BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

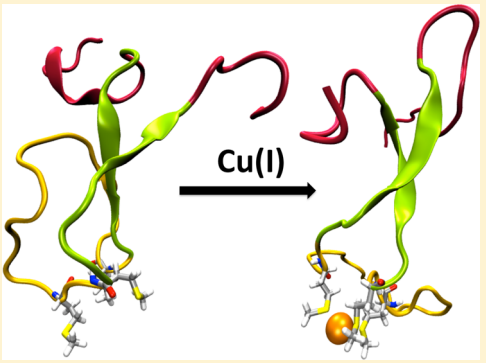
Structure and Dynamics of the N-Terminal Domain of the Cu(I) Binding Protein CusB

Melek N. Ucisik, Dhruva K. Chakravorty,[†] and Kenneth M. Merz, Jr.*[,‡]

Department of Chemistry and Quantum Theory Project, University of Florida, 2328 New Physics Building, P.O. Box 118435, Gainesville, Florida 32611-8435, United States

Supporting Information

ABSTRACT: CusCFBA is one of the metal efflux systems in *Escherichia coli* that is highly specific for its substrates, Cu(I) and Ag(I). It serves to protect the bacteria in environments that have lethal concentrations of these metals. The membrane fusion protein CusB is the periplasmic piece of CusCFBA, which has not been fully characterized by crystallography because of its extremely disordered N-terminal region. This region has both structural and functional importance because it has been experimentally proven to transfer the metal by itself from the metallochaperone CusF and to induce a structural change in the rest of CusB to increase Cu(I)/Ag(I) resistance. Understanding metal uptake from the periplasm is critical to gain insight into the mechanism of the whole CusCFBA pump, which makes resolving a structure for the N-terminal region necessary because it contains the metal binding site. We ran extensive molecular dynamics simulations to reveal the structural and dynamic properties of both the apo and Cu(I)-bound versions of the CusB N-terminal region. In contrast to its functional companion CusF, Cu(I) binding to the N-terminus of CusB causes only a slight, local stabilization around the metal site. The trajectories were analyzed in detail, revealing extensive structural disorder in both the apo and holo forms of the protein. CusB was further analyzed by breaking the protein up into three subdomains according to the extent of the observed disorder: the N- and C-terminal tails, the central beta strand motif, and the M21−M36 loop connecting the two metal-coordinating methionine residues. Most of the observed disorder was traced back to the tail regions, leading us to hypothesize that the latter two subdomains (residues 13−45) may form a functionally competent metal-binding domain because the tail regions appear to play no role in metal binding.



In most organisms, copper-binding enzymes are usually membrane bound or found in the periplasm.[1] *Escherichia coli* is one such organism. In *E. coli*, the periplasm can contain significantly higher copper concentrations than the cytoplasm because copper is needed for periplasmic or inner-membrane enzymes such as multicopper oxidases, Cu,Zn-superoxide dismutases, and cytochrome c oxidase, the final enzyme of the respiratory chain.[2] Still, periplasmic copper concentrations must be kept at low levels because Cu(I) can initiate the generation of reactive oxygen species (ROS) under aerobic conditions.[2] This delicate balance is achieved through various metal-resistance systems that function against excess metal concentrations including the RND-type (resistance, nodulation, division) transporters, which are common defense mechanisms not only in *E. coli* but also in all Gram-negative bacteria. Thus, they play major roles in the intrinsic and acquired antibiotic resistance of Gram-negative bacteria by facilitating their survival under otherwise lethal concentrations of drugs and metal ions.[3] They also aid in expelling bacterial products such as siderophores, peptides, and quorum-sensing signals.[4,5] Considering that antibiotic-resistant pathogens represent a growing threat to human health, clearer insight into these efflux mechanisms is of significant importance.

RND-type efflux systems consist of three fundamental components: an energy-requiring inner-membrane protein,[6] an outer-membrane factor, and a periplasmic component.[7] Proton-substrate antiporters of the RND protein superfamily serve as the inner-membrane components, and their exported substrate determines the subclass to which they belong. Accordingly, the inner-membrane proteins ejecting heavy metals make up the heavy-metal efflux subfamily, and they are highly substrate-specific with the ability to differentiate even between the charge of the ions.[6] The cytoplasm or periplasm provides the substrates, depending on the properties of the particular efflux system and the substrate.[8]

The CusCFBA efflux system in *E. coli* is responsible for extrusion of Cu(I) and Ag(I) and consists of CusA, the inner-membrane proton/substrate antiporter of the heavy-metal efflux RND family, CusB, the periplasmic protein, and CusC, the outer-membrane protein.[9−11] The Cus system has an additional fourth component, CusF, which acts as a periplasmic Cu(I)/Ag(I) metallochaperone and is vital for maximal metal resistance.[10,12] The periplasmic protein CusB is a member of the membrane fusion protein (MFP) family.[13] It is hypothesized to stabilize the tripartite intermembrane complex through

Received: May 14, 2013
Revised: August 26, 2013
Published: August 29, 2013

dx.doi.org/10.1021/bi400606b | *Biochemistry* 2013, 52, 6911−6923

its interactions with CusA and CusC, which is similar to the role that AcrA plays in the well-characterized multidrug efflux pump AcrAB-TolC.[14−17] Four domains were identified from the available crystal structures: the membrane proximal, beta-barrel, lipoyl, and α-helical domains. However, the most important part containing the three conserved metal-binding Met residues could not be resolved with crystallographic techniques.[14,18] Mutation studies conducted on these three Met residues have resulted in the loss of metal resistance in vivo.[19] Moreover, CusB and CusF interact temporarily in the presence of metal and have similar metal-binding affinities inferred from experimental findings where the metal in the medium gets distributed approximately equally between them when mixed in equimolar concentrations in vitro.[19,20] Direct metal transfer between these proteins has also been shown experimentally.[21] Chemical cross-linking/mass spectrometry experiments captured the CusB-CusF interaction, and underlined the significance of the N-terminal region in terms of protein-protein interactions and metal transfer.[22]

Recently the N-terminal region of CusB (CusB-NT) was experimentally found to exhibit metal transfer from CusF by itself, although not as effectively as the full-length CusB, which rules out the hypothesis that CusB acts simply as a metal chelator. It must be inducing a structural change in the rest of the chain that gives rise to higher resistance. This change was actually captured experimentally, implying conformational flexibility.[23] Yet, the retention of the metal-transfer ability by truncated CusB encouraged us to examine this region by itself to unravel the dynamics and structure of this highly disordered entity. Being highly disordered in their N- and C-terminal regions is not an uncommon feature of membrane fusion proteins. The N- and C-termini of the multidrug efflux MFP's MexA and MacA could not be crystallographically resolved either.[24−26] These regions appear to be flexible and dynamic in nature, as the observed disorder suggests.

Structural and dynamic properties of proteins can be probed simultaneously by molecular dynamics (MD) simulations, which make them especially useful for studying protein structure and folding.[27] Recent attempts to examine the folding pathways of several peptides, including villin headpiece,[28−37] bovine acyl-coenzyme A,[38] Trp-cage,[39−44] and Alzheimer amyloid beta 10−35 peptide,[45] resulted in success, suggesting that this technique has the potential to provide molecular-level insights into the structure of CusB.

As stated above, the available crystal structures of the periplasmic protein CusB lack the N-terminus. The absence of this very important piece hampers the ultimate modeling of the full metal efflux pump CusCFBA, where metal uptake in the periplasm cannot be modeled. Determining a viable structural ensemble for CusB, whose members could transfer the metal from the open form of CusF, remains essential to reach our ultimate goal: simulating the entire metal-extrusion process via CusCFBA. Shedding light on this highly disordered domain would provide researchers with an ensemble of full model structures of CusB that have been missing so far. Moreover, the insights gained would facilitate further insight into other disordered proteins.[46]

In this article, we examine the structural and dynamic changes of the N-terminal region of CusB upon Cu(I) binding by simulating the apo and holo CusB-NT using MD simulations and analyzing the sampled phase space in each case with various analysis tools.

METHODS

We ran extensive MD simulations of 25 different initial models of the N-terminal region of CusB that were obtained with the protein structure prediction tools Quark,[47] I-TASSER,[48,49] and Sparks-X.[50] Quark is an ab inito protein folding and structure prediction algorithm, which is accessible through a web server. It constructs the 3D protein model from the amino acid sequence only with no global template information. Thus, it is useful for proteins without homologous templates. It was ranked as the leading server in free modeling in the CASP9 experiment.[47,51,52] The second folding server that we used, I-TASSER, aims to predict protein structure and function by building 3D models on the basis of multiple threading alignments and assembly simulations. The predicted models are then matched with the BioLiP protein function database.[53] The latest CASP experiments ranked it as the number one server for protein structure prediction in addition to its first rank for function prediction in CASP9.[54] The third server that we employed was Sparks-X, which is one of the best performers in the CASP9 experiment for single-method fold recognition. It is a template-based modeling algorithm which was established by weighted matching of multiple profiles and provides an improved scoring after taking into account the errors in the predicted 1D structural properties.[50] We refer to residues 29−79 of chain C in the PDB structure 3NE5[18] as the N-terminal region of CusB. Residues 1−28 constitute the signaling region and were not included in this study. The residue numbering is reported after subtraction of the signaling peptide.

Both the apo and Cu(I)-bound versions of the N-terminal region of CusB were simulated where all the protein and solvent atoms were treated explicitly. Each model system was solvated with the TIP3P triangulated water model[55] in a periodically replicated rectangular water box whose sides were at least 10 Å away from the solute atoms. The systems were neutralized by the addition of Na^+ ions, and charged amino acids were modeled in the protonated states obtained with the H++ protonation state server at physiological pH.[56] The apo models were run through an energy-minimization protocol of seven stages involving the minimization of only the solvent atoms and the counterions (stage 1), the minimization of the hydrogen atoms (stage 2), the minimization of the side chains by gradually decreasing the harmonic positional restraints acting on them (stages 3−6), and finally the energy minimization of the whole system with no positional restraints (stage 7). A total of 157 000 energy-minimization steps were performed in total, 47 000 of which used the steepest descent protocol.[57] The remaining 110 000 steps utilized the conjugate gradient method for minimization.[57] A two-stage equilibration protocol followed: first, the systems were heated slowly from 0 to 300 K over 200 ps of MD within the canonical ensemble (NVT) by maintaining a weak harmonic restraint on the protein and then they were simulated for 10 ns at 300 K to check for the stability of the peptide chains after removal of the harmonic restraints at a constant pressure of 1.0 bar for an isobaric, isothermal ensemble (NPT) using Langevin dynamics with a collision frequency of 1.0 ps^{-1}.[58] Periodic boundary conditions were imposed on the systems during the calculation of nonbonded interactions at all minimization and equilibration stages, while the lengths of the covalent bonds involving hydrogen were constrained and their interactions were omitted with the SHAKE algorithm while the systems were heated.[59] All

of the restraints on the systems were released before we started with the MD production runs with the apo models.

The metal-coordination site of CusB was experimentally found to consist of three Met residues (M21, M36, and M38) aligned in an almost equilateral triangle through their sulfur atoms, where the S−Cu(I) distance was measured to be 2.3 Å in EXAFS experiments.[21,60] To orient correctly the metal-binding residues for Cu(I) binding in the apo systems, we employed 2 ns of steered MD on each of the 25 systems to bring the metal-binding residues (M21, M36, and M38) into a hypothetical preorganized state where the sulfur atoms would be 4.15 Å from each other in an equilateral triangular alignment using a harmonic restraint of 1000 kcal/mol·Å^2. This distance value was obtained from our optimization efforts involving methionine side chains and a Cu(I) ion at the density functional theory (DFT) QM level which were performed with Gaussian 09[61] using the M06L DFT functional[62] in conjunction with the double-ζ-quality LANL2DZ pseudopotential basis set for Cu(I)[63] and the Pople-type basis set 6-31G* for all the other atoms.[64] In addition to providing the needed S−S distances and a general positioning in the metal site for the steered MD, these calculations confirmed the experimentally observed Cu(I)−S distance of 2.3 Å. Once we obtained the individual preorganized geometries for the metal site in 25 models, we introduced the Cu(I) ion in the center of mass of the three sulfur atoms of M21, M36, and M38 after we had stripped all the solvent and counterions from the final recorded snapshot at the end of the 2 ns simulation time, which yielded 25 holo models. The metal-bound systems were represented by a bonded model where the metal parameters were obtained with MTK++/MCPB functionality[65] of AmberTools version 1.5.[66] A frequency calculation was run on the optimized structure (M062X/LANL2DZ-6-31G*) of the binding site using Gaussian 09 to collect the bond and angle parameters.[67] The charges for the Cu(I)-bound ligands were calculated using RESP (M062X/6-31G*) within the MTK++ program.[68,69]

Having equilibrated the apo models and holo models, we proceeded with 200 ns of conventional MD and 400 ns of accelerated MD (aMD)[70] for each of these 50 systems, which adds up to a total simulation time of 5 microseconds of MD and 10 microseconds of aMD for both the apo and holo models. The MD simulations were run with the ff99SBildn force field[71] of the AMBER11 package,[66] whereas for the aMD simulations, the same force field in the AMBER12 suite was utilized.[72] The aMD parameters were obtained from the 200 ns MD runs for each of the 50 systems, as described in the AMBER12 manual, and an extra boost was applied to the torsions (Table SI.2). Throughout the production simulations, the SHAKE algorithm was used to constrain covalent bonds involving hydrogen. The Particle mesh Ewald (PME) method was utilized for long-range electrostatic interactions, and an 8 Å nonbonded cutoff was applied to limit the direct space sum in PME.[73] The temperature of the systems was maintained at 300 K with Langevin dynamics (collision frequency of 1.0 ps^{-1}). Frames were collected every 2 ps. A selected subset of these snapshots were employed in distance, angle, dihedral, root-mean-square deviation (rmsd), root-mean-square fluctuation (rmsf), correlation analyses, entropy calculations, and clustering of trajectory frames with the ptraj utility of AmberTools version 1.5. The average linkage algorithm with a maximum eccentricity of 5 Å was used in clustering the frames. NMR chemical shifts were assigned to backbone atoms extracted from snapshots separated by 0.5 ns with SPARTA+.[74] A smaller subset of these snapshots with a frame separation of 20 ns was employed in secondary-structure predictions with the Stride program.[75] Visual molecular dynamics (VMD) was utilized to visualize the molecular structures,[76] whereas Gnuplot[77] and Grace software were used to plot the data.

RESULTS AND DISCUSSION

The N-terminus of CusB was found to be mostly disordered in previous circular dichroism and NMR experiments involving its apo and Ag(I)-bound versions.[23] Before we started with the MD simulations, we assessed the expected level of disorder in the CusB N-terminus with SPINE-D, an online artificial neural network server that classifies regions of a given protein chain as ordered or disordered on the basis of the protein sequence only.[53] The results are shown in Figure 1 where CusB-NT is

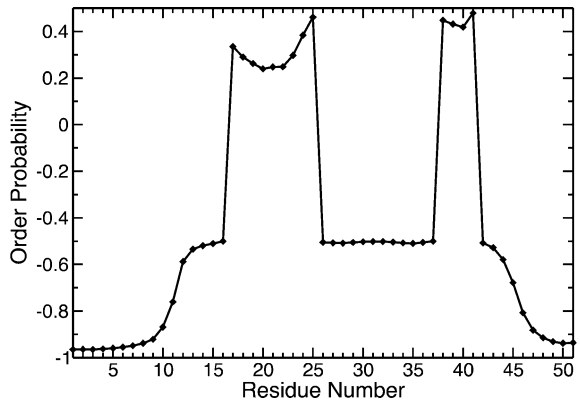


Figure 1. Predicted level of order for each residue by the online server SPINE-D. The scale for the order probability is from +1 to −1, where +1 implies a maximum probability for order, namely, complete order, and −1 expresses a maximum probability for disorder, in other words, total disorder.

predicted to have significant disorder, whereas some order was estimated for two regions around the residues 16−25 and 38−41, which roughly corresponded to the metal-coordination sites. This result, combined with the experimental findings, prepared us for the fact that a very careful analysis was needed to extricate insights into the intrinsically disordered protein.

Both in the apo and holo trajectories of the CusB N-terminal domain there are some motifs that were visually observed repeatedly: an antiparallel β-sheet configuration in the central part, a small β-sheet structure at the N-terminus of the CusB-NT chain that sometimes aligns with the antiparallel β sheets in the center to form a triple antiparallel β-sheet conformation, the N- and C-terminal tails folding into short helices, and the 1-turn α helix in the loop connecting the metal-binding residues M21 and M36. Figure 2 gives sample structures observed in the simulations containing some of these features. Naturally, the Cu(I) ion bound to the three Met residues, M21, M36 and M38, provides some degree of constraint to the chain motion that renders the holo form slightly more ordered around the metal-binding region. Hence, the apo protein adopts a larger number of conformations. For both forms, we observed the maximum mobility in the CusB-NT chain ends. The N-terminal end (residues 1−10) is even more mobile than the C-terminal end, and the latter would normally be attached to the rest of the CusB chain whose crystal structure had been already resolved by X-ray crystallography techniques. In both the holo and apo forms, the N-terminus of the CusB-NT chain often

dx.doi.org/10.1021/bi400606b | Biochemistry 2013, 52, 6911−6923

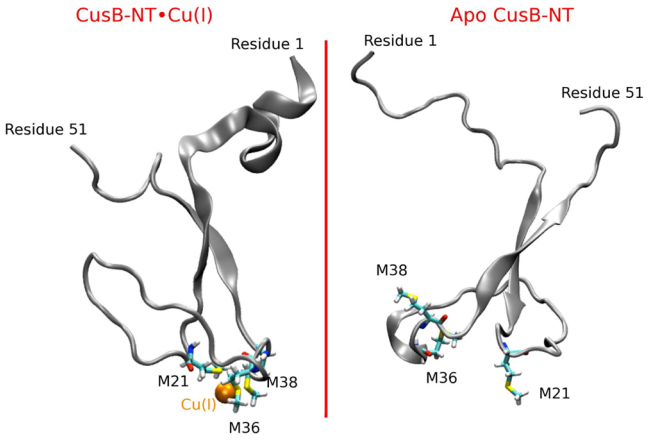


Figure 2. Cu(I)-bound version of the CusB N-terminal region (left) and its apo version (right). The central β motif with two antiparallel β strands, the loop connecting metal-binding residues M21 and M36, and the disordered tail regions are indicated in these sample structures. The short α helix on the N-terminal end of CusB-NT forms very frequently in both the holo and apo versions.

folds into a small α helix in addition to the β sheet that it forms occasionally. Although, as our rmsd analysis suggests (Figures 4 and SI.2 and discussion below), substantial structural changes occur during the simulations. We observe numerous structural motifs ranging from rather extended to more globular where the termini and the M21−M36 loop fold against the antiparallel β-sheet motif in the center. Hence, the β sheets act like a core region, leading to a rather compact shape along with the M21−M36 loop and the transient secondary structure (β strand or α helix) in the C-terminus of CusB-NT.

We also made use of secondary-structure predictions to quantify the observed structural similarities between the apo and Cu(I)-bound protein models. These were established using the Structural Identification (STRIDE) algorithm that uses atomic coordinates for assignment.[75] The frames were isolated from the MD and aMD trajectories, and the secondary-structure elements assigned to each frame were combined into a matrix that was displayed with color coding, as seen in Figure 3. We analyzed the MD and aMD data separately in each case. Since we simulated the holo version for a longer time using aMD, we ended up with more frames contributing to this matrix resulting from the frame separation was kept at 20 ns at all times. It must be noted that the covered phase space with aMD going from one frame to the next is larger because of the enhanced sampling of aMD. Hence, greater structural changes are to be expected. Actually, our test runs indicate an acceleration of 3−45-fold in sampling the phase space depending on the starting point (Figure SI.4), as tracked

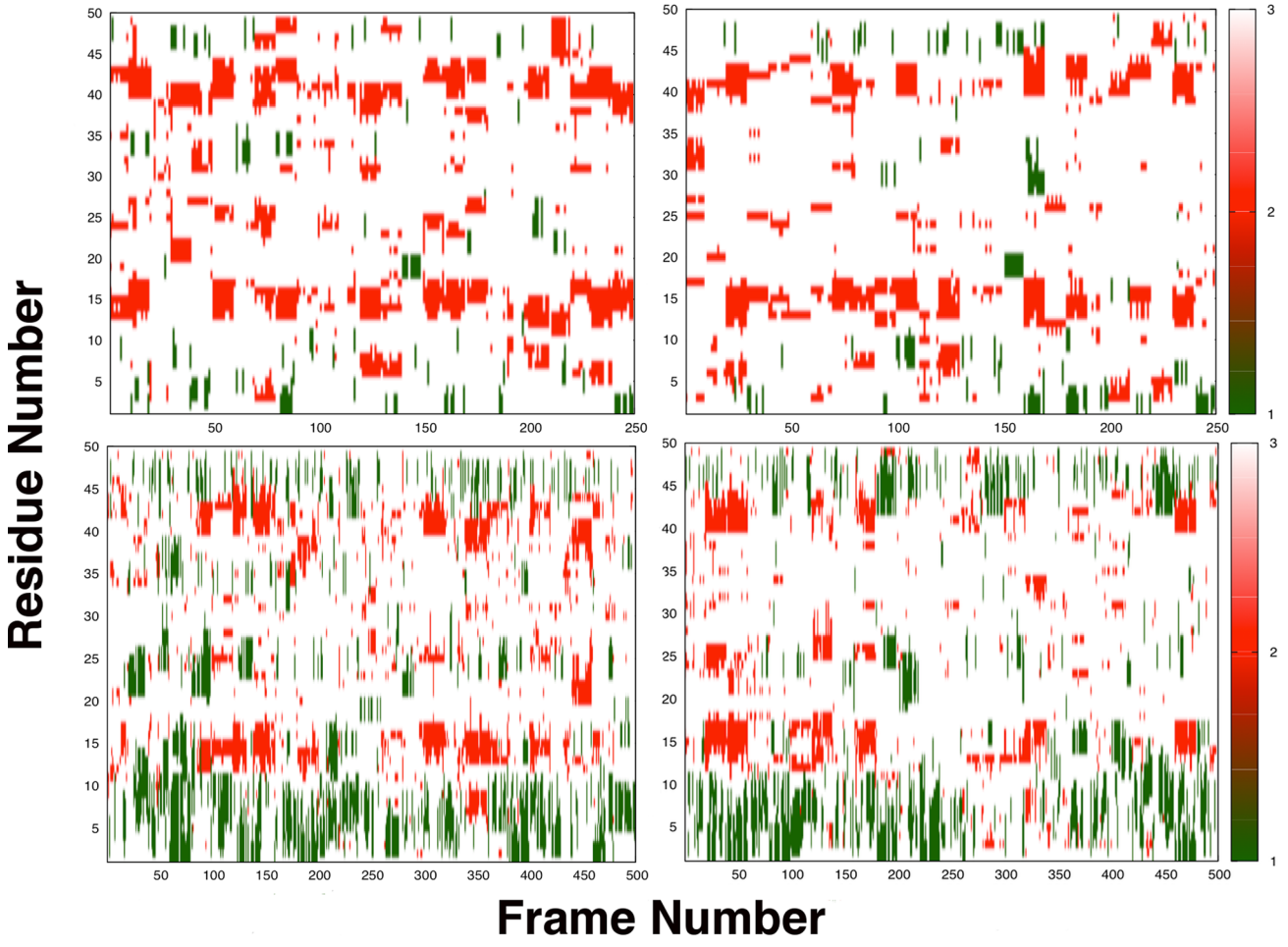


Figure 3. STRIDE secondary-structure predictions for the apo CusB-NT, MD trajectories (top left); Cu(I)-bound CusB-NT, MD trajectories (top right); apo CusB-NT, aMD trajectories (bottom left); and Cu(I)-bound CusB-NT, aMD trajectories (bottom right). Color scale: 1/green, α helix; 2/red, β sheet; and 3/white, turn, coil, or not assigned.

dx.doi.org/10.1021/bi400606b | *Biochemistry* 2013, 52, 6911−6923

from the backbone rmsd profiles of arbitrary model structures. Thus, the MD trajectories tend to retain the secondary-structure elements over consecutive frames. However, the aMD trajectories give a more striped appearance resulting from the increased conformational variability of the snapshots and do not stack up as seen in MD matrices. The bottom panels reveal this effect very clearly: the longer the separation between the snapshots, the less continuous the matrix representation.

For both the apo and holo simulations, turns and coils shown in white dominate in terms of the observed secondary structures, which implies a large extent of disorder throughout the protein structure. In addition to these, the persistence of the β-sheet motif can be seen in the red spots around residues 13−16 and 40−44. In the aMD trajectories, the α helix in the beginning part of the chain (first 10 residues) dominates the picture. Helices are known to require time frames around 200 ns to form, which in turn makes them more rare in our MD trajectories[78,79] in comparison to the aMD trajectories. The small helix forming in the loop region connecting M21 and M36 sometimes appears in the apo aMD simulation, but it is not a persistent structural element for the holo version. The most significant insight gained through these matrix representations is how much disorder is seen throughout the protein chain. Nonetheless, this analysis gives structural and time-scale insights into the formation and unraveling of secondary-structural elements in a disordered protein.[80]

Root-mean-square deviation data served as another tool to assess the extent of structural variability of the apo and holo versions of the CusB N-terminus. The rmsd ranges and distributions observed in the MD trajectories for individual apo and holo models do not differ drastically, as seen in Figures 4 and SI.2. The mean rmsd for the apo aMD trajectories is 10.86 ± 2.11 Å, whereas for the holo aMD trajectories it is 12.32 ± 2.76 Å. The rmsd values of the apo models span a range of 5−18 Å where changes within one trajectory are minor (usually about 2 Å, although it reaches 6 Å for some models) (Figure SI.2).

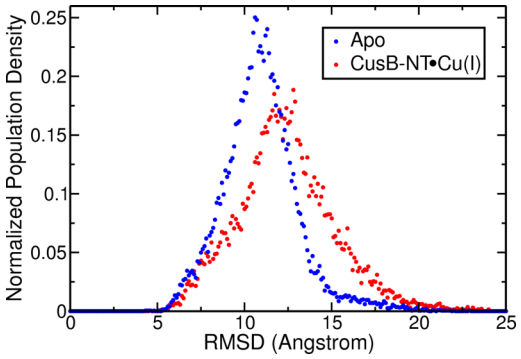


Figure 4. Distribution of rmsd values over 10 μs of aMD for the apo and Cu(I)-bound versions of the CusB-NT chain. The means and the standard deviations for both versions do not differ significantly, implying comparable structural mobilities before and after metal binding.

Surprisingly, the rmsd spread for the holo models is 6−22 Å, whereas the rmsd within one trajectory remains mostly at the much lower values of 2−3 Å. Here, we experience the benefits of having started our MD runs from 25 distinct starting points to end up with good overall sampling. The rmsd values for each starting structure with respect to an arbitrary starting structure were computed to determine how dissimilar the starting models were, and the resulting plot is given in the Figure SI.5. The aMD data broadens the observed rmsd ranges slightly: 5−23 Å for the apo model, and 4−24 Å for the holo model. More importantly, the rmsd ranges spanned by single trajectories exceed 10 Å in certain cases, which indicates better sampling, as expected, and would push the different conformations from different portions of the potential energy surface to other regions. In this way, energetic traps were avoided by switching to different starting points.

The Cu(I)-bound structures experience greater structural changes both in the MD and aMD trajectories compared to their apo counterparts, which is counterintuitive with what one would anticipate on the basis of the reasonable hypothesis that the metal ion would restrain the chain motion. This is observed both in the individual rmsd profiles and in the distribution of the rmsd values for the apo and holo trajectories. In Figure 4, the distribution of the rmsd values associated with the holo trajectories is broader and its mode is greater than the one for the apo distribution. Although the analyzed trajectories are from the production phase, it appears that the polypeptide chain is still adapting to this extra constraint by fluctuating to a greater extent and searching for stable conformations. More detailed rmsd profiles are shown in Figure SI.2. This is in contrast to recent work on CzrA, where metal binding (Zn(II)) lead to a reduction in the number of states sampled by CzrA relative to its apo state.[81]

The rmsf profile is a measure of the mobility of the α carbons of the protein chain. In this case, the analysis shows that the end residues are more mobile compared to the rest of the molecule for both the apo and holo conformations. The individual plots obtained from the individual MD trajectories contain various minima and local maxima in the interior parts of the chain (Figure SI.3). The rmsf profiles from the aMD trajectories provided useful insight in which some trends were observable: for the apo models, the local minima occur at residues T9, R12, I14, F27, S33, and K43, as seen in Figure 5. The reduced mobility around residue T9 is associated mostly with the formation of the α helix in the N-terminal tail of CusB-NT, which ceases to some extent upon the addition of metal, as indicated in the lower panels of Figure 3. I14 and K43 fall in the

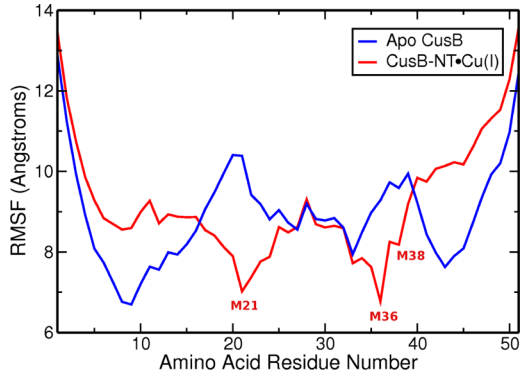


Figure 5. Atomic fluctuations for the apo (blue) and Cu(I)-bound (red) versions of the N-terminal region of CusB. Data were collected over a total of 10 μs of aMD simulations from 25 different starting points. Cu(I) binding seems to impact only the metal-binding residues, M21, M36, and M38. The cumulative rmsf graph obtained using the MD data is shown in Figure SI.8. General trends are conserved in the MD versus aMD results.

dx.doi.org/10.1021/bi400606b | Biochemistry 2013, 52, 6911−6923

β sheets that make up the central β-sheet motif, which emphasizes the tendency for this secondary-structural element to form and stay stable over long time frames. R12 is adjacent to the first β strand and is stabilized by the central β motif. F27 and S33, however, are involved in extensive hydrogen bonding with other residues in the M21−M36 loop (Figure 10 and discussion below). This intermolecular interaction network reduces the mobility of these two residues. The loop region between M21 and M36 becomes almost as mobile as the tails of the chain toward the middle for some of the trajectories (Figure SI.3). The incorporation of the metal introduces a constraint to free motion and diminishes the atomic fluctuations of the metal-ligating residues (indicated in Figure 5). The minima of the averaged atomic fluctuations for the holo version of the CusB N-terminal region are found at M21 and M36. Interestingly, the Cu(I)-binding residues, M21, M36, and M38, in the holo models do not necessarily show minimal fluctuations in some of the individual trajectories. We ensured that this does not trace back to the disruption of the metal site; the S−Cu(I) distances have been analyzed throughout and were confirmed to conserve the bonding distances at all times. Visual inspection aids in confirming these seemingly contradictory observations: The loop connecting the M21 and M36 residues moves to both sides of the central β-sheet motif similar to a flap, which enhances the mobility of the Met residues bound to Cu.

Comparing the atomic fluctuation profiles of single models obtained from the aMD trajectories reveals that the minima are less spread in the holo chain (Figure SI.3). The fact that the immobilization effect remains only local around the metal-binding residues becomes even more obvious in the cumulative rmsf profiles in Figure 5. When looking at the individual profiles of the apo models, on the other hand, the motions seem to be dispersed along the chains more evenly (Figure SI.3). However, this artifact disappears in the cumulative graph.

We next calculated the NMR chemical shifts for the apo and holo forms of CusB-NT using SPARTA+.[74] These shifts were calculated for a total of 20 000 snapshots of the apo and holo models and averaged. The program could only calculate 43 backbone N atoms and their associated H atoms because the first amino acid residue is exempt from predictions and there are seven Pro residues without an amide hydrogen, leaving 43 residues in total. The average shifts for the backbone nitrogens (N) and their hydrogens (HN) were plotted with and without their standard deviations (Figure 6). The conclusion drawn from these calculations is that no striking structural differences other than the local organization of the metal-binding site occur between the apo and the Cu(I)-bound versions of the CusB N-terminus, which is consistent with the experimental observations involving Ag(I).[23] Additionally, this resembles the metallochaperone CusF, where the apo and metal-bound structures were similar to each other.[12,82−84]

In addition to the above, we investigated numerous structural properties to identify the structural changes induced by metal binding. To determine these properties, we focused on the prominent structural motifs. The first was the β-sheet motif in the center of the chain. The distance between the two antiparallel β strands measured from the center of mass of the residues 13−16 (strand 1) to the center of mass of the residues 41−44 (strand 2) emerged as a good coordinate to estimate the extent to which these two β sheets maintained their antiparallel alignment. This distance was measured to be around 5−6 Å when the two β sheets are found to be in an antiparallel

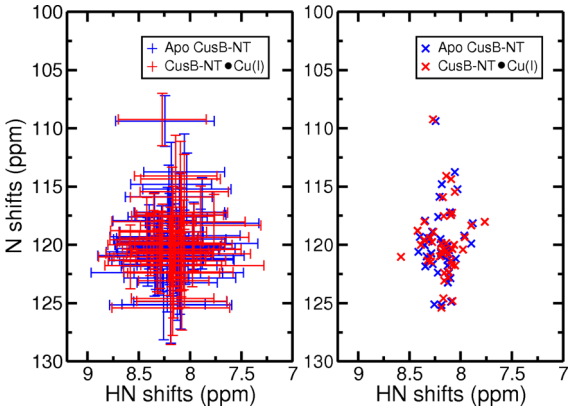


Figure 6. Average NMR shifts for backbone N atoms and their associated H atoms for the apo (blue) and Cu(I)-bound (red) versions of the CusB N-terminus predicted by SPARTA+ (right) and their standard deviations (left).

configuration. The inset of the Figure 7 further illustrates this distance. The histogram showing the distribution of this

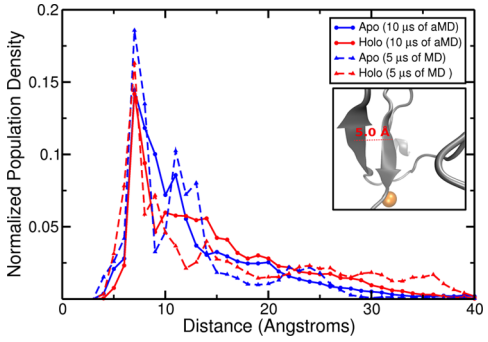


Figure 7. Distribution of the distances between the two β sheets forming the central β motif. The distribution from the apo MD and aMD trajectories are shown in blue, and the data from the holo MD and aMD trajectories are displayed in red. The inset visualizes this property. In this particular conformational alignment, the distance is exactly 5.0 Å. The vast range of distances suggests that we have extensively sampled the available phase space.

distance over the apo and holo trajectories shows that the majority of the observed snapshots locates the two β strands as being 5−6 Å apart from each other, which implies that this central motif is one of the key structural elements in this disordered protein chain (Figure 2). The range for this specific distance spans values from 3 to 66 Å, suggesting that most, if not all, of the available phase space is visited.

Second, to check the relative positioning of the two β strands belonging to the central β motif, the dihedral angle formed by the centers of mass of the L15, I14, K43, and P42 residues was probed (Figure 8, right). Dihedral angles of 0° imply an antiparallel conformation. As seen in the histogram in Figure 8, the global maximum occurs at around 0° for both the apo and the holo trajectories of CusB-NT, suggesting that the antiparallel β conformation is the most favored. Other peaks occurring around −120, −30, 60, and 90° indicate other common alignments in the trajectories. The similarity in behavior of the apo and holo models underlines the fact that this β motif is one of the key structural features of the CusB N-terminal region, as suggested by the data shown in Figure 7.

dx.doi.org/10.1021/bi400606b | *Biochemistry* 2013, 52, 6911−6923

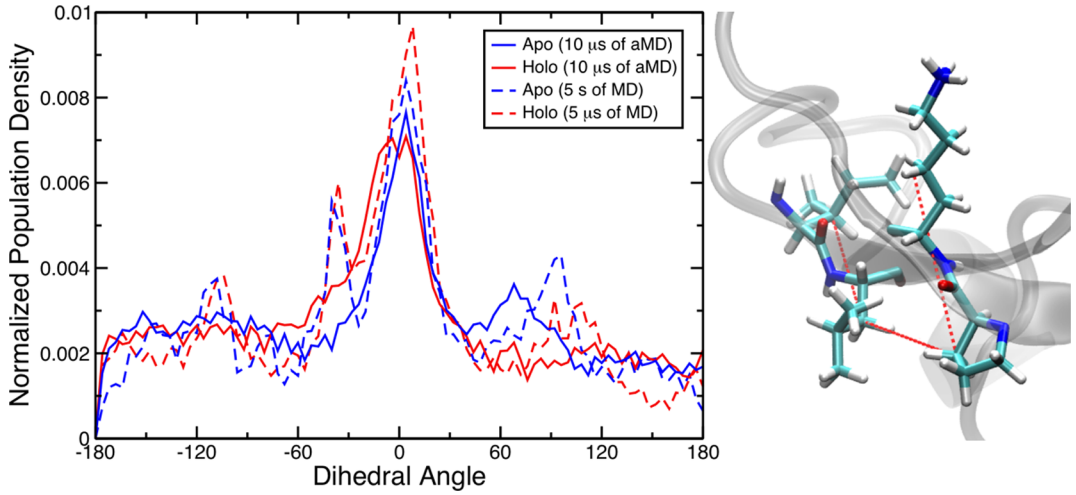


Figure 8. Dihedral angle formed by L15, I14, K43, and P42 as a measure of the relative β-sheet positions in the central motif (right). The distribution of this property is shown in the plot over the apo (blue) and holo (red) trajectories.

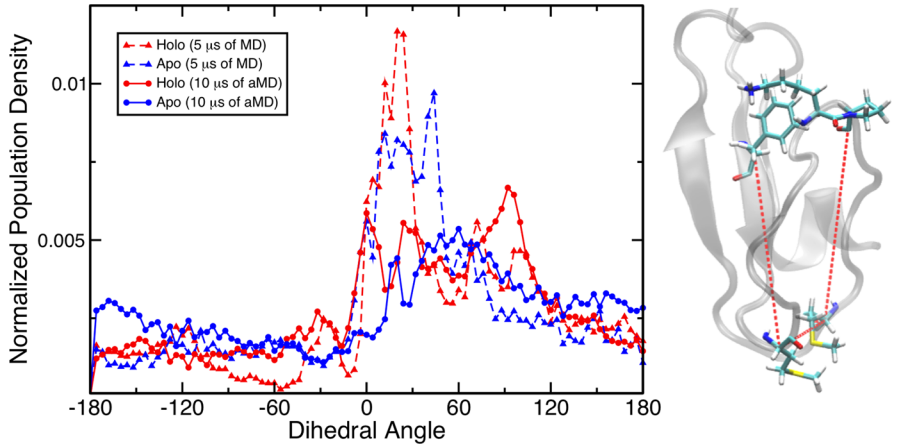


Figure 9. Dihedral involving the center of mass of F16 (in the central β conformation), M21, and M36 as well as the common center of mass of residues D28, K29, and P30 is shown in the inset. The distribution of this property is shown in the plot over the apo (blue) and holo (red) trajectories.

Further analysis aimed to explore the conformational preferences of the loop connecting M21 and M36 relative to the central β motif, whose importance we have highlighted. The dihedral involving the center of mass of F16 (in the central β sheet), M21, and M36 as well as the common center of mass of residues D28, K29, and P30, which falls on the midpoint of the big loop connecting M21 and M36, was examined for this purpose (Figure 9, inset). Not surprisingly, the apo and Cu(I)-bound models do not differ significantly. It appears that this loop tends to locate itself somewhere on the range of 0−120° from the central β motif, whereas for the holo models, angles of 0−30, 70−80, and 90−100° dominate. For the apo models, dihedrals of 0−60° are more common (Figure 9). As more sampling is included in the analysis with the aMD trajectories, the peaks covering the range of 0−60° flatten out for both the apo and the holo models, giving rise to a more dispersed distribution. The distribution of the data obtained from the holo aMD trajectories shows two maxima at around 0−30 and 90°. The one associated with the apo aMD data peaks at around 30−70°.

To understand further the spatial alignment of the loop, a distance analysis was conducted on the residues that constitute the loop (residues 22−35) and the ones close to the loop (residues 17, 19, 37, and 39), which may interact with the loop residues. Figure 10 shows these distance trends: hydrogen bonding between the backbone oxygen of S33 and the amide hydrogen of F35 occurs in the majority of the snapshots taken from both the apo and holo trajectories. This is by far the most prevalent interaction inferred from this analysis. In a great extent of the snapshots, this hydrogen bond is a strong one because the carbonyl oxygen−amide nitrogen distance is less than 3 Å. S33 and F35 are located at the hinge of the loop when it folds against the β motif, and they favor this type of folded alignment. The side chain oxygen of S33 and the amide hydrogen of F35 also interact in a significant portion of the frames in the apo and Cu(I)-bound simulations. This interaction alternates with the previous S33O−F35NH interaction from time to time, emphasizing the importance of F35 and S33 in determining the spatial arrangement of the loop. Alternatively, the side chain oxygen of S33 interacts with the amide hydrogen of M36 in some of the snapshots taken from both the apo and holo trajectories. This adds to the structural importance of residue M36: binding the Cu(I) ion and aiding in structural organization of the domain by facilitating the folding of the big loop with hydrogen bonds at the hinge point. In addition to these interactions, the

dx.doi.org/10.1021/bi400606b | *Biochemistry* 2013, 52, 6911−6923

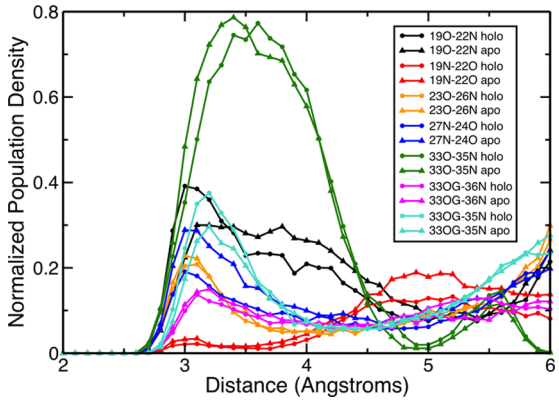


Figure 10. Distance analysis plot for residues in and adjacent to the big loop connecting residues M21 and M36. The triangles indicate data obtained from the apo trajectories, whereas the dots indicate the holo trajectories. The numbers in the inset are the residue numbers. O is the backbone oxygen, N is the amide nitrogen, and OG is the side chain oxygen of S33.

backbone carbonyl oxygen of D19 and the amide H of Y22 form a backbone hydrogen bond on the other end of the loop, and this acts as a hinge as well, which is observed in some of the apo and even more frequently in the holo conformations. Very rarely, the amide H of D19 and the backbone O of Y22 form hydrogen bonds. Another interesting point revealed in the distance analysis is the formation of the 1-turn helix within the M21−M36 loop, whose formation is made possible through the backbone hydrogen bonding involving residues P23−R26 and N24−F27. This is observed in both of the apo and holo models.

An additional structural property was examined to gain insight into the general shape of the M21−M36 loop. The distance between the Cu(I) ion and the center of mass of each residue constituting the loop connecting M21 to M36 was calculated for every snapshot taken from the holo trajectories and averaged (Figure 11). This analysis reveals that the loop is

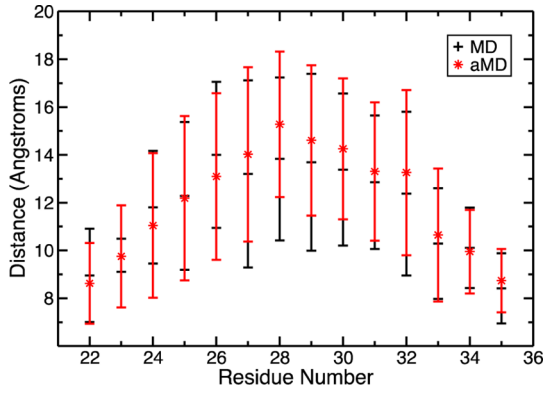


Figure 11. Distance between the Cu(I) ion and the centers of mass of each residue constituting the loop connecting M21 to M36 along with error bars.

mostly oval-shaped. The standard deviations increase for residues closer to the middle of the loop, implying that they are more mobile as they move farther away from the hinge points.

Finally, to examine the reversibility of the structural effects caused by metal binding, we randomly picked one of the 25 models and removed the Cu(I) ion from snapshots at 100 ns (snapshot 1) and 200 ns (snapshot 2) into the conventional MD simulation. We solvated, minimized, and equilibrated these two structures using the protocol described in the Methods section and ran 200 ns of conventional MD on them. We tracked the backbone motions of the new trajectories both visually and with root-mean-square fluctuation (rmsf) and root-mean-square deviation (rmsd) analyses. These were compared to the corresponding analyses of the original apo trajectory of the same model. The comparisons are shown in Figures SI.9 and SI.10. Accordingly, both the rmsd and rmsf profiles of snapshot 1 track the backbone motions of the original apo trajectory very closely, which suggests that they may be covering similar regions of phase space. However, for snapshot 2, the recovery of apo behavior was not as clear. Although the original apo behavior of this particular model is not obtained in the snapshot 2 trajectory, the main secondary-structural elements, like the central β motif and the loop between M21 and M36, emerge in the middle of the trajectory, which is a point visited by other apo models, as the STRIDE analysis qualitatively demonstrates.

DECOMPOSING THE DISORDER

Our original hypothesis envisioned that we would be able to extract probable structures for the Cu(I)-bound version of the CusB N-terminal region, which would further guide experimental analysis of this domain. Through visual inspection and the rmsf analyses, we observed the chain ends to be, not surprisingly, more mobile than the rest of the chain. This led to the idea of splitting the chain into three "subdomains" and analyzing these individually: the tails, the central β motif, and the Cu(I) site along with the big loop. These are shown in Figures 12 and 13. The consideration of the atomic fluctuations

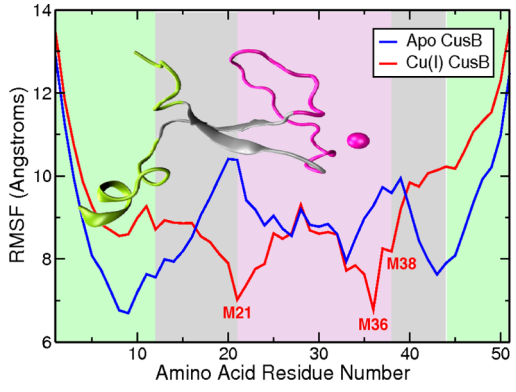


Figure 12. Root-mean-square fluctuation profiles of the apo and holo CusB N-terminus revisited with background color coding. Green indicates the tails, gray, the central β motif, and pink, the M21−M36 loop. The inset shows these three subdomains.

of the α carbon atoms presents a useful quantitative way to justify how we determined the subdomains. As seen in Figure 12, the fluctuations behave consistently in the designated subdomains. Both tails show elevated backbone motion in the Cu(I)-bound version compared to the apo version, which is designated as the first subdomain and displayed with a green background. In the central β motif there is a changing trend in the rmsf: the β region of the holo models become less mobile toward the metal-binding site, whereas this is inverted for the apo models. In other words, the chain fluctuations decrease in

dx.doi.org/10.1021/bi400606b | Biochemistry 2013, 52, 6911−6923

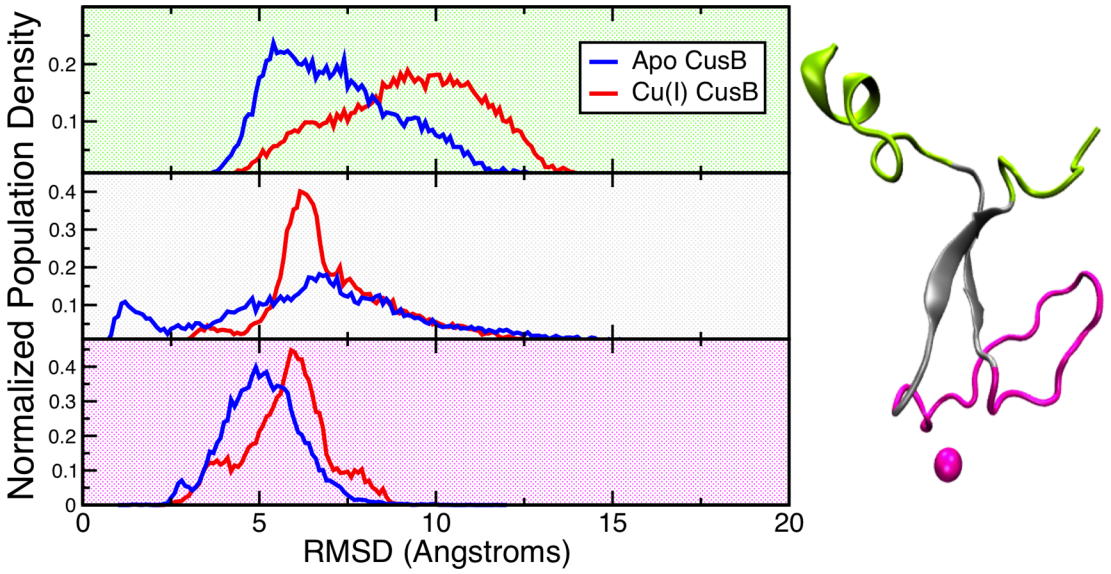


Figure 13. N-terminal domain of CusB split into three subdomains to decompose the structural disorder (right): the tails (green), the central β motif (gray), and the Cu(I)-site and M21−M36 loop (magenta). The rmsd distributions associated with these subdomains extracted from 10 μs of aMD runs on the Cu(I)-bound protein (left). Green background, subdomain 1; gray background, subdomain 2; and pink background, subdomain 3.

the holo models as one approaches the Cu(I) site, whereas the opposite behavior is observed in the apo models. Thus, the region can also be considered as a distinct subdomain and is indicated with a gray background in Figure 12. The M21−M36 loop region constitutes the third subdomain, as it displays a local stabilization upon metal binding and a higher mobility in the apo models (pink background).

In this partitioning, the most stable subdomain was the Cu(I) site and the M21−M36 loop followed by the central β-sheet motif and then the tails, as seen in the rmsd distribution in Figure 13. The rmsd distribution profile shows that the tails have the highest extent of structural changes, yielding the widest rmsd distribution with the greatest mode value of almost 10 Å, further confirming that the most significant chain motion happens there. Interestingly, the biggest change in the rmsd distributions upon metal binding is detected in the β-motif region, which points to the largest acquired ordering in contrast to what might have been expected to happen in the third subdomain containing the metal site and the loop. The β motif is likely to assist in generating a preorganized state suitable for metal binding, and this hypothesis emphasizes its structural significance. The average rmsf profiles in Figure 12 also support this finding, where the α carbons of the residues constituting the Cu(I) site and the big loop display the lowest atomic fluctuations, whereas the highest atomic fluctuations stem from the tails of the protein chain. Additionally, clustering the split trajectories yielded a much lower number of clusters in general, whereas the lowest number of clusters were produced by the most stable subdomain constituting of the Cu(I) site and the M21−M36 loop. More specifically, a subset of 6000 frames from each subdomain was clustered using the ptraj utility of AmberTools version 1.5 using the average linkage algorithm, which resulted in 303, 22, and 4 clusters for subdomains 1, 2, and 3, respectively. Hence, we could observe that the tails are mostly responsible for the disordered state of the N-terminal region of CusB. This led us to isolate the tails (subdomain 1) from the CusB-NT and cluster the remainder of the N-terminus, namely, the subdomains 2 and 3. This protocol provided us with 261 clusters, which is a large decrease from the 1646 clusters obtained when the entire CusB-NT chain was analyzed in an analogous manner. Thus, the observed disorder can be largely attributed to subdomain 1, the N- and C-termini, of the CusB-NT domain. On the basis of this observation, we propose that even the smaller construct of subdomains 2 and 3 might form a functionally competent metal-binding domain.

■ CONCLUSIONS

In this study, we explored the conformational dynamics of the apo and Cu(I)-bound CusB N-terminal region to assess the impact of metal binding on this mostly disordered protein domain that lacked an experimentally resolved structure. To obtain viable starting points, we utilized the online protein-folding servers Quark,[47] I-TASSER,[48,49] and Sparks-X,[50] which gave promising results in CASP competitions. Twenty-five starting models were obtained and used in extensive MD simulations to study the structure and dynamics of the CusB N-terminal region. Moreover, we applied the accelerated MD enhanced sampling method of Pierce et al.[70] for another 10 μs of sampling in addition to the 5 μs of classical MD simulations. The Cu(I) ion was inserted into the apo models after generating preorganized metal-binding sites according to the available experimental information regarding the binding site. A bonded force field was created that attached the three coordinating Met residues (M21, M36, and M38) to the Cu(I) ion in a trigonal planar (triangular) alignment over the course of these simulations. The holo models generated in this way were simulated for the same amount of simulation time, which provided us with sufficient conformational dynamics and structural information to compare the apo and metal-bound versions of the CusB N-terminal domain.

Our analysis of the secondary-structure assignments, rmsd, atomic fluctuations of the α carbons, NMR chemical shift assignments, and various structural coordinates focusing on the more prominent motifs like the central β sheet and the M21−M36 loop showed that significant disorder was present in both the apo and Cu(I)-bound forms of the protein. Metal binding led to only a local ordering around the metal site, causing a modest impact on the disorder of the whole domain. To reduce

dx.doi.org/10.1021/bi400606b | *Biochemistry* 2013, 52, 6911−6923

the complexity of this general disorder, the protein was broken into three subdomains utilizing logically selected structural elements. The first subdomain consisted of both tails (residues 1−12 and 46−51), the second contained the two β strands in the center of the structure (residues 13−20 and 39−45), and the third subdomain consisted of the M21−M36 loop in the apo chain along with the Cu(I) ion in the holo version. The rmsd analysis carried out on these regions showed that the first subdomain with the N- and C-terminal tails demonstrated the most disorder and was not affected by metal binding. Surprisingly, the biggest structural impact upon Cu(I) binding occurs in the β-motif region, namely, the second subdomain, in contrast to our hypothesis that the binding site and the loop (the third subdomain) would be influenced the most. This finding suggests that the β-motif region is of the utmost structural importance in the function of CusB as a metal chelator in the CusCFBA metal-extrusion mechanism. Indeed, our simulations suggest that residues 13−45 would also function in a similar manner to the entire chain because the two tail regions do not form any significant structure during the simulations, and this prediction can be tested experimentally.

The role of CusB-NT for metal delivery offers a conundrum. The concept of preorganization[85] suggests that to have tight and selective metal-ion binding the receptor should adopt a structure similar to that of the complexed structure. Experimentally, CusB-NT appears disordered both in the apo and metal-loaded forms, as determined by NMR.[23] Computationally, there appears to be a modest level of structural preorganization in CusB-NT in residues 13−45, which was also seen for CusF,[12,82−84] but it is still remarkable that this protein, which performs an important and essential function, has characteristics that suggest that it might not sequester Cu(I) from the environment at levels typically seen in Cu(I) binding proteins.[86] This disorder appears to be carried over into the intact complex, as observed in the X-ray structures of the CusBA system.[18,87] Perhaps the association with CusF induces further order or once metal is bound in the intact complex the interactions between the metal-binding domain and the intact transporter induces further stabilization.

Our simulations concluded that CusB-NT largely remains disordered, with only the central region of this protein (residues 13−45) forming secondary-structure elements. Interestingly, this situation is not significantly affected by metal binding, in opposition to our initial biases. This study has set the stage to simulate this domain with its functional companion CusF, the metallochaperone of the CusCFBA Ag(I)/Cu(I) pump. Another goal is to attach the CusB N-terminus to the remainder of CusB to see how it interacts with the intact copper-transport system. The presence of these companion proteins might result in an increased order of this region as well as further insights into the metal-ion transport mechanism.

■ ASSOCIATED CONTENT

Supporting Information

Cu(I) binding site with the three coordinating residues, M21, M36, and M38; rmsd profiles of the apo and holo models with respect to an arbitrary structure from one of the trajectories; atomic fluctuations of alpha C's in the N-terminal region of CusB; comparison of the rmsd profiles for trajectories of arbitrary individual models extracted from MD and aMD simulations; rmsd of each starting structure with respect to an arbitrary starting structure; full CusB containing six peptide chains that lack their N-terminal regions; two chains of CusB shown with a hypothetical disordered N-terminus and the metallochaperone CusF; atomic fluctuations for the apo and Cu(I)-bound versions of the N-terminal region of CusB; rmsd and rmsf profiles of the original apo model and snapshots 1 and 2; parameters for the methionine residues (M21, M36, and M38) and the Cu(I) ion comprising the bonded model; entropy values calculated from the MD trajectories of the 25 Cu(I)-bound models; and aMD parameters used in the calculations of the apo and Cu(I)-bound models. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 517-884-2540. Fax: 517-353-7248. E-mail: kmerz1@gmail.com.

Present Addresses

[†]Department of Chemistry, University of New Orleans, 2000 Lake Shore Drive, New Orleans, Louisiana 70148, United States.

[‡]Institute for Cyber Enabled Research, Department of Chemistry and Department of Biochemistry and Molecular Biology, Michigan State University, 578 South Shaw Lane, East Lansing, Michigan 48824-1322, United States.

Funding

This study was funded by National Institutes of Health grant nos. GM044974 and GM066859.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Ross Walker, Dr. Romelia Salomon, and Antonia Mey for useful scientific discussions. We acknowledge the University of Florida High Performance Computing Center for providing technical support.

■ ABBREVIATIONS

E. coli, *Escherichia coli*; ROS, reactive oxygen species; RND, resistance-nodulation-division; MFP, membrane fusion protein; CASP, Critical Assessment of protein Structure Prediction; AMBER, Assisted Model Building with Energy Refinement; TIP3P, transferable intermolecular potential 3P; MD, molecular dynamics; aMD, accelerated molecular dynamics; EXAFS, Extended X-ray Absorption Fine Structure; DFT, density functional theory; QM, quantum mechanics; RESP, restrained electrostatic potential; MCPB, metal center parameter builder; rmsd, root-mean-square deviation; rmsf, root-mean-square fluctuation; NMR, nuclear magnetic resonance; SPARTA, Shifts Prediction from Analogy in Residue Type and Torsion Angle; VMD, visual molecular dynamics; CusB-NT, CusB N-terminus

■ REFERENCES

(1) Zhang, Y., and Gladyshev, V. N. (2009) Comparative genomics of trace elements: Emerging dynamic view of trace element utilization and function. *Chem. Rev. 109*, 4828−4861.

(2) Macomber, L., Rensing, C., and Imlay, J. A. (2007) Intracellular copper does not catalyze the formation of oxidative DNA damage in *Escherichia coli*. *J. Bacteriol. 189*, 1616−1626.

(3) Poole, K., and Srikumar, R. (2001) Multidrug efflux in *Pseudomonas aeruginosa*: Components, mechanisms and clinical significance. *Curr. Top. Med. Chem. 1*, 59−71.

(4) Piddock, L. J. V. (2006) Multidrug-resistance efflux pumps - not just for resistance. *Nat. Rev. Microbiol. 4*, 629−636.
(5) Yang, S., Lopez, C. R., and Zechiedrich, E. L. (2006) Quorum sensing and multidrug transporters in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A. 103*, 2386−2391.
(6) Tseng, T. T., Gratwick, K. S., Kollman, J., Park, D., Nies, D. H., Goffeau, A., and Saier, M. H. (1999) The RND permease superfamily: An ancient, ubiquitous and diverse family that includes human disease and development proteins. *J. Mol. Microbiol. Biotechnol. 1*, 107−125.
(7) Dinh, T., Paulsen, I. T., and Saier, M. H. (1994) A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of Gram-negative bacteria. *J. Bacteriol. 176*, 3825−3831.
(8) Zgurskaya, H. I., and Nikaido, H. (2000) Multidrug resistance mechanisms: Drug efflux across two membranes. *Mol. Microbiol. 37*, 219−225.
(9) Franke, S., Grass, G., and Nies, D. H. (2001) The product of the ybdE gene of the *Escherichia coli* chromosome is involved in detoxification of silver ions. *Microbiology (Reading, U.K.) 147*, 965−972.
(10) Franke, S., Grass, G., Rensing, C., and Nies, D. H. (2003) Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli*. *J. Bacteriol. 185*, 3804−3812.
(11) Munson, G. P., Lam, D. L., Outten, F. W., and O'Halloran, T. V. (2000) Identification of a copper-responsive two-component system on the chromosome of *Escherichia coli* K-12. *J. Bacteriol. 182*, 5864−5871.
(12) Loftin, I. R., Franke, S., Roberts, S. A., Weichsel, A., Heroux, A., Montfort, W. R., Rensing, C., and McEvoy, M. M. (2005) A novel copper-binding fold for the periplasmic copper resistance protein CusF. *Biochemistry 44*, 10533−10540.
(13) Saier, M. H., Tam, R., Reizer, A., and Reizer, J. (1994) Two novel families of bacterial-membrane proteins concerned with nodulation, cell-division and transport. *Mol. Microbiol. 11*, 841−847.
(14) Su, C. C., Yang, F., Long, F., Reyon, D., Routh, M. D., Kuo, D. W., Mokhtari, A. K., Van Ornam, J. D., Rabe, K. L., Hoy, J. A., Lee, Y. J., Rajashankar, K. R., and Yu, E. W. (2009) Crystal structure of the membrane fusion protein CusB from *Escherichia coli*. *J. Mol. Biol. 393*, 342−355.
(15) Tikhonova, E. B., and Zgurskaya, H. I. (2004) AcrA, AcrB, and TolC of *Escherichia coli* form a stable intermembrane multidrug efflux complex. *J. Biol. Chem. 279*, 32116−32124.
(16) Zgurskaya, H. I., and Nikaido, H. (1999) Bypassing the periplasm: Reconstitution of the AcrAB multidrug efflux pump of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A. 96*, 7190−7195.
(17) Zgurskaya, H. I., and Nikaido, H. (2000) Cross-linked complex between oligomeric periplasmic lipoprotein AcrA and the inner-membrane-associated multidrug efflux pump AcrB from *Escherichia coli*. *J. Bacteriol. 182*, 4264−4267.
(18) Su, C. C., Long, F., Zimmermann, M. T., Rajashankar, K. R., Jernigan, R. L., and Yu, E. W. (2011) Crystal structure of the CusBA heavy-metal efflux complex of *Escherichia coli*. *Nature 470*, 558−U153.
(19) Bagai, I., Liu, W., Rensing, C., Blackburn, N. J., and McEvoy, M. M. (2007) Substrate-linked conformational change in the periplasmic component of a Cu(I)/Ag(I) efflux system. *J. Biol. Chem. 282*, 35695−35702.
(20) Kittleson, J. T., Loftin, I. R., Hausrath, A. C., Engelhardt, K. P., Rensing, C., and McEvoy, M. M. (2006) Periplasmic metal-resistance protein CusF exhibits high affinity and specificity for both Cu-I and Ag-I. *Biochemistry 45*, 11096−11102.
(21) Bagai, I., Rensing, C., Blackburn, N. J., and McEvoy, M. M. (2008) Direct metal transfer between periplasmic proteins identifies a bacterial copper chaperone. *Biochemistry 47*, 11408−11414.
(22) Mealman, T. D., Bagai, I., Singh, P., Goodlett, D. R., Rensing, C., Zhou, H., Wysocki, V. H., and McEvoy, M. M. (2011) Interactions between CusF and CusB identified by NMR spectroscopy and chemical cross-linking coupled to mass spectrometry. *Biochemistry 50*, 2559−2566.
(23) Mealman, T. D., Zhou, M. W., Affandi, T., Chacon, K. N., Aranguren, M. E., Blackburn, N. J., Wysocki, V. H., and McEvoy, M. M. (2012) N-terminal region of CusB is sufficient for metal binding and metal transfer with the metallochaperone CusF. *Biochemistry 51*, 6767−6775.
(24) Akama, H., Matsuura, T., Kashiwagi, S., Yoneyama, H., Narita, S. I., Tsukihara, T., Nakagawa, A., and Nakae, T. (2004) Crystal structure of the membrane fusion protein, MexA, of the multidrug transporter in *Pseudomonas aeruginosa*. *J. Biol. Chem. 279*, 25939−25942.
(25) Higgins, M. K., Bokma, E., Koronakis, E., Hughes, C., and Koronakis, V. (2004) Structure of the periplasmic component of a bacterial drug efflux pump. *Proc. Natl. Acad. Sci. U.S.A. 101*, 9994−9999.
(26) Yum, S. W., Xu, Y. B., Piao, S. F., Sim, S. H., Kim, H. M., Jo, W. S., Kim, K. J., Kweon, H. S., Jeong, M. H., Jeon, H. S., Lee, K., and Ha, N. C. (2009) Crystal structure of the periplasmic component of a tripartite macrolide-specific efflux pump. *J. Mol. Biol. 387*, 1286−1297.
(27) Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on. *Science 338*, 1042−1046.
(28) Beauchamp, K. A., Ensign, D. L., Das, R., and Pande, V. S. (2011) Quantitative comparison of villin headpiece subdomain simulations and triplet-triplet energy transfer experiments. *Proc. Natl. Acad. Sci. U.S.A. 108*, 12734−12739.
(29) Bowman, G. R., Beauchamp, K. A., Boxer, G., and Pande, V. S. (2009) Progress and challenges in the automated construction of Markov state models for full protein systems. *J. Chem. Phys. 131*, 124101-1−124101-11.
(30) Bowman, G. R., and Pande, V. S. (2010) Protein folded states are kinetic hubs. *Proc. Natl. Acad. Sci. U.S.A. 107*, 10890−10895.
(31) Bowman, G. R., Voelz, V. A., and Pande, V. S. (2011) Taming the complexity of protein folding. *Curr. Opin. Struct. Biol. 21*, 4−11.
(32) Duan, Y., and Kollman, P. A. (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science 282*, 740−744.
(33) Freddolino, P. L., Harrison, C. B., Liu, Y. X., and Schulten, K. (2010) Challenges in protein-folding simulations. *Nat. Phys. 6*, 751−758.
(34) Jayachandran, G., Vishal, V., and Pande, V. S. (2006) Using massively parallel simulation and Markovian models to study protein folding: Examining the dynamics of the villin headpiece. *J. Chem. Phys. 124*, 164902-1−164902-12.
(35) Lei, H. X., Wu, C., Liu, H. G., and Duan, Y. (2007) Folding free-energy landscape of villin headpiece subdomain from molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A. 104*, 4925−4930.
(36) Pande, V. S., Baker, I., Chapman, J., Elmer, S. P., Khaliq, S., Larson, S. M., Rhee, Y. M., Shirts, M. R., Snow, C. D., Sorin, E. J., and Zagrovic, B. (2003) Atomistic protein folding simulations on the submillisecond time scale using worldwide distributed computing. *Biopolymers 68*, 91−109.
(37) Piana, S., Lindorff-Larsen, K., and Shaw, D. E. (2012) Protein folding kinetics and thermodynamics from atomistic simulation. *Proc. Natl. Acad. Sci. U.S.A. 109*, 17845−17850.
(38) Voelz, V. A., Jager, M., Yao, S. H., Chen, Y. J., Zhu, L., Waldauer, S. A., Bowman, G. R., Friedrichs, M., Bakajin, O., Lapidus, L. J., Weiss, S., and Pande, V. S. (2012) Slow unfolded-state structuring in acyl-CoA binding protein folding revealed by simulation and experiment. *J. Am. Chem. Soc. 134*, 12565−12577.
(39) Juraszek, J., and Bolhuis, P. G. (2006) Sampling the multiple folding mechanisms of Trp-cage in explicit solvent. *Proc. Natl. Acad. Sci. U.S.A. 103*, 15859−15864.
(40) Lei, H. X., Wang, Z. X., Wu, C., and Duan, Y. (2009) Dual folding pathways of an alpha/beta protein from all-atom ab initio folding simulations. *J. Chem. Phys. 131*, 165105-1−165105-7.
(41) Ozkan, S. B., Wu, G. A., Chodera, J. D., and Dill, K. A. (2007) Protein folding by zipping and assembly. *Proc. Natl. Acad. Sci. U.S.A. 104*, 11987−11992.
(42) Paschek, D., Nymeyer, H., and Garcia, A. E. (2007) Replica exchange simulation of reversible folding/unfolding of the Trp-cage

miniprotein in explicit solvent: On the structure and possible role of internal water. *J. Struct. Biol. 157*, 524−533.
(43) Qiu, L. L., Pabit, S. A., Roitberg, A. E., and Hagen, S. J. (2002) Smaller and faster: The 20-residue Trp-cage protein folds in 4 μs. *J. Am. Chem. Soc. 124*, 12952−12953.
(44) Zhou, R. H. (2003) Trp-cage: Folding free energy landscape in explicit water. *Proc. Natl. Acad. Sci. U.S.A. 100*, 13280−13285.
(45) Baumketner, A., and Shea, J. E. (2007) The structure of the Alzheimer amyloid beta 10−35 peptide probed through replica-exchange molecular dynamics simulations in explicit solvent. *J. Mol. Biol. 366*, 275−285.
(46) Babu, M. M., Kriwacki, R. W., and Pappu, R. V. (2012) Versatility from protein disorder. *Science 337*, 1460−1461.
(47) Xu, D., and Zhang, Y. (2012) Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins: Struct., Funct., Bioinf. 80*, 1715−1735.
(48) Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc. 5*, 725−738.
(49) Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinf. 9*, 40-1−40-8.
(50) Yang, Y. D., Faraggi, E., Zhao, H. Y., and Zhou, Y. Q. (2011) Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native properties of templates. *Bioinformatics 27*, 2076−2082.
(51) Kinch, L., Shi, S. Y., Cong, Q., Cheng, H., Liao, Y. X., and Grishin, N. V. (2011) CASP9 assessment of free modeling target predictions. *Proteins: Struct., Funct., Bioinf. 79*, 59−73.
(52) Xu, D., Zhang, J., Roy, A., and Zhang, Y. (2011) Automated protein structure modeling in CASP9 by I-TASSER pipeline combined with QUARK-based ab initio folding and FG-MD-based structure refinement. *Proteins: Struct., Funct., Bioinf. 79*, 147−160.
(53) Yang, J. Y., Roy, A., and Zhang, Y. (2013) BioLiP: A semi-manually curated database for biologically relevant ligand-protein interactions. *Nucleic Acids Res. 41*, D1096−D1103.
(54) Zhang, Y. (2009) I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins: Struct., Funct., Bioinf. 77*, 100−113.
(55) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(56) Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S., and Onufriev, A. (2005) H++: A server for estimating pK_as and adding missing hydrogens to macromolecules. *Nucleic Acids Res. 33*, W368−W371.
(57) Leach, A. R. (2001) *Molecular Modelling: Principles and Applications*, 2nd ed., Prentice Hall, New York.
(58) Allen, M. P., and Tildesley, D. J. (1987) *Computer Simulation of Liquids*, Clarendon Press, Oxford.
(59) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys. 23*, 327−341.
(60) Kim, E.-H., Rensing, C., and McEvoy, M. M. (2010) Chaperone-mediated copper handling in the periplasm. *Nat. Prod. Rep. 27*, 711−719.
(61) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. . (2009) *Gaussian 09, Revision A.02*, Gaussian, Inc., Wallingford, CT.
(62) Zhao, Y., and Truhlar, D. G. (2006) A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys. 125*, 194101-1−194101-18.
(63) Hay, P. J., and Wadt, W. R. (1985) Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys. 82*, 299−310.
(64) Ditchfie, R., Hehre, W. J., and Pople, J. A. (1971) Self-consistent molecular-orbital methods. IX. Extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys. 54*, 724−728.
(65) Peters, M. B., Yang, Y., Wang, B., Fusti-Molnar, L., Weaver, M. N., and Merz, K. M. (2010) Structural survey of zinc-containing proteins and development of the zinc AMBER force field (ZAFF). *J. Chem. Theory Comput. 6*, 2935−2947.
(66) Case, D. A., Darden, T. A., T.E. Cheatham, I., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossvai, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2010) *AMBER 11*, University of California, San Francisco, CA.
(67) Zhao, Y., and Truhlar, D. G. (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc. 120*, 215−241.
(68) Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem. 97*, 10269−10280.
(69) Cieplak, P., Cornell, W. D., Bayly, C., and Kollman, P. A. (1995) Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem. 16*, 1357−1377.
(70) Pierce, L. C. T., Salomon-Ferrer, R., de Oliveira, C. A. F., McCammon, J. A., and Walker, R. C. (2012) Routine Access to millisecond time scale events with accelerated molecular dynamics. *J. Chem. Theory Comput. 8*, 2997−3002.
(71) Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Bioinf. 78*, 1950−1958.
(72) Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A. W., Kolossvai, I., Wong, K. F., Paesani, F., Vanicek, J., Wolf, R. M., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2012) *AMBER 12*, University of California, San Francisco, CA.
(73) York, D. M., Darden, T. A., and Pedersen, L. G. (1993) The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods. *J. Chem. Phys. 99*, 8345−8348.
(74) Shen, Y., and Bax, A. (2010) SPARTA plus: A modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J. Biomol. NMR 48*, 13−22.
(75) Heinig, M., and Frishman, D. (2004) STRIDE: A web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res. 32*, W500−W502.
(76) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics 14*, 33−38.

(77) Williams, T., and Kelley, C. (2010) Gnuplot 4.4.
(78) Duan, L. L., Mei, Y., Zhang, D. W., Zhang, Q. G., and Zhang, J. Z. H. (2010) Folding of a helix at room temperature is critically aided by electrostatic polarization of intraprotein hydrogen bonds. *J. Am. Chem. Soc. 132*, 11159−11164.
(79) Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H., and Dyer, R. B. (1996) Fast events in protein folding: Helix melting and formation in a small peptide. *Biochemistry 35*, 691−697.
(80) Uversky, V. N. (2013) A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci. 22*, 693−724.
(81) Chakravorty, D. K., Wang, B., Lee, C. W., Giedroc, D. P., and Merz, K. M. (2012) Simulations of allosteric motions in the zinc sensor CzrA. *J. Am. Chem. Soc. 134*, 3367−3376.
(82) Chakravorty, D. K., Wang, B., Ucisik, M. N., and Merz, K. M. (2011) Insight into the cation-π interaction at the metal binding site of the copper metallochaperone CusF. *J. Am. Chem. Soc. 133*, 19330−19333.
(83) Loftin, I. R., Franke, S., Blackburn, N. J., and McEvoy, M. M. (2007) Unusual Cu(I)/Ag(I) coordination of *Escherichia coli* CusF as revealed by atomic resolution crystallography and X-ray absorption spectroscopy. *Protein Sci. 16*, 2287−2293.
(84) Xue, Y., Davis, A. V., Balakrishnan, G., Stasser, J. P., Staehlin, B. M., Focia, P., Spiro, T. G., Penner-Hahn, J. E., and O'Halloran, T. V. (2008) Cu(I) recognition via cation-pi and methionine interactions in CusF. *Nat. Chem. Biol. 4*, 107−109.
(85) Cram, D. J. (1988) The design of molecular hosts, guests, and their complexes. *Science 240*, 760−767.
(86) Reyes-Caballero, H., Campanello, G. C., and Giedroc, D. P. (2011) Metalloregulatory proteins: Metal selectivity and allosteric switching. *Biophys. Chem. 156*, 103−114.
(87) Su, C. C., Long, F., Lei, H. T., Bolla, J. R., Do, S. V., Rajashankar, K. R., and Yu, E. W. (2012) Charged amino acids (R83, E567, D617, E625, R669, and K678) of CusA are required for metal ion transport in the Cus efflux system. *J. Mol. Biol. 422*, 429−441.