THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL

Letter

Uncovering the Conformational Distribution of a Small Protein with Nanoparticle-Aided Cryo-Electron Microscopy Sampling

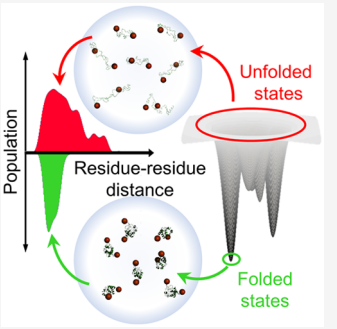
Hosung Ki, Junbeom Jo, Youngmin Kim, Tae Wu Kim, Changin Kim, Yeeun Kim, Chang Woo Kim, Srinivasan Muniyappan, Sang Jin Lee, Yonggwan Kim, Ho Min Kim, Yongsoo Yang, Young Min Rhee, and Hyotcherl Ihee*

Cite This: *J. Phys. Chem. Lett.* 2021, 12, 6565−6573

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Here, we introduce the nanoparticle-aided cryo-electron microscopy sampling (NACS) method to access the conformational distribution of a protein molecule. Two nanogold particles are labeled at two target sites, and the interparticle distance is measured as a structural parameter via cryo-electron microscopy (cryo-EM). The key aspect of NACS is that the projected distance information instead of the global conformational information is extracted from each protein molecule. This is possible because the contrast provided by the nanogold particles is strong enough to provide the projected distance, while the protein itself is invisible due to its low contrast. We successfully demonstrate that various protein conformations, even for small or disordered proteins, which generally cannot be accessed via cryo-EM, can be captured. The demonstrated method with the potential to directly observe the conformational distribution of such systems may open up new possibilities in studying their dynamics at a single-molecule level.



Proteins that function in unstructured states as well as in well-structured states play meaningful roles in various biological processes.[1−3] Due to the significance of the unstructured proteins, their structural characterizations have been actively carried out.[4−18] However, the intrinsic flexibility of such proteins makes it difficult to identify diverse instantaneous conformations, while such identification is essential to understanding the structural natures and the dynamic behaviors of the proteins. Generally, to characterize the structural propensities and the dynamics of these unstructured proteins, nuclear magnetic resonance spectroscopy has served as the most powerful tool.[10,19] However, this technique provides only short distance information and an ensemble-averaged structure of various instantaneous states, which is often insufficient for determining the conformational ensemble itself.

As an alternative in delineating the conformational distribution, a single-molecule experimental concept of single-object scattering sampling (SOSS) was proposed.[20−22] There, the scattering pattern from each object (a molecule or a complex) is measured by irradiation with a coherent, ultrashort, and ultraintense X-ray pulse. The scattering pattern contains the information regarding a structural parameter of the object in its instantaneous and specific conformation. Collecting a sufficient number of such patterns yields information about many conformations, which can then be used to extract a distribution of the structural parameter.

Here, as a step to further substantiate the concept of SOSS, especially its application to biological molecules such as proteins, we demonstrate an experimental method named nanoparticle-aided cryo-electron microscopy sampling (NACS), which can estimate various protein conformations through cryo-electron microscopy (cryo-EM). Cryo-EM collects images of individual particles fixed by vitrification, which are used to determine the three-dimensional structure.[23,24] Due to the low contrast of protein molecules in cryo-EM, small proteins generally are not visible in cryo-EM images. Even for large proteins that can be detected by cryo-EM, nanogold particles are sometimes used as a contrast agent as commonly employed in imaging biological systems.[25−28] This labeling strategy has in fact been extensively applied to studying the structural dynamics of DNA[29−31] and its reaction accompanying a reversible conformational change.[32] In these studies, nanogold particles were labeled at both ends of the double-stranded DNA (ds-DNA), and EM or electron tomography measurements[29,33] were performed to extract the interparticle distance and ultimately the length of ds-DNA. There, all ds-DNA molecules (10 nm[31,32] to 28 nm[29]) were substantially shorter than the persistence length of ds-DNA (∼40 nm).[34] Thus, the target molecules were much more rigid than the protein, and direct labeling with relatively large nanogold particles (5 nm[29] to 36 nm[31] in diameter) did not largely disturb DNA native structures. Here, in this study, we

Received: April 19, 2021
Accepted: July 6, 2021

https://doi.org/10.1021/acs.jpclett.1c01277
J. Phys. Chem. Lett. 2021, 12, 6565−6573

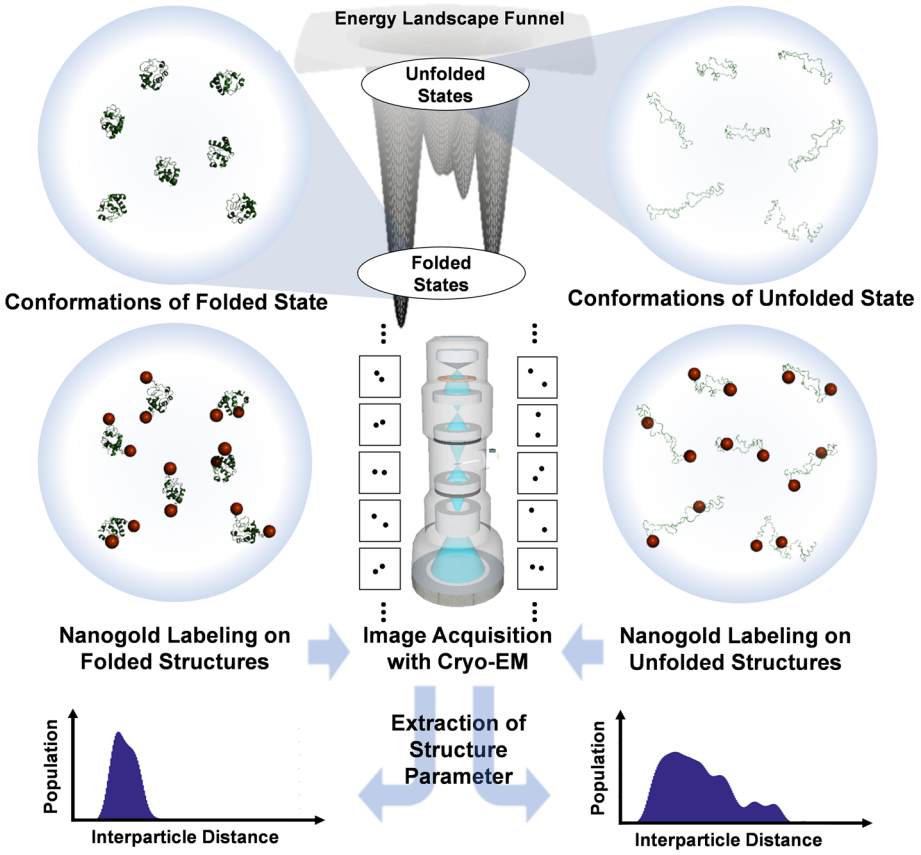


Figure 1. Schematic representation of the concept of NACS (nanoparticle-aided cryo-EM sampling) with two-point nanogold-labeled protein molecules. Both folded and unfolded states of the protein can have various instantaneous conformations. Nanogold labeling of two selected sites of the protein combined with the single-object imaging capability of cryo-EM allows the acquisition of micrographs with a clear contrast of nanogold particles. The distribution of various conformations of the nanogold-labeled protein can be constructed by measuring interparticle distances from obtained micrographs. The shape of the distribution depends on whether the protein is folded or unfolded.

apply the nanogold labeling strategy to a protein. With a protein, in contrast to DNA, it has been reported that the structure can be significantly altered by direct labeling of a single nanogold particle even at a relatively small size (1.5 nm in diameter) due to the electronic charges at the surface of the particle.[35] To avoid this problem, we adopt a nanogold particle that has a negligible surface charge thanks to organic surfactants on the surface. In addition, a particle having a linker that can be conjugated to a specific position of the protein is used to avoid direct contact of the protein with the surface of the particle. We first confirm that this strategy works properly in reducing the structural distortion of the protein. The linker in principle can be flexible, which can add a certain level of uncertainty to the distance between two labeled sites. Therefore, we take into account the influence of the flexible linker in our analysis.

As a model system for describing conformational distributions of both folded and unfolded states, we chose equine heart cytochrome *c* (cyt-c), which has been widely employed to investigate the structural aspect of the unfolded state as well as the native state, in relation to its function.[36−40] On the basis of the cryo-EM imaging of gold nanoparticle-labeled cyt-c, we succeeded in experimentally constructing its conformational distribution in the unfolded state. As such, this work presents the prospect of using nanogold labeling for protein structure characterizations as an alternative method of deducing conformational distributions of proteins.

For the NACS idea to work properly, the angular orientation of the nanogold-labeled protein molecules with respect to the direction of the incident electron beam needs to be determined. For example, a recent work with cryo-electron tomography[41] reported that approximately half of the tested proteins (12 of 28) showed an apparent preferred orientation, although there was no clear regularity with respect to predicting whether a certain protein will align along with a preferred orientation. Thus, unless the orientation of the cyt-c labeled with two nanogold particles is directly measured, one cannot predict whether they will be in a random or preferred orientation. We applied cryo-electron multitilt analysis to determine the actual distribution of the orientation of nanogold particles attached to proteins. Indeed, the measured orientation distribution turned out to be highly anisotropic instead of being random (Supporting Method S1 and Figure S1). Then, we used the measured orientational distribution to analyze the distribution of the projection distances, which contains the distribution of the actual three-dimensional distance between two labeled sites convoluted with the orientation distribution, from the cryo-EM images.

We performed experiments for both folded and unfolded states of cyt-c to compare their conformational distributions. The nanogold-labeled cyt-c was synthesized by the direct conjugation of the mutated cyt-c and two gold nanoparticles through the thiol group in cysteine residue, acting as the linker moiety. The details for the synthesis and characterization of

https://doi.org/10.1021/acs.jpclett.1c01277
J. Phys. Chem. Lett. 2021, 12, 6565−6573

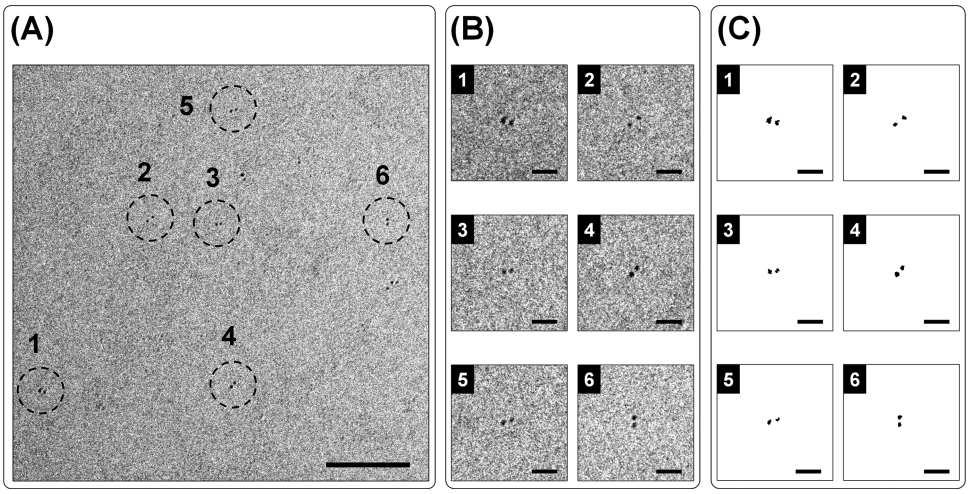


Figure 2. Cryo-EM images. (A) Representative cryo-EM images obtained from the nanogold-labeled G1C-G23C mutant of cyt-c. (B) Cropped regions (marked with circles in panel A) from panel A. (C) Background-removed images of panel B. Scale bars of 50 nm in panel A and 10 nm in panels B and C.

the nanogold-labeled cyt-c are described in Supporting Methods S2 and S3 and Figure S2. The experimental scheme is depicted in Figure 1. For both folded and unfolded states, because cyt-c consists of light atoms and is not sufficiently large for detection via cryo-EM, electron micrographs in general cannot provide its three-dimensional structures. Labeling nanogold particles on two desired sites, however, enables us to obtain images with clearly visible nanogold particles. Taking a sufficient number of images and then extracting interparticle distances from them allow us to construct a statistical distribution whose shape depends on the structural flexibility of cyt-c. By analyzing this distribution, we can gain insight into conformational heterogeneity.

As a proof of this concept, we first applied the scheme to the folded state. Three pairs of labeling sites with differing pair distances were selected to check whether the different yet already known distances can be reconstructed.[35,42] All three pairs of labeling sites were selected among the residues exposed on the surface of the protein so that the nanogold particles can readily access the sites to achieve proper labeling. In addition, we avoided the residues expected to play important roles in the structure and function of cyt-c, such as Gly6, Phe10, Cys14, Cys17, His18, Gly29, Pro30, Gly41, Asn52, Trp59, Tyr67, Leu68, Pro71, Pro76, Thr78, Met80, Phe82, Leu94, and Tyr97.[42] The chosen pairs are Gly1 and Gly23, Gly1 and Gly45, and Gly1 and Glu104, and they have C_α−C_α distances of 2.69, 3.07, and 1.79 nm, respectively, according to its structure (Protein Data Bank entry 1HRC).[43] As a labeling reagent, we chose commercially available 1.4 nm monomaleimido nanogold particles.[44] Because the maleimide functional group can covalently cross-link with the thiol group in cysteine, we can achieve labeling by mutating the target residues with cysteine, namely, as three mutants, G1C-G23C, G1C-G45C, and G1C-E104C.

For the mutants and wild-type cyt-c, circular dichroism (CD) spectra were recorded to check the integrity of the labeled protein structure, as shown in Figure S2. In light of secondary structures, the site-directed mutation did not affect the native conformation. The CD spectra after nanogold labeling were also recorded (Figure S2). The CD results indicate that the nanogold labeling did not cause any severe distortion from the native structure. Contrary to previous studies[35,45] that showed significant denaturation of cyt-c after labeling by negatively charged nanogold particles, we used particles with a neutral charge and labeled them through a maleimide linker. Our results confirm that this charge neutrality of the particle and the presence of the linker help maintain the native structure even with the nanogold conjugation.

Then, we measured cryo-EM images of the nanogold-labeled cyt-c. The details of the sample preparation and the measurement conditions are described in Supporting Method S4. A representative cryo-EM image is shown in Figure 2A. To conveniently post-process the images, we cropped the appearances of well-defined pairs from the acquired raw images so that they were at the centers of the cropped images (Figure 2B). The selection criteria for the well-defined particle pairs are described in Supporting Method S5. For each cropped image, edges between the particle and the background were determined (Figure 2C). Subsequently, the center-of-mass coordinates of both particles were determined for the calculation of the interparticle distance. The distance values were then collected to plot the distribution. All of the cropped images and corresponding processed images of nanogold-labeled cyt-c are presented in Figures S8−S11. The detailed procedures for the processing of cryo-EM images are described in Supporting Method S5.

As depicted in Figure 3A, the interparticle distances measured directly from TEM images are values of actual distances projected onto the plane perpendicular to the direction of the electron beam. Accordingly, we need to convert the measured distribution into the actual one. If the actual distance (d) and the angle (θ) between the sample and the electron beam are known, the projected distance (p) can be easily obtained by using $p = d\sin(\theta)$. In this study, however, both the actual distances and the corresponding angles are unknown parameters, and only the projected distances are measured experimentally. In fact, the three-dimensional position of the other particle with respect to one particle in a pair of particles can be expressed with three variables, d and two angles (θ and an additional azimuthal angle, φ), in a spherical coordinate system. Nevertheless, as the relation between p and d [i.e., $p = d\sin(\theta)$] shows, φ is irrelevant for the reconstruction of d from p, and thus φ will be ignored

https://doi.org/10.1021/acs.jpclett.1c01277
J. Phys. Chem. Lett. 2021, 12, 6565−6573

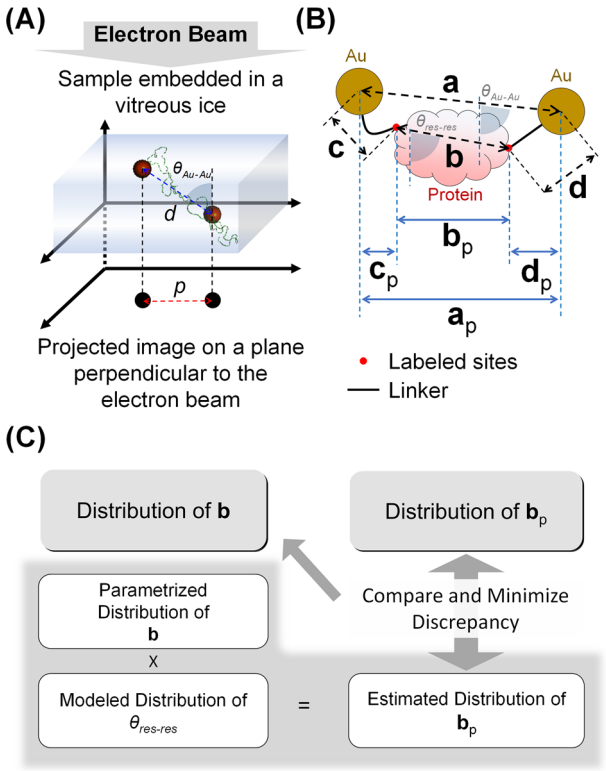


Figure 3. Schematic of data analysis. (A) Geometric relation among the direction of an electron beam, a nanogold-labeled protein molecule embedded in the vitreous ice, and the corresponding cryo-EM image acquired. The parameters d, θ_{Au-Au}, and p represent the distance between two particles, the angle between the axis connecting two particles and the direction of the electron beam, and the projected distance, respectively. (B) Effect of the linker length and the radius of the nanogold particle on the interparticle distance. **a** is the actual interparticle distance. $\mathbf{a}_p$ is the projected interparticle distance, which is the experimental observable. **b** is the actual intersite distance. $\mathbf{b}_p$ is the projected intersite distance. **c** and **d** each represent an extra distance, which is the sum of the linker length and the radius of a nanogold particle. $\mathbf{c}_p$ and $\mathbf{d}_p$ are each the sum of the projected linker length and the projected radius of the nanogold particle. Accordingly, we defined $\mathbf{c}_p + \mathbf{d}_p$ as the projected extra distance. Two angles, θ_{Au-Au} and the angle between the axis connecting two labeled sites and the direction of the electron beam, $\theta_{res-res}$, that influence the projection distances are also shown. Note that θ_{Au-Au} and $\theta_{res-res}$ are different. The protein, nanoparticles, and linkers shown in this figure reflect their actual relative sizes. (C) Scheme for reconstructing the actual distance distribution from the projected distance distribution. The distribution of **b** was described by a histogram consisting of some representative values (bins) of **b** and the population fraction of each bin of **b**. We assigned the population of each bin of **b** as a variable and used a proper assumption about the distributions of projection angles, $P(\theta_{Au-Au})$ and $P(\theta_{res-res})$. From the distribution of projection angles and **b**, the distribution of $\mathbf{b}_p$ can be theoretically calculated. This calculated distribution of $\mathbf{b}_p$ can be compared with the experimentally obtained distribution of $\mathbf{b}_p$. The variables describing **b** and $P(\theta_{res-res})$ are then optimized by minimizing the discrepancy between the experimental and calculated distributions of $\mathbf{b}_p$. A more detailed procedure is described throughout Supporting Methods S5–S8.

throughout this study. To reconstruct the actual distance distribution from the projected one, we used the following analysis scheme.

Before describing the reconstruction step, we should note that the experimentally obtained projected interparticle distance ($\mathbf{a}_p$ in Figure 3B, which is the projection of the actual interparticle distance, **a**) includes the projections of extra distances ($\mathbf{c}_p$ and $\mathbf{d}_p$ in Figure 3B, which are the projections of extra distances **c** and **d**, respectively), in addition to the projection of the intersite distance ($\mathbf{b}_p$ in Figure 3B, which is the projection of the intersite distance, **b**), which we are aiming to determine. **c** and **d** each represent an extra distance, which is the sum of the linker length and the radius of a nanogold particle. Accordingly, we defined the sum of the projections of the extra distances, $\mathbf{c}_p + \mathbf{d}_p$, as the projected extra distance. We assumed that the linker can be flexible within its maximum length (~1.3 nm)[44] and a flexible orientation with respect to both the nanogold particle and the protein (Figure S4D,F). Depending on the spatial orientation of the linker with respect to the protein molecule, $\mathbf{a}_p$ can deviate from $\mathbf{b}_p$ by the projected extra distance ($\mathbf{c}_p + \mathbf{d}_p$ in Figure 3B) in a range from −4 to 4 nm. To reconstruct the actual intersite distance (**b** in Figure 3B) distribution, $\mathbf{b}_p$ should first be extracted by removing $\mathbf{c}_p + \mathbf{d}_p$ from experimentally obtained $\mathbf{a}_p$. To estimate the distribution of $\mathbf{c}_p + \mathbf{d}_p$, for each labeled site, we performed simulations as detailed in Supporting Method S6. The estimated distributions of $\mathbf{c}_p + \mathbf{d}_p$ for all mutants are shown in Figure S4G–I. By using the distribution, we subtracted $\mathbf{c}_p + \mathbf{d}_p$ from the experimentally measured $\mathbf{a}_p$ to obtain the distribution of $\mathbf{b}_p$ as detailed in Supporting Method S7.

After obtaining the distribution of $\mathbf{b}_p$, we reconstructed the distribution of the actual intersite distance, **b**, according to the analytic scheme based on a previously reported method[31,32] as illustrated in Figure 3C. In short, the population fraction at each bin of the distribution of **b** was set as an independent parameter. By taking into account the orientations of the samples with respect to the electron beam, which are strongly anisotropic (Figure S1), we can obtain the parametric equation of the population fraction at each bin of the distribution of $\mathbf{b}_p$ (eqs S3 and S4). The parameters were iteratively fit by minimizing the difference between the calculated distribution of $\mathbf{b}_p$ using the parametric equation and the experimental distribution of $\mathbf{b}_p$. More information about the analysis, including the equations, is presented in Supporting Method S8.

The distributions of $\mathbf{a}_p$s for the three mutants are plotted in Figure 4A–C. A notable point in the distributions of $\mathbf{a}_p$ is that the population fraction is zero for the short $\mathbf{a}_p$ distances ($\mathbf{a}_p$ < 2 nm). This observation may seem unnatural at first glance, but it is in fact consistent with the result of the multitilt measurement that the orientation of the particle pairs has a highly anisotropic distribution instead of a random distribution (Supporting Method S1 and Figure S1). According to the result, the vector connecting two nanoparticles is most likely perpendicular to the incident electron beam. More specifically, the probabilities of observing θ_{Au-Au} angles of < 30°, < 45°, and < 60° are 0.0001%, 0.06%, and 2.7%, respectively. Therefore, the probability that $\mathbf{a}_p$ has a short distance of < 2 nm becomes negligible. The distribution of $\mathbf{b}_p$s after subtracting the projected extra distances is presented in Figure 4E–G. The distributions of **b**s satisfying these $\mathbf{b}_p$ distributions were reconstructed and are shown in Figure 4I–K. Red circles with crosses in Figure 4E–G indicate fit population fractions to produce the distributions of **b** in Figure 4I–K. In these reconstructed distributions, the most probable distance values are ~2.3 nm for G1C-G23C, ~3.3 nm for G1C-G45C, and ~1.9 nm for G1C-E104C. These values are in good agreement with the estimated distances between the sulfur atoms of the mutated positions, namely, 2.83, 3.35, and 2.02 nm, respectively.

https://doi.org/10.1021/acs.jpclett.1c01277
J. Phys. Chem. Lett. 2021, 12, 6565−6573

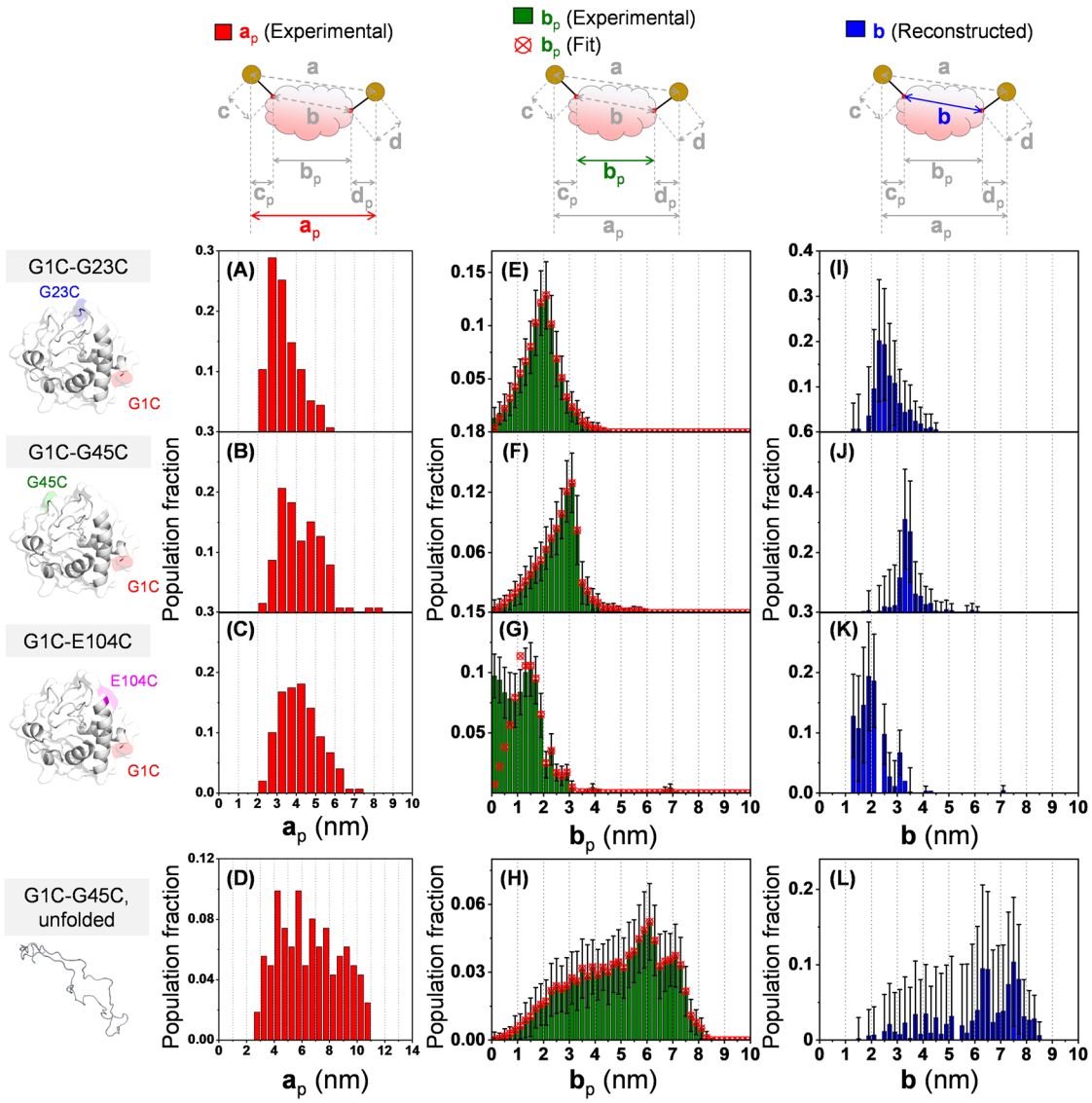


Figure 4. Experimental data and results of the analysis of the folded and unfolded states. Experimental distributions of $\mathbf{a}_p$ obtained from the nanogold-labeled (A) G1C-G23C, (B) G1C-G45C, and (C) G1C-E104C mutants of cyt-c in the folded state and (D) the G1C-G45C mutant of cyt-c in the unfolded state. The locations of labeled sites are represented on the left side of each panel. (E−H) Distributions of $\mathbf{b}_p$ obtained by subtraction of the projected extra distances, $\mathbf{c}_p + \mathbf{d}_p$, from $\mathbf{a}_p$. Error bars represent standard deviations of the population at each bin. Red circles with crosses are fit population fractions for reconstructing the distribution of $\mathbf{b}$ shown in panels I−L. (I−L) Reconstructed distributions of $\mathbf{b}$ for each mutant and state. For the reconstruction, the distribution of one of the projection angles, $P(\theta_{Au-Au})$, was determined from a separate experiment (see Figure S1). The distribution of the other projection angle, $P(\theta_{res-res})$, was modeled and optimized together with the reconstruction of the distribution of $\mathbf{b}$ (see Figure S6 and Supporting Method S8 for details). Error bars represent errors of the reconstruction that were estimated using the MINUIT software package with the MINOS error estimation method.

Now, let us move on to investigate the conformational distribution of the unfolded state. In the case of the unfolded protein, the conformational distribution is difficult to characterize, and direct measurement has rarely been reported. We constructed its conformational distribution with the same strategy taken with the folded state. To generate the unfolded state, guanidine hydrochloride (GdnHCl) was added to the solutions of labeled mutants, with a final GdnHCl concentration of 4 M,[46] and the cryo-EM experiments of the unfolded, labeled mutants were performed in the same manner as for the folded proteins. For the unfolded G1C-G45C, the $\mathbf{a}_p$ distribution was obtained by analyzing 162 images as displayed in Figure 4D, together with its post-processed $\mathbf{b}_p$ in Figure 4H and $\mathbf{b}$ in Figure 4L (cryo-EM images can be found in Figure S11). Here, the same distribution of $\mathbf{c}_p + \mathbf{d}_p$ as in the folded state of the G1C-G45C mutant was assumed to generate the distribution of $\mathbf{b}_p$. As expected, compared to the folded state cases shown in Figure 4I−K, $\mathbf{b}$ in the unfolded state exhibits a broader distribution, implying that the unfolded state has a more flexible and extended structure than the structure in the folded state. In addition, we can observe a group of small populations at relatively short distances (at $\mathbf{b}$ < 5.5 nm) and two distinct peaks at ∼6.3 and ∼7.5 nm rather than a monotonous distribution. This may indicate that the unfolded state may be composed of three conformational groups with some residual regular structural motifs[47] rather than being in just a single, completely disordered state, i.e., a randomly coiled state. In our analysis, ∼27% of the population is within a

https://doi.org/10.1021/acs.jpclett.1c01277
J. Phys. Chem. Lett. 2021, 12, 6565−6573

b of 5.5 nm, ~34% of the population is at a **b** of around 6.3 nm, and the remaining ~39% is at a **b** of around 7.5 nm.

For comparison, we performed molecular dynamics (MD) simulations to obtain structural ensembles of the folded and unfolded states. The conformations of the folded ensemble were obtained from trajectories with explicit solvation, while the unfolded configurations were sampled from high-temperature simulations with implicit solvation accompanied by a relatively small solvent viscosity. Other details for the MD simulations are described in Supporting Method S9. Figure 5 compares the distributions of **b** derived from both the experiment and MD simulations.

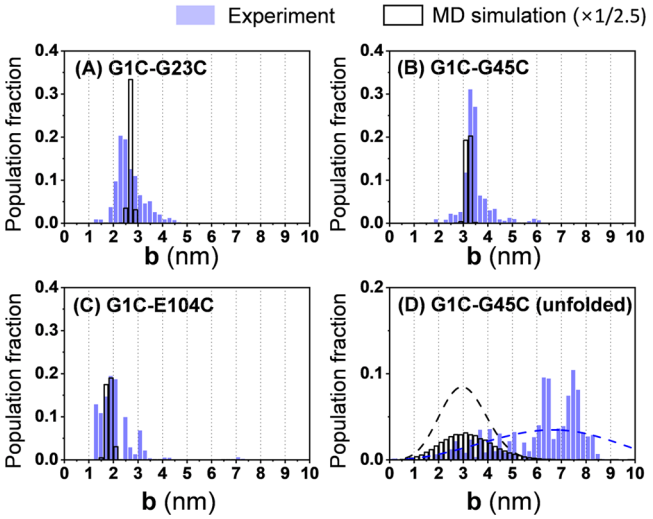


Figure 5. Comparison of the intersite distance distributions. Distance distributions obtained from experiments and MD simulations for (A) G1C-G23C in the folded state, (B) G1C-G45C in the folded state, (C) G1C-E104C in the folded state, and (D) G1C-G45C in the unfolded state. Experimental results are shown as blue filled bars, and MD simulation results are shown as black empty bars. The population fraction (*y*-axis in the plots) of each bin from the MD simulation is scaled down by a factor of 2.5 for the plots to compare the experimental and theoretical distributions on a similar *y*-axis scale. The dashed curves in panel D are the distance distributions estimated by using two representative models (black for the random walk model,[48] with an excluded volume, and blue for the wormlike chain model[49]) of random coil polymers. These two models produce distributions that fit well with the MD simulation and the NACS experimental results, respectively.

The comparison of **b** from the experiment and that from the MD simulation provides a clear-cut demonstration to validate the structural sensitivity of NACS and the relevance of the theoretical approach for elucidating the structural parameter of unfolded cyt-c. For the folded state, the distributions display decent to good agreements, as shown in Figure 5A–C. For the unfolded state, the agreement is not as good. Major peaks on the distribution beyond 5 nm from the experiment are not visible with the simulated data, however, hinting that details in the structural ensembles are different. To some extent, the discrepancy is not surprising as the force field parameters we have adopted were not really developed for describing unfolded states. In this sense, we hope that the experimental distribution can serve as a benchmark data set for the development of force fields suitable for unfolded proteins in the future.

As a baseline for comparison with the distance distribution of the unfolded state, we also estimated the distributions by assuming that the unfolded protein can be regarded as a random coil polymer. Specifically, we considered two models: (i) a random walk model[48] with a completely flexible chain but with an excluded volume for avoiding intrachain steric clashes and (ii) a wormlike chain model[49] in which the chain is not completely flexible but has a degree of stiffness due to the residue–residue interaction. We confirmed that the distribution from the first random walk model (Figure 5D, black dashed line) agrees well with the distribution from MD simulations when the excluded volume parameter was set to $4\pi/3 \times (1.3\ \text{Å})^3$. Because the residue–residue interactions in MD simulations at sufficiently high temperatures become negligible in comparison with the thermal energy, this agreement is indeed quite understandable. Interestingly, this random walk model-based distribution could not fit with the experimental data. Indeed, the experimental data from the NACS method could fit much better with the second one, namely the wormlike chain model (Figure 5D, blue dashed line), when the persistence length was set to 22.8 Å. The fact that the experimental distribution is more consistent with the second model signifies that the residue–residue interaction in the unfolded state should not be ignored. The existence of distinct peaks in the experimental distribution also suggests that the unfolded cyt-c is not a completely random coil but retains some residual structures. More details regarding the estimation of the distance distribution using the random coil polymer models are described in Supporting Method S10.

One important aspect that we have to note here with the data about the large separation regime is the fact that the unfolded state can easily extend with 7 nm or even 8 nm of the intersite distances, and cryo-EM-based measurements can readily detect such conformations on a single-molecule level, unlike other conventional experimental techniques. For example, in earlier FRET studies of the unfolded state of cyt-c,[37,38] six different positions (Lys4, His39, Asp50, Glu66, Leu85, and Lys99) were labeled with a fluorescent donor and the heme group was regarded as an acceptor, and the donor–acceptor (D–A) distance was regarded as the structural parameter. However, in the case of the heme–Leu85 and –Lys99 distances, most populations could not be detected via FRET and were judged to be beyond 5.9 nm. While these studies could provide valuable structural information about the unfolded state, it is quite limited within the detection range of FRET. In comparison, our method could determine the whole conformational distribution even beyond 6 nm. In addition, the distance information from every single object is utilized to construct the distribution, whereas with FRET, judging the size of the fraction of the undetected population is not straightforward. Meanwhile, from the viewpoint of conformational heterogeneity, our result implicates a coexistence of extended structures and relatively compact structures. Even though the two nanogold particles labeled as the probe can alter the nature of the unfolded conformations, we could still extract the structural information about the unfolded state as well as the folded state.

We believe that the issue about the perturbation induced by the attached nanoparticles can be analyzed in a clearer way by further investigating the influences of the nanogold labeling from both theoretical and experimental perspectives. In this work, we alleviated the ambiguity associated with the orientation of the sample, especially for a projection angle,

https://doi.org/10.1021/acs.jpclett.1c01277
J. Phys. Chem. Lett. 2021, 12, 6565–6573

θ_{Au-Au}, by using the multitilt analysis and thereby reduced the overall inherent uncertainty of the final distribution of **b**. Nevertheless, it should be noted that some uncertainty in the distribution of θ_{Au-Au} still exists because we applied the cryo-electron multitilt analysis for only some of the G1C-G45C mutants in the folded state. It is possible that the degree of the preferred orientation is different for each mutant and a mutant in the folded state and the unfolded state. Furthermore, the degree of the preferred orientation might differ for each grid, and in the worst case, it might vary depending on the position within the same grid. Considering this, the reconstructed distributions of **b** presented herein would possess a certain degree of inaccuracies originating from the uncertainty in the distribution of θ_{Au-Au}. Ideally, determining the distribution of θ_{Au-Au} by performing the multitilt analysis on all of the measured samples is desirable, but we did not take such a rigorous approach in this work but rather focused on demonstrating a new method. The rather long linker and its conformational and orientational flexibility with respect to the nanogold particle and the protein may still pose some ambiguity. In particular, this problem becomes remarkable in the case of the G1C-E104C mutant of cyt-c where nanogold particles are labeled on two nearby residues. One can see that, for the G1C-E104C mutant, our method failed to properly reproduce the experimental distribution of $\mathbf{b}_p$ at short (< 1.5 nm) distances, probably because the distribution of $\mathbf{c}_p + \mathbf{d}_p$ was not precisely described (Figure 4G). In contrast, the experimental distribution of $\mathbf{b}_p$ can be excellently reproduced for the two other mutants where nanogold particles are labeled on two distant residues. Considering this, upon application of NACS methods, alleviating the problem of the linker flexibility by selecting labeling sites as far apart as possible is desirable.

We have demonstrated that the nanogold labeling scheme can be adapted to access the intersite distance information about various conformational states of a protein. Using nanogold particles with an organic linker group, we labeled three different pairs of residue sites located on the surface of cyt-c in its folded state and characterized the pair distances with cryo-EM. Also, by adding GdnHCl as a denaturant to the labeled protein solution, we could generate the unfolded state with a minimal structural influence by the nanogold labels. By processing the cryo-EM images of the nanogold labels, we analyzed the conformational characters of the folded and unfolded states at a molecule-specific level. Quite expectedly, the distance distribution extracted from the unfolded state is wider than the one from the folded state, representing its intrinsic flexibility. Also, the conformational distribution exhibits multiple peaks, suggesting the coexistence of a few structurally distinct sub-ensembles. Considering the experimental difficulties associated with the characterization of unstructured states of proteins, the proposed experimental scheme has the potential to become a useful tool for investigating conformational distributions of unstructured proteins in general and elucidating their dynamic behaviors involving their various conformations, potentially under interactions with their binding partners.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.1c01277.

Supporting methods for the determination of the distribution of projection angle θ_{Au-Au}, preparation of the protein sample, nanogold labeling of cyt-c, acquisition of cryo-EM images, processing of cryo-EM images, construction of the projected extra distance distribution, subtraction of the extra distance from the experimental data, reconstruction of the actual intersite distance distribution from the projected intersite distance distribution, calculation of the intersite distance distribution from the MD simulations, and estimation of the intersite distance distribution by using random coil polymer models, a table of the sequences of the oligonucleotides used to construct cyt-c mutants, and figures showing the distribution of θ_{Au-Au}, the angle between the axis connecting two particles and the direction of the electron beam, results of CD measurements and size-exclusion chromatography, representative raw cryo-EM images for nanogold-labeled G1C-G45C cyt-c, construction of the projected extra distance distribution, estimation of the range of the projected extra distance values and their correlation with the projected interparticle distance values, comparison of the distribution of the actual intersite distance, **b**, reconstructed under two different assumptions of the distribution of projection angle, $P(\theta_{res-res})$, and the corresponding best-fit distribution of the projected intersite distance, $\mathbf{b}_p$, to the experimentally derived distribution of $\mathbf{b}_p$, effect of bubbles in the sample solution containing guanidine hydrochloride (GdnHCl), and cryo-EM images obtained from nanogold-labeled cyt-c (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Hyotcherl Ihee − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; KI for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea;* orcid.org/0000-0003-0397-5965; Phone: +82-42-350-2844; Email: hyotcherl.ihee@kaist.ac.kr

Authors

Hosung Ki − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; KI for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea;* orcid.org/0000-0001-8273-6596

Junbeom Jo − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; KI for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea;* Present Address: LED Business, Samsung Electronics, Gyeonggi-do 18448, Republic of Korea

https://doi.org/10.1021/acs.jpclett.1c01277
J. Phys. Chem. Lett. 2021, 12, 6565−6573

Youngmin Kim − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; KI for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea;* orcid.org/0000-0001-8538-8435

Tae Wu Kim − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; KI for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea;* Present Address: Department of Chemistry, Mokpo National University, Muan-gun, Jeollanam-do, Republic of Korea; orcid.org/0000-0002-6370-0907

Changin Kim − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; KI for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea*

Yeeun Kim − *Department of Physics, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea*

Chang Woo Kim − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea*

Srinivasan Muniyappan − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; KI for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea*

Sang Jin Lee − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; KI for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea*

Yonggwan Kim − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; KI for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Center for Nanomaterials and Chemical Reactions, Institute for Basic Science (IBS), Daejeon 34141, Republic of Korea*

Ho Min Kim − *Center for Biomolecular & Cellular Structure, Institute for Basic Science (IBS), Daejeon 34126, Republic of Korea; Graduate School of Medical Science & Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea*

Yongsoo Yang − *Department of Physics, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0001-8654-302X

Young Min Rhee − *Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-2392-3962

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.1c01277

Author Contributions

H.K., J.J., Youngmin Kim, and T.W.K. contributed equally to this work. H.K., J.J., and H.I. conceived the project. H.K., J.J., Youngmin Kim, S.M., S.J.L., and Yonggwan Kim prepared protein samples. J.J. carried out cryo-EM measurements. C.K., Yeeun Kim, and Y.Y. performed the cryo-electron multitilt measurements and analysis. H.K., J.J., T.W.K., and H.I. analyzed the data. T.W.K., C.W.K., and Y.M.R. performed MD simulations. H.K., J.J., T.W.K., Y.M.R., and H.I. wrote the manuscript, and all authors commented on the manuscript. H.I. supervised the whole project.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Some of the TEM measurements were conducted using a Thermo Fisher Glacios instrument at the KAIST Analysis Center for Research Advancement (KARA). H.K., J.J., Youngmin Kim, T.W.K., C.K., S.M., S.J.L., Yonggwan Kim, and H.I. were supported by the Institute for Basic Science (IBS-R004). Yeeun Kim and Y.Y. were supported by the Samsung Science and Technology Foundation (SSTF-BA1802−05). C.W.K. and Y.M.R. were supported by the National Research Foundation (Creative Materials Discovery Program, Grant 2018M3D1A1058813), funded by the Ministry of Science and ICT of Korea.

■ REFERENCES

(1) Tompa, P. The Interplay between Structure and Function in Intrinsically Unstructured Proteins. *FEBS Lett.* **2005**, *579*, 3346−3354.

(2) Dyson, H. J.; Wright, P. E. Intrinsically Unstructured Proteins and their Functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197−208.

(3) Wright, P. E.; Dyson, H. J. Intrinsically Unstructured Proteins: Re-Assessing the Protein Structure-Function Paradigm. *J. Mol. Biol.* **1999**, *293*, 321−331.

(4) van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M. Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **2014**, *114*, 6589−6631.

(5) Jakob, U.; Kriwacki, R.; Uversky, V. N. Conditionally and Transiently Disordered Proteins: Awakening Cryptic Disorder to Regulate Protein Function. *Chem. Rev.* **2014**, *114*, 6779−6805.

(6) Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V. N. Introducing Protein Intrinsic Disorder. *Chem. Rev.* **2014**, *114*, 6561−6588.

(7) Gsponer, J.; Futschik, M. E.; Teichmann, S. A.; Babu, M. M. Tight Regulation of Unstructured Proteins: From Transcript Synthesis to Protein Degradation. *Science* **2008**, *322*, 1365−1368.

(8) Liu, J. G.; Perumal, N. B.; Oldfield, C. J.; Su, E. W.; Uversky, V. N.; Dunker, A. K. Intrinsic Disorder in Transcription Factors. *Biochemistry* **2006**, *45*, 6873−6888.

(9) Iakoucheva, L. M.; Brown, C. J.; Lawson, J. D.; Obradovic, Z.; Dunker, A. K. Intrinsic Disorder in Cell-Signaling and Cancer-associated Proteins. *J. Mol. Biol.* **2002**, *323*, 573−584.

(10) Jensen, M. R.; Zweckstetter, M.; Huang, J. R.; Blackledge, M. Exploring Free-energy Landscapes of Intrinsically Disordered Proteins at Atomic Resolution using NMR Spectroscopy. *Chem. Rev.* **2014**, *114*, 6632−6660.

(11) Eliezer, D. Biophysical Characterization of Intrinsically Disordered Proteins. *Curr. Opin. Struct. Biol.* **2009**, *19*, 23−30.
(12) Roy, R.; Hohng, S.; Ha, T. A Practical Guide to Single-Molecule FRET. *Nat. Methods* **2008**, *5*, 507−516.
(13) Muller, D. J.; Dufrene, Y. F. Atomic Force Microscopy as a Multifunctional Molecular Toolbox in Nanobiotechnology. *Nat. Nanotechnol.* **2008**, *3*, 261−269.
(14) Hofmann, H. Understanding Disordered and Unfolded Proteins using Single-Molecule FRET and Polymer Theory. *Methods Appl. Fluoresc.* **2016**, *4*, 042003.
(15) Schuler, B.; Eaton, W. A. Protein Folding Studied by Single-Molecule FRET. *Curr. Opin. Struct. Biol.* **2008**, *18*, 16−26.
(16) Rounsevell, R.; Forman, J. R.; Clarke, J. Atomic Force Microscopy: Mechanical Unfolding of Proteins. *Methods* **2004**, *34*, 100−111.
(17) Bornschlogl, T.; Rief, M. Single-Molecule Protein Unfolding and Refolding using Atomic Force Microscopy. *Methods Mol. Biol.* **2011**, *783*, 233−250.
(18) He, Y. F.; Lu, M. L.; Cao, J.; Lu, H. P. Manipulating Protein Conformations by Single-Molecule AFM-FRET Nanoscopy. *ACS Nano* **2012**, *6*, 1221−1229.
(19) Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S. B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L. J.; Dobson, C. M.; Schwalbe, H. Long-Range Interactions within a Nonnative Protein. *Science* **2002**, *295*, 1719−1722.
(20) Ki, H.; Kim, K. H.; Kim, J.; Lee, J. H.; Kim, J.; Ihee, H. Prospect of Retrieving Vibrational Wave Function by Single-Object Scattering Sampling. *J. Phys. Chem. Lett.* **2013**, *4*, 3345−3350.
(21) Ihee, H. Novel Single-Molecule Technique by Single-Object Scattering Sampling (SOSS). *Bull. Korean Chem. Soc.* **2011**, *32*, 1849−1850.
(22) Kim, J.; Kim, K. H.; Lee, J. H.; Ihee, H. Ultrafast X-ray Diffraction in Liquid, Solution and Gas: Present Status and Future Prospects. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2010**, *66*, 270−280.
(23) Fernandez-Leiro, R.; Scheres, S. H. Unravelling Biological Macromolecules with Cryo-Electron Microscopy. *Nature* **2016**, *537*, 339−346.
(24) Cheng, Y. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell* **2015**, *161*, 450−457.
(25) Ackerson, C. J.; Powell, R. D.; Hainfeld, J. F. Site-Specific Biomolecule Labeling with Gold Clusters. *Methods Enzymol.* **2010**, *481*, 195−230.
(26) Sperling, R. A.; Rivera Gil, P.; Zhang, F.; Zanella, M.; Parak, W. J. Biological Applications of Gold Nanoparticles. *Chem. Soc. Rev.* **2008**, *37*, 1896−1908.
(27) Hainfeld, J. F.; Powell, R. D. New Frontiers in Gold Labeling. *J. Histochem. Cytochem.* **2000**, *48*, 471−480.
(28) Jahn, W. Review: Chemical Aspects of the Use of Gold Clusters in Structural Biology. *J. Struct. Biol.* **1999**, *127*, 106−112.
(29) Zhang, L.; Lei, D.; Smith, J. M.; Zhang, M.; Tong, H.; Zhang, X.; Lu, Z.; Liu, J.; Alivisatos, A. P.; Ren, G. Three-Dimensional Structural Dynamics and Fluctuations of DNA-nanogold Conjugates by Individual-Particle Electron Tomography. *Nat. Commun.* **2016**, *7*, 11083.
(30) Chen, Q.; Smith, J. M.; Park, J.; Kim, K.; Ho, D.; Rasool, H. I.; Zettl, A.; Alivisatos, A. P. 3D Motion of DNA-Au Nanoconjugates in Graphene Liquid Cell Electron Microscopy. *Nano Lett.* **2013**, *13*, 4556−4561.
(31) Busson, M. P.; Rolly, B.; Stout, B.; Bonod, N.; Larquet, E.; Polman, A.; Bidault, S. Optical and Topological Characterization of Gold Nanoparticle Dimers Linked by a Single DNA Double Strand. *Nano Lett.* **2011**, *11*, 5060−5065.
(32) Lermusiaux, L.; Sereda, A.; Portier, B.; Larquet, E.; Bidault, S. Reversible Switching of the Interparticle Distance in DNA-Templated Gold Nanoparticle Dimers. *ACS Nano* **2012**, *6*, 10992−10998.
(33) Zhang, X.; Zhang, L.; Tong, H. M.; Peng, B.; Rames, M. J.; Zhang, S. L.; Ren, G. 3D Structural Fluctuation of IgG1 Antibody Revealed by Individual Particle Electron Tomography. *Sci. Rep.* **2015**, *5*, 9803.
(34) Gross, P.; Laurens, N.; Oddershede, L. B.; Bockelmann, U.; Peterman, E. J. G.; Wuite, G. J. L. Quantifying How DNA Stretches, Melts and Changes Twist under Tension. *Nat. Phys.* **2011**, *7*, 731−736.
(35) Aubin-Tam, M. E.; Hwang, W.; Hamad-Schifferli, K. Site-Directed Nanoparticle Labeling of Cytochrome c. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 4095−4100.
(36) Pletneva, E. V.; Gray, H. B.; Winkler, J. R. Nature of the Cytochrome c Molten Globule. *J. Am. Chem. Soc.* **2005**, *127*, 15370−15371.
(37) Pletneva, E. V.; Gray, H. B.; Winkler, J. R. Snapshots of Cytochrome c Folding. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 18397−18402.
(38) Pletneva, E. V.; Gray, H. B.; Winkler, J. R. Many Faces of the Unfolded State: Conformational Heterogeneity in Denatured Yeast Cytochrome c. *J. Mol. Biol.* **2005**, *345*, 855−867.
(39) Ritter, A. R.; Shi, Y.; Fisher, A.; Ferretti, J. A.; Taniuchi, H. On the Nature of the Partially Unfolded State of Yeast Iso-2-Cytochrome c Proteins. *Biophys. J.* **1996**, *70*, Tu231.
(40) Kim, T. W.; Lee, S. J.; Jo, J.; Kim, J. G.; Ki, H.; Kim, C. W.; Cho, K. H.; Choi, J.; Lee, J. H.; Wulff, M.; Rhee, Y. M.; Ihee, H. Protein Folding From Heterogeneous Unfolded State Revealed by Time-Resolved X-ray Solution Scattering. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 14996.
(41) Noble, A. J.; Dandey, V. P.; Wei, H.; Brasch, J.; Chase, J.; Acharya, P.; Tan, Y. Z.; Zhang, Z. N.; Kim, L. Y.; Scapin, G.; Rapp, M.; Eng, E. T.; Rice, W. J.; Cheng, A. C.; Negro, C. J.; Shapiro, L.; Kwong, P. D.; Jeruzalmi, D.; des Georges, A.; Potter, C. S.; Carragher, B. Routine Single Particle CryoEM Sample and Grid Characterization by Tomography. *eLife* **2018**, *7*, e34257.
(42) Zaidi, S.; Hassan, M. I.; Islam, A.; Ahmad, F. The Role of Key Residues in Structure, Function, and Stability of Cytochrome-c. *Cell. Mol. Life Sci.* **2014**, *71*, 229−255.
(43) Bushnell, G. W.; Louie, G. V.; Brayer, G. D. High-Resolution Three-Dimensional Structure of Horse Heart Cytochrome c. *J. Mol. Biol.* **1990**, *214*, 585−595.
(44) Hainfeld, J. F.; Furuya, F. R. A 1.4-nm Gold Cluster Covalently Attached to Antibodies Improves Immunolabeling. *J. Histochem. Cytochem.* **1992**, *40*, 177−184.
(45) Aubin-Tam, M. E.; Hamad-Schifferli, K. Gold Nanoparticle-Cytochrome c Complexes: The Effect of Nanoparticle Ligand Charge on Protein Structure. *Langmuir* **2005**, *21*, 12080−12084.
(46) Nall, B. T.; Landers, T. A. Guanidine Hydrochloride Induced Unfolding of Yeast Iso-2 Cytochrome c. *Biochemistry* **1981**, *20*, 5403−5411.
(47) Zagrovic, B.; Pande, V. S. Structural Correspondence between the Alpha-Helix and the Random-Flight Chain Resolves How Unfolded Proteins Can Have Native-Like Properties. *Nat. Struct. Mol. Biol.* **2003**, *10*, 955−961.
(48) Edwards, S. F. The Statistical Mechanics of Polymers with Excluded Volume. *Proc. Phys. Soc., London* **1965**, *85*, 613−624.
(49) Thirumalai, D.; Ha, B.-Y. Statistical Mechanics of Semiflexible Chains: A Mean-Field Variational Approach. In *Theoretical and Mathematical Models in Polymer Research*; Grosberg, A., Ed.; Academic Press: San Diego, 1998; pp 1−35.

https://doi.org/10.1021/acs.jpclett.1c01277
J. Phys. Chem. Lett. 2021, 12, 6565−6573