order to avoid the signal from the damaged sample. The X-ray probe pulse at 12 keV, generated from the 324-bunch mode operation at APS, was delivered to the protein sample, and X-ray scattering patterns were recorded using a Mar165 CCD detector (Rayonix). The energy profile of X-ray pulse used for the collection of scattering pattern was measured by using a photodiode and was used for the structural analysis of X-ray scattering data in order to correct the polychromaticity in theoretical scattering curve (see *SI Appendix*, Note 1 and Fig. S15, for details).

TRXSS data were collected at the ID09B beamline at the European Synchrotron Radiation Facility (ESRF) by following an established experimental protocol based on the pump-and-probe scheme (45–47, 50, 51). For the measurement, a 2-mM cyt-c solution dissolved in the buffer, which is the same as that used in the steady-state measurement, was flowed into the capillary system by using a syringe pump. To prevent reoxidation during the time-resolved measurement, we used an airtight circulation system where the bulk solution was fully degassed. Furthermore, to fundamentally remove the possibility that the folded form (reduced form) generated as photoproducts can undergo reoxidation, the sample portion exposed to the excitation laser and the X-ray pulses were flown through and discarded. Namely, the protein sample was exposed for laser and X-ray pulses only once. The cyt-c solution was irradiated by circularly polarized nanosecond (ns) laser pulses (Opotek Vibrant) with the fluence of 1 mJ/mm$^2$ at 355 nm and was probed by X-ray pulses incident at well-defined time delays. Polychromatic X-ray probe pulse with the energy of 18 keV was generated from the 16-bunch mode operation at ESRF. Time-dependent X-ray scattering patterns were recorded using a Mar165 CCD detector (Rayonix), covered up to the wide-angle region. The scattering pattern at a negative time delay (−50 μs) was also collected and used as a reference for calculating time-resolved difference scattering curves. The scattering pattern at −50-μs time delay contains structural information of the initial state, while the data at positive time delays contain the contributions from a mixture of the initial state and photoproducts. The time-resolved difference scattering curves, $\Delta S(q, t) = S(q, t) - S(q, -50\ \mu s)$, were generated by subtracting the scattering curve at −50 μs from the curves at positive time delays. The $\Delta S(q, t)$ curves provide the information on the change caused by the laser photoexcitation. We measured the TRXSS curves in a wide time range from 31.6 μs to 316 ms. At each time delay, more than 300 scattering curves were averaged to achieve a high signal-to-noise ratio.

**Data Availability.** All data are included in the manuscript and *SI Appendix*.

**ACKNOWLEDGMENTS.** We acknowledge inspiring discussion with Dr. Hyun Woo Kim of Korea Research Institute of Chemical Technology, extensive support from Dr. Irina Kosheleva and Dr. Robert Henning of 14-ID beamline at Advanced Photon Source during data collection there, and the initial experimental help from Marco Cammarata and Theyencheri Narayanan at ESRF. The time-resolved and static X-ray solution scattering experiments in this study were performed at the ID09 beamline at ESRF and at the BioCARS 14-ID beamline at the Advanced Photon Source (United States), respectively. Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under Contract DE-AC02-06CH11357. Use of the BioCARS Sector 14 was also supported by the National Institutes of Health, National Institute of General Medical Sciences Grant R24GM111072. This work was supported by the National Research Foundation of Korea (2017R1A2B3004946). This work was supported by the Institute for Basic Science (IBS-R004).

1. C. B. Anfinsen, Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
2. C. Levinthal, ""How to fold graciously"" in *Mossbauer Spectroscopy in Biological Systems*, P. Debrunner, J. Tsibris, E. Münck, Eds. (University of Illinois Press, Urbana, 1969), Vol. 67, pp. 22–24.
3. A. Sali, E. Shakhnovich, M. Karplus, How does a protein fold? *Nature* **369**, 248–251 (1994).
4. J. D. Bryngelson, J. N. Onuchic, N. D. Socci, P. G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins* **21**, 167–195 (1995).
5. J. N. Onuchic, P. G. Wolynes, Z. Luthey-Schulten, N. D. Socci, Toward an outline of the topography of a realistic protein-folding funnel. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3626–3630 (1995).
6. N. D. Socci, J. N. Onuchic, P. G. Wolynes, Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.* **104**, 5860–5868 (1996).
7. K. A. Dill, H. S. Chan, From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10–19 (1997).
8. C. M. Dobson, A. Šali, M. Karplus, Protein folding: A perspective from theory and experiment. *Angew. Chem. Int. Ed. Engl.* **37**, 868–893 (1998).
9. S. S. Plotkin, J. N. Onuchic, Investigation of routes and funnels in protein folding by free energy functional methods. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6509–6514 (2000).
10. C. Hardin, M. P. Eastwood, M. Prentiss, Z. Luthey-Schulten, P. G. Wolynes, Folding funnels: The key to robust protein structure prediction. *J. Comput. Chem.* **23**, 138–146 (2002).
11. P. E. Leopold, M. Montal, J. N. Onuchic, Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8721–8725 (1992).
12. P. A. Ellison, S. Cavagnero, Role of unfolded state heterogeneity and en-route ruggedness in protein folding kinetics. *Protein Sci.* **15**, 564–582 (2006).
13. D. V. Fedyukina, S. Cavagnero, Protein folding at the exit tunnel. *Annu. Rev. Biophys.* **40**, 337–359 (2011).
14. J. Sabelko, J. Ervin, M. Gruebele, Observation of strange kinetics in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6031–6036 (1999).
15. B. Gillespie, K. W. Plaxco, Nonglassy kinetics in the folding of a simple single-domain protein. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12014–12019 (2000).
16. M. M. Garcia-Mira, M. Sadqi, N. Fischer, J. M. Sanchez-Ruiz, V. Muñoz, Experimental identification of downhill protein folding. *Science* **298**, 2191–2195 (2002).
17. B. A. Krantz, L. Mayne, J. Rumbley, S. W. Englander, T. R. Sosnick, Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *J. Mol. Biol.* **324**, 359–371 (2002).
18. S. Saigo, N. Shibayama, Highly nonexponential kinetics in the early-phase refolding of proteins at low temperatures. *Biochemistry* **42**, 9669–9676 (2003).
19. W. Y. Yang, J. W. Pitera, W. C. Swope, M. Gruebele, Heterogeneous folding of the trpzip hairpin: Full atom simulation and experiment. *J. Mol. Biol.* **336**, 241–251 (2004).
20. H. Ma, M. Gruebele, Kinetics are probe-dependent during downhill folding of an engineered lambda6-85 protein. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2283–2287 (2005).
21. P. Li, F. Y. Oliva, A. N. Naganathan, V. Muñoz, Dynamics of one-state downhill protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 103–108 (2009).
22. J. Stigler, F. Ziegler, A. Gieseke, J. C. M. Gebhardt, M. Rief, The complex folding network of single calmodulin molecules. *Science* **334**, 512–516 (2011).
23. A. J. Lee, W. B. Asher, H. A. Stern, K. L. Bren, T. D. Krauss, Single-molecule analysis of cytochrome *c* folding by monitoring the lifetime of an attached fluorescent probe. *J. Phys. Chem. Lett.* **4**, 2727–2733 (2013).
24. R. M. Abaskharon, R. M. Culik, G. A. Woolley, F. Gai, Tuning the attempt frequency of protein folding dynamics via transition-state rigidification: Application to Trp-cage. *J. Phys. Chem. Lett.* **6**, 521–526 (2015).
25. C. M. Jones *et al.*, Fast events in protein folding initiated by nanosecond laser photolysis. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11860–11864 (1993).
26. E. Chen, C. J. Abel, R. A. Goldbeck, D. S. Kliger, Non-native heme-histidine ligation promotes microsecond time scale secondary structure formation in reduced horse heart cytochrome c. *Biochemistry* **46**, 12463–12472 (2007).
27. S. Nishida, T. Nada, M. Terazima, Kinetics of intermolecular interaction during protein folding of reduced cytochrome c. *Biophys. J.* **87**, 2663–2675 (2004).
28. E. Chen, R. A. Goldbeck, D. S. Kliger, The earliest events in protein folding: A structural requirement for ultrafast folding in cytochrome C. *J. Am. Chem. Soc.* **126**, 11175–11181 (2004).
29. E. F. Chen, R. A. Goldbeck, D. S. Kliger, Earliest events in protein folding: Submicrosecond secondary structure formation in reduced cytochrome c. *J. Phys. Chem. A* **107**, 8149–8155 (2003).
30. S. Akiyama *et al.*, Conformational landscape of cytochrome c folding studied by microsecond-resolved small-angle x-ray scattering. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1329–1334 (2002).
31. T. Pascher, Temperature and driving force dependence of the folding rate of reduced horse heart cytochrome c. *Biochemistry* **40**, 5812–5820 (2001).
32. J. R. Telford, F. A. Tezcan, H. B. Gray, J. R. Winkler, Role of ligand substitution in ferrocytochrome c folding. *Biochemistry* **38**, 1944–1949 (1999).
33. E. F. Chen, P. Wittung-Stafshede, D. S. Kliger, Far-UV time-resolved circular dichroism detection of electron-transfer-triggered cytochrome c folding. *J. Am. Chem. Soc.* **121**, 3811–3817 (1999).
34. J. R. Telford, P. Wittung-Stafshede, H. B. Gray, J. R. Winkler, Protein folding triggered by electron transfer. *Acc. Chem. Res.* **31**, 755–763 (1998).
35. T. Pascher, J. P. Chesick, J. R. Winkler, H. B. Gray, Protein folding triggered by electron transfer. *Science* **271**, 1558–1560 (1996).
36. D. Rimmerman *et al.*, Probing cytochrome c folding transitions upon phototriggered environmental perturbations using time-resolved X-ray scattering. *J. Phys. Chem. B* **122**, 5218–5224 (2018).
37. S. J. Hagen, J. Hofrichter, A. Szabo, W. A. Eaton, Diffusion-limited contact formation in unfolded cytochrome c: Estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 11615–11617 (1996).
38. D. Thirumalai, D. K. Klimov, S. A. Woodson, Kinetic partitioning mechanism as a unifying theme in the folding of biomolecules. *Theor. Chem. Acc.* **96**, 14–22 (1997).
39. C. K. Chan *et al.*, Submillisecond protein folding kinetics studied by ultrarapid mixing. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1779–1784 (1997).
40. T. R. Sosnick, L. Mayne, R. Hiller, S. W. Englander, The barriers in protein folding. *Nat. Struct. Biol.* **1**, 149–156 (1994).
41. M. C. Shastry, H. Roder, Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nat. Struct. Biol.* **5**, 385–392 (1998).
42. A. K. Bhuyan, J. B. Udgaonkar, Folding of horse cytochrome c in the reduced state. *J. Mol. Biol.* **312**, 1135–1160 (2001).

43. H. Ki, K. Y. Oang, J. Kim, H. Ihee, Ultrafast X-ray crystallography and liquidography. *Annu. Rev. Phys. Chem.* **68**, 473–497 (2017).
44. Y. Chen *et al*., Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 334–339 (2017).
45. T. W. Kim *et al*., Combined probes of X-ray scattering and optical spectroscopy reveal how global conformational change is temporally and spatially linked to local structural perturbation in photoactive yellow protein. *Phys. Chem. Chem. Phys.* **18**, 8911–8919 (2016).
46. J. G. Kim, T. W. Kim, J. Kim, H. Ihee, Protein structural dynamics revealed by time-resolved X-ray solution scattering. *Acc. Chem. Res.* **48**, 2200–2208 (2015).
47. T. W. Kim *et al*., Protein structural dynamics of photoactive yellow protein in solution revealed by pump-probe X-ray solution scattering. *J. Am. Chem. Soc.* **134**, 3145–3153 (2012).
48. S. Westenhoff *et al*., Time-resolved structural studies of protein reaction dynamics: A smorgasbord of X-ray approaches. *Acta Crystallogr. A* **66**, 207–219 (2010).
49. H. S. Cho *et al*., Protein structural dynamics in solution unveiled via 100-ps time-resolved x-ray scattering. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 7281–7286 (2010).
50. M. Cammarata *et al*., Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering. *Nat. Methods* **5**, 881–886 (2008).
51. M. Cammarata *et al*., Impulsive solvent heating probed by picosecond x-ray diffraction. *J. Chem. Phys.* **124**, 124504 (2006).
52. H. D. T. Mertens, D. I. Svergun, Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **172**, 128–141 (2010).
53. A. Nasedkin *et al*., Deconvoluting protein (un)folding structural ensembles using X-ray scattering, nuclear magnetic resonance spectroscopy and molecular dynamics simulation. *PLoS One* **10**, e0125662 (2015).
54. P. Bernadó, D. I. Svergun, Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Mol. Biosyst.* **8**, 151–167 (2012).
55. P. Bernadó, E. Mylonas, M. V. Petoukhov, M. Blackledge, D. I. Svergun, Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **129**, 5656–5664 (2007).
56. G. Jones, *Genetic and Evolutionary Algorithms*, (Wiley, Chichester, United Kingdom, 1998).
57. D. J. Segel, A. L. Fink, K. O. Hodgson, S. Doniach, Protein denaturation: A small-angle X-ray scattering study of the ensemble of unfolded states of cytochrome c. *Biochemistry* **37**, 12443–12451 (1998).
58. Y. Orii, Immediate reduction of cytochrome c by photoexcited NADH: Reaction mechanism as revealed by flow-flash and rapid-scan studies. *Biochemistry* **32**, 11910–11914 (1993).
59. K. Y. Oang, C. Yang, S. Muniyappan, J. Kim, H. Ihee, SVD-aided pseudo principal-component analysis: A new method to speed up and improve determination of the optimum kinetic model from time-resolved data. *Struct. Dyn.* **4**, 44013 (2017).
60. R. Kohlrausch, Theorie des elektrischen Rückstandes in der Leidener Flasche. *Ann. Phys.* **72**, 393–398 (1854).
61. A. T. N. Kumar, L. Zhu, J. F. Christian, A. A. Demidov, P. M. Champion, On the rate distribution analysis of kinetic data using the maximum entropy method: Applications to myoglobin relaxation on the nanosecond and femtosecond timescales. *J. Phys. Chem. B* **105**, 7847–7856 (2001).
62. M. Kataoka, Y. Hagihara, K. Mihara, Y. Goto, Molten globule of cytochrome c studied by small angle X-ray scattering. *J. Mol. Biol.* **229**, 591–596 (1993).
63. A. K. Churg, A. Warshel, Control of the redox potential of cytochrome c and microscopic dielectric effects in proteins. *Biochemistry* **25**, 1675–1681 (1986).
64. Z. Guo, D. Thirumalai, Kinetics of protein folding: Nucleation mechanism, time scales, and pathways. *Biopolymers* **36**, 83–102 (1996).
65. S. J. Hagen, J. Hofrichter, W. A. Eaton, Rate of intrachain diffusion of unfolded cytochrome c. *J. Phys. Chem. B* **101**, 2352–2365 (1997).
66. J. A. Ridge, R. L. Baldwin, A. M. Labhardt, Nature of the fast and slow refolding reactions of iron(III) cytochrome c. *Biochemistry* **20**, 1622–1630 (1981).
67. S. Piana, A. G. Donchev, P. Robustelli, D. E. Shaw, Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **119**, 5113–5123 (2015).
68. J. A. Riback *et al*., Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. *Science* **358**, 238–241 (2017).