JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Hidden States within Disordered Regions of the CcdA Antitoxin Protein

Virginia M. Burger,[†] Alexandra Vandervelde,[‡,§] Jelle Hendrix,[∥,⊥] Albert Konijnenberg,[#] Frank Sobott,[#,∇,○] Remy Loris,[‡,§] and Collin M. Stultz*[,†,◆]

[†]Research Laboratory of Electronics, [◆]Electrical Engineering and Computer Science & Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

[‡]Structural Biology Brussels, Department of Biotechnology, Vrije Universiteit Brussel, B-1050 Brussels, Belgium

[§]Molecular Recognition Unit, Structural Biology Research Center, Vlaams Instituut voor Biotechnologie, B-1050 Brussels, Belgium

[∥]Laboratory for Photochemistry and Spectroscopy, Department of Chemistry, University of Leuven, B-3000 Leuven, Belgium

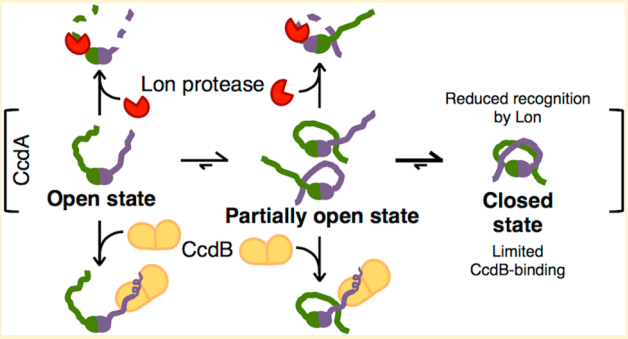
[⊥]Faculty of Medicine and Life Sciences and Biomedical Research Institute, Hasselt University, B-3500 Hasselt, Belgium

[#]Biomolecular & Analytical Mass Spectrometry, Department of Chemistry, University of Antwerp, B-2020 Antwerp, Belgium

[∇]Astbury Centre for Structural Molecular Biology, [○]School of Molecular and Cellular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom

Supporting Information

ABSTRACT: The bacterial toxin−antitoxin system CcdB−CcdA provides a mechanism for the control of cell death and quiescence. The antitoxin protein CcdA is a homodimer composed of two monomers that each contain a folded N-terminal region and an intrinsically disordered C-terminal arm. Binding of the intrinsically disordered C-terminal arm of CcdA to the toxin CcdB prevents CcdB from inhibiting DNA gyrase and thereby averts cell death. Accurate models of the unfolded state of the partially disordered CcdA antitoxin can therefore provide insight into general mechanisms whereby protein disorder regulates events that are crucial to cell survival. Previous structural studies were able to model only two of three distinct structural states, a closed state and an open state, that are adopted by the C-terminal arm of CcdA. Using a combination of free energy simulations, single-pair Förster resonance energy transfer experiments, and existing NMR data, we developed structural models for all three states of the protein. Contrary to prior studies, we find that CcdA samples a previously unknown state where only one of the disordered C-terminal arms makes extensive contacts with the folded N-terminal domain. Moreover, our data suggest that previously unobserved conformational states play a role in regulating antitoxin concentrations and the activity of CcdA's cognate toxin. These data demonstrate that intrinsic disorder in CcdA provides a mechanism for regulating cell fate.



INTRODUCTION

Bacterial toxin−antitoxin (TA) modules regulate cell death and quiescence in nearly all free-living bacteria.[1,2] In these systems, an antitoxin inhibits its cognate protein toxin, thereby preventing the toxin from disrupting essential cellular processes. When antitoxin levels fall, either through increased degradation or decreased production, the free toxin is activated to kill the cell or halt cell growth. A key feature of antitoxins is their lack of a stable structure, which causes them to have shorter lifetimes than the toxins they are inhibiting. TA modules were originally ascribed the role of plasmid maintenance, as plasmid-encoded TA modules ensure that only cells containing the corresponding plasmid (and able to produce antitoxin) are viable.[3,4] Additional roles in programmed cell death and persistence in response to stress have since been conjectured. For example, the ability of several cells within a colony to become dormant in response to life-threatening stresses allows those persistent cells to survive and repopulate the colony after the stress has ended.[5,6]

Bacteria cells can contain many coexisting TA modules; e.g., *Mycobacterium tuberculosis* contains almost 80 distinct TA modules and *Escherichia coli* contains over 30.[7,8] It has been hypothesized that bacteria that move between different environments have more TA systems, which allow precise tuning of a cell's response to environmental stress via the ratio of antitoxins to their cognate toxins.[9] Altogether, TA systems form an intricate and versatile system for regulating cell fate.

Received: November 3, 2016
Published: January 26, 2017

DOI: 10.1021/jacs.6b11450
J. Am. Chem. Soc. 2017, 139, 2693−2701

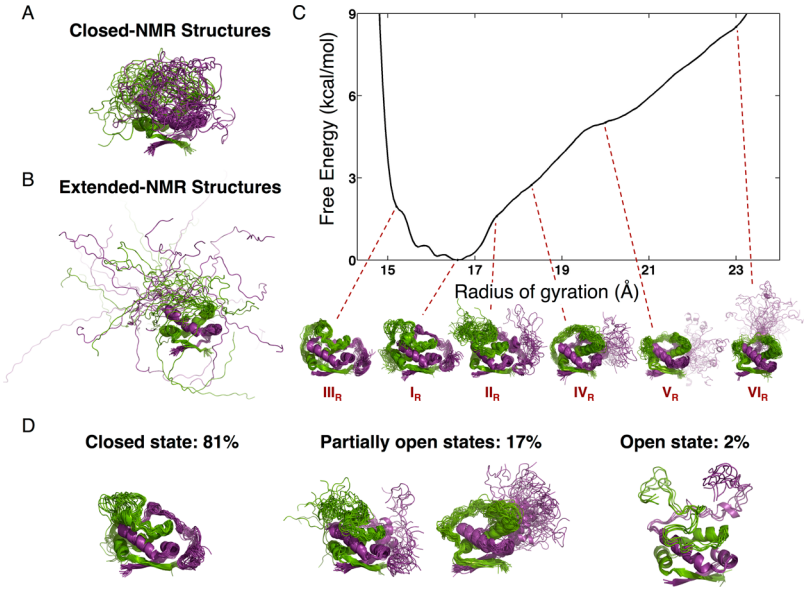


Figure 1. Models of CcdA. (A) Superimposed NMR models of the closed state of CcdA (PDB IDs 2h3c and 2adn[19]). For each conformation, monomer 1 (residues 1−72) is colored green and monomer 2 (residues 1′−72′) is colored purple. This color scheme is used for all subsequent figures. (B) Superimposed NMR models of the extended state of CcdA (PDB IDs 2h3a and 2adl[19]). In parts A and B, DNA has been removed from the 2h3c and 2h3a PDB models. (C) PMF for CcdA with respect to its radius of gyration. Representative conformations at radii of gyration 16.6 Å (I_R, the global free energy minimum), 17.5 Å (II_R), 15.2 Å (III_R), 18.3 Å (IV_R), 20.0 Å (V_R), and 23.0 Å (VI_R) are shown below the PMF, where the Roman numeral indicates increasing free energy. (D) Representative structures assigned to the closed (both C-terminal arms contact the NTD), partially open (exactly one C-terminal arm contacts the NTD), and open (neither C-terminal arm contacts the NTD) states are shown.

The *ccd* (control of cell death) gene system, carried on the *E. coli* F plasmid, codes for the CcdB toxin and the CcdA antitoxin.[10] *Ccd* regulates cell fate via several mechanisms: (1) Under conditions conducive to cell growth, CcdA binds to CcdB, forming a complex that prevents CcdB from binding and inhibiting DNA gyrase;[11−13] (2) CcdA can disrupt existing complexes between CcdB and DNA gyrase, thereby freeing DNA gyrase (in this manner CcdA can "rejuvenate" DNA gyrase molecules made defective by CcdB);[13−15] (3) CcdA and CcdA−CcdB complexes can bind the *ccd* promoter−operator region along the F plasmid to regulate transcription of both CcdA and CcdB.[16−18] Whether a bacterium lives, dies, or enters a quiescent state depends largely on the relative concentrations of both CcdA and CcdB, and the cell uses a variety of mechanisms to regulate the ratio of antitoxin to toxin.

The CcdA antitoxin exists as a homodimer formed from two 72-residue monomers. The dimer consists of a folded N-terminal domain (NTD) involving residues 1−40 of each monomer and two intrinsically disordered C-terminal arms formed by residues 41−72 of each monomer. The NTD has a ribbon-helix-helix fold that binds DNA,[19] and the C-terminal arms are responsible for binding CcdB; i.e., each C-terminal arm can bind and thereby inhibit one CcdB molecule.[15] The C-terminal arms also play a role in CcdA catabolism, as they contain the major cleavage sites for Lon protease, the enzyme responsible for CcdA degradation.[20] Structural studies of CcdA, and the disordered C-terminal arms in particular, can therefore provide insight into mechanisms that regulate cell fate.

In light of the important regulatory roles performed by the intrinsically disordered C-terminal arms of CcdA, our goal is to characterize the thermally accessible states in the native ensemble of apo CcdA. Due to their flexibility, intrinsically disordered regions, like the C-terminal arms of CcdA, present a challenge for traditional structure-determination methods. For intrinsically disordered proteins (IDPs) in general, recent computational and experimental advances have provided insights into a number of normal and pathological cellular processes.[21−26] In the present study, we build off of these advances to map the conformational free energy surface of CcdA, paying particular attention to the disordered C-terminal arms. Our free energy simulations are compared to experimental data from single-pair Förster resonance energy transfer (spFRET) and existing NMR measurements. We then propose how the conformations within CcdA's free energy surface enable it to both bind CcdB and regulate cleavage by Lon protease.

RESULTS AND DISCUSSION

CcdA Preferentially Adopts Closed and Partially Open States. CcdA fluctuates between several conformational states when not bound to CcdB.[19] Specifically, Madl et al. observed a folded NTD and multiple sets of resonances for several C-terminal residues in the $^{15}N-^{1}H$ HSQC spectrum for a R70K mutant of CcdA, suggesting that CcdA adopts several conformations on the NMR time scale. One set of resonances had chemical shifts that were in the range expected for a folded protein and are associated with long-range NOEs, whereas the chemical shifts of the other two sets of resonances were in the range expected for a random coil and were associated with only trivial and short-range NOEs. Using these data, Madl et al. constructed two structural models for CcdA. We refer to these models as the closed-NMR structures and the extended-NMR structures (Figure 1A,B). In the closed-NMR structures, both C-terminal arms fold back against the structured NTD (Figure 1A). To be consistent with the observed NOEs, the closed-NMR structures have contacts between C-terminal residues Ala66 and Asp67 and N-terminal residues Tyr20 and Val22 (from the same monomer) and residue Leu39′ (from the other monomer).[19] In the extended-NMR structures, the C-termini are modeled as having no contacts with the NTD and adopt an

DOI: 10.1021/jacs.6b11450
J. Am. Chem. Soc. 2017, 139, 2693−2701

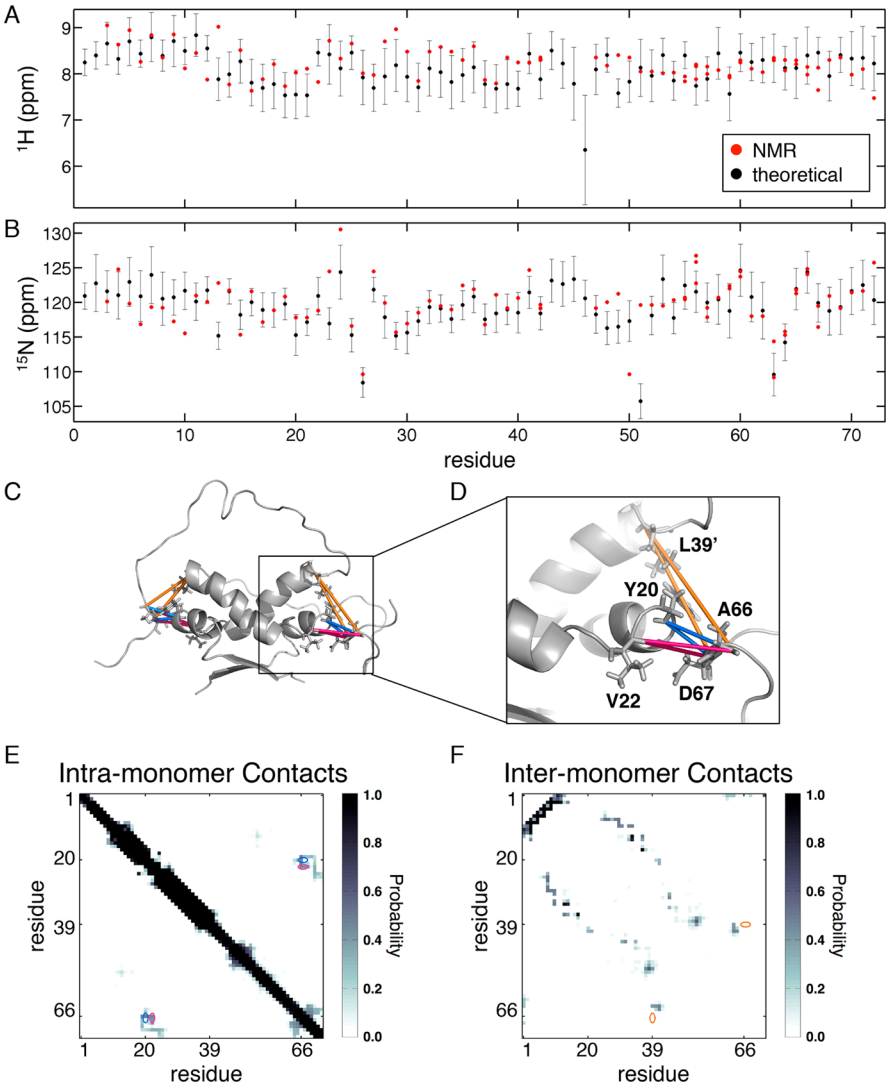


Figure 2. Comparison of calculated and measured NMR data. (A) Ensemble-averaged ^{1}H chemical shifts. The ensemble mean chemical shift is shown as a black dot for each residue, with error bars indicating standard deviation combined additively with the root-mean-squared error in the SHIFTX2 predictions. Red dots indicate the experimental ^{1}H NMR chemical shifts extracted from the HSQC spectrum for CcdAR70K.[19] (B) Ensemble-averaged ^{15}N chemical shifts, analogous to (A). (C) Contacts corresponding to the experimentally observed NOEs.[19] Residues involved in NOEs are shown as sticks and labeled. Orange lines indicate contacts with residue Leu39′, blue lines indicate contacts with residue Tyr20, and pink lines indicate contacts with residue Val22. (D) Close-up view of contacts corresponding to the experimentally observed NOEs[19] shown in panel C. (E) Intramonomer contact maps derived from the PMF for the closed state of CcdA. The residue pairs corresponding to detected NOEs are circled in the contact maps in panels E and F and colored as in panel D. (F) Intermonomer contact maps derived from the PMF for the closed state of CcdA, analogous to (E).

ensemble of extended conformations, corresponding to the random coil-like set of chemical shifts for which no long-range NOEs were observed (Figure 1B).[19] Although three sets of resonances were observed, the data were insufficient to build a NMR model for the third state.[19]

In this work, we employ a combined computational/experimental approach to study the conformational ensemble of CcdA. To determine the ensemble of structures sampled by CcdA in solution, we first calculated the free energy of CcdA as a function of the radius of gyration (R_g) using umbrella sampling coupled with explicit-solvent molecular dynamics simulations (see the Supporting Information). The resulting free energy surface has a well-defined global energy minimum and several shallow local minima (Figure 1C). Structures within the lowest energy state are compact in the sense that both C-termini arms (C1, residues Ala41−Trp72 monomer 1; and C2, residues Ala41′−Trp72′ monomer 2) are folded back against the structured NTD (Figure 1C, conformer I_R). Outside of the free energy minima, several less compact conformations are sampled (e.g., Figure 1C, conformers II_R and IV_R−VI_R).

We classified the structures sampled by apo CcdA into three representative states, in a manner similar to what was done in the aforementioned NMR study by Madl et al.[19] For this classification, we required a definition for a contact between a C-terminal arm and the structured N-terminal. Since C-terminal residues Ala66 and Asp67 are involved in long-range NOEs with residues in the NTD,[19] we focused on the region around these residues in defining a contact. Specifically, we say that a C-terminal arm contacts the NTD when the C_α atom of either residue Ala66 or Asp67 is within 8.5 Å of any N-terminal C_α atom (Figure S1). With this definition, we assigned sampled CcdA conformations to either (1) a closed state, in which both

DOI: 10.1021/jacs.6b11450
J. Am. Chem. Soc. 2017, 139, 2693−2701

C-terminal arms contact the structured NTD (this is similar to the closed-NMR structure), (2) an open state, in which neither C-terminal arm contacts the NTD (similar to the extended-NMR structure), or (3) a partially open state, in which exactly one C-terminal arm contacts the structured NTD. Using this classification, CcdA samples closed, partially open, and open states 81, 17, and 2% of the time, respectively (Figure 1D). We emphasize that the same trends are preserved over a wide range of cutoff distances (see Figure S1).

Free Energy Surface of CcdA Clarifies Underdetermined NMR Data. To determine how the theoretical free energy surface of CcdA compares to the aforementioned NMR studies, we computed ensemble-averaged chemical shifts for CcdA based on the free energy surface. Overall, the theoretical chemical shifts are in good agreement with the experimentally determined values (Figure 2A,B).

Additionally, because long-range NOEs were observed between C- and N-terminal residues, we examined the inter-residue contacts observed in the theoretical closed state models (Figure 2C−F). Specifically, for conformers belonging to the closed state, we determined average intra- and intermonomer contact maps using the Boltzmann probabilities derived from the free energy profile (also known as the potential of mean force or PMF) (see Experimental Procedures in the Supporting Information). The closed-state contact maps agree qualitatively with the contacts deduced from the experimental NOEs (circled in Figure 2E,F). That is, the general trend of C-terminal residues Ala66−Arg70 contacting N-terminal residues Ala19−Val22 is consistent with the experimentally observed NOEs, while the specific residues vary slightly from those residues for which strong NOEs were observed (residues Ala66 and Asp67 to residues Tyr20, Val22, and Leu39′; Figure 2C,D). We note that the NMR study was performed on a CcdAR70K mutant and, although this corresponds to a conservative mutation, arginine to lysine mutations can have destabilizing effects on protein structure.[27,28] Moreover, this substitution, which occurs near the C-terminal residues that are involved in long-range NOEs (i.e., residues A66 and A67), may affect the specific contacts involved in the closed state. Nevertheless, although the interaction with residue Leu39′ was not observed in the intermonomer contact map, contacts were detected between C-terminal residues Asn62−Ser64 of one monomer with N-terminal residues Leu39′−Asn42′ of the other monomer (Figure 2F).

Since the open state does not, by definition, have any contacts between residues 66 and 67 and residues in the NTD, no NOEs would be observed experimentally for this state. However, ensemble-averaged inter- and intramonomer contact maps for the partially open state are similar to the contact maps from the closed state (Figure 3). Thus, the closed and partially open states are not distinguishable through measurement of NOEs alone.

The ensemble average contact maps for the partially open state do not fully capture the range of partially open structures that the protein can adopt. Since there are two disordered C-terminal arms, and each one can contact the folded NTD, we can, in principle, distinguish between two partially open conformations. In the first conformation, C1 is open and only the C2 arm contacts the NTD (partially open substate A shown in Figure 4A), and in the second conformation, C2 is open and only the C1 arm contacts the NTD (partially open substate B shown in Figure 4B). In both substates, however, the C-terminal arm that contacts the NTD has interactions with the

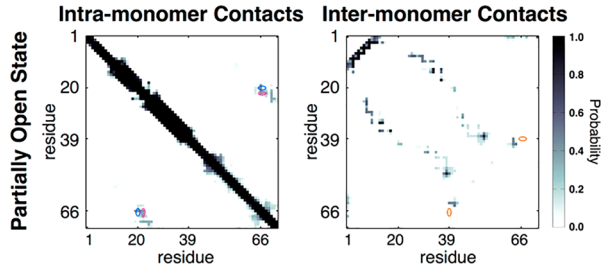


Figure 3. Intramonomer (left) and intermonomer (right) contact maps derived from the PMF for the partially open state of CcdA. The residue pairs corresponding to detected NOEs are circled in each contact map and colored as in Figure 2C−F.

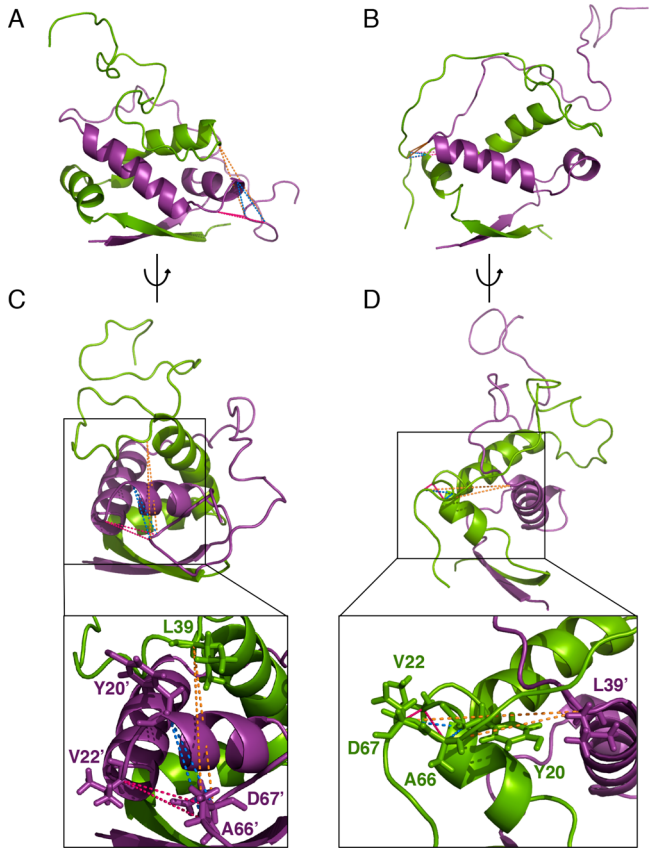


Figure 4. Representative structures from partially open substates. (A) Arm C1 from monomer 1 (green) is open and the arm C2 from monomer 2 (purple) contacts the N-terminal domain; (B) arm C1 from monomer 1 contacts the N-terminal domain. (C, D) Alternate views of structures in A and B. Side chains of residues for which NOEs were observed are shown as sticks in the bottom panels. Inter-residue contacts between residues that have NOEs are shown as dotted lines between C_α atoms in A−D. The intermonomer distances for the partially open substates are similar to those in the NMR-closed state structures (Table S2).

NTD that are similar to what is observed in the closed-NMR structures (Figure 4C and D).

Lastly, we note that the closed state has an ensemble-averaged R_g of 16.4 Å, which is equal to the average R_g of the closed-NMR structures. By contrast, the extended-NMR structures have an average R_g of 26.9 Å, which corresponds to a highly unfavorable free energy (greater than 12 kcal/mol) according to our free energy surface (Figure 1C). Since long-range NOEs were not observed for residues in the C-terminal

DOI: 10.1021/jacs.6b11450
J. Am. Chem. Soc. 2017, 139, 2693−2701

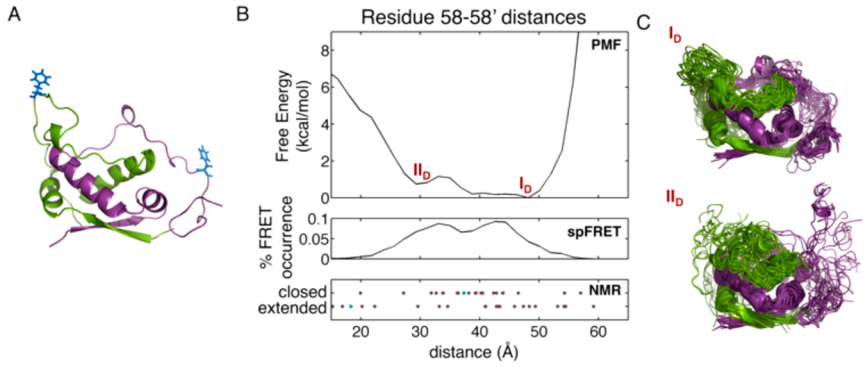


Figure 5. Intermonomer Phe58−Phe58′ distances sampled by CcdA. (A) Representative structure of CcdA with residues Phe58 (selected for mutation and labeling for FRET) on each monomer shown as blue sticks. (B) Top panel: the PMF of CcdA transformed onto an axis describing the intramolecular distance between residues Phe58 and Phe58′; middle panel: distribution of distances between the fluorophores tagged to F58C detected by FRET measurements; bottom panel: intra- molecular distance between residues Phe58 and Phe58′ computed from the closed-NMR (top) and extended-NMR (bottom) structure models of CcdA. (C) Representative conformers from the minimum energy wells (I_D, II_D) in panel B.

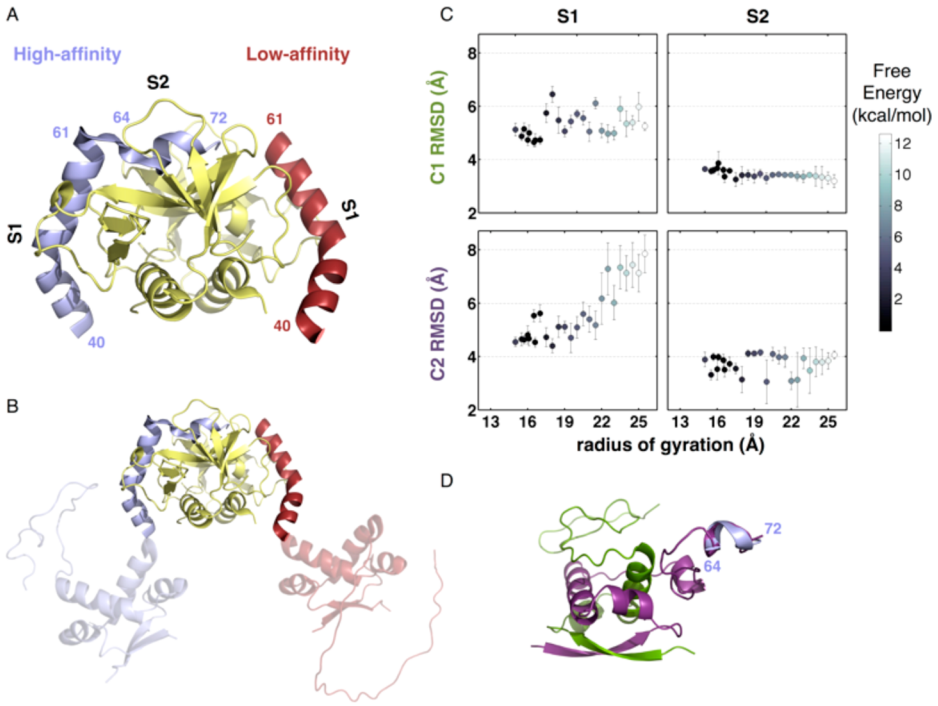


Figure 6. Sampling of bound structure by free CcdA. (A) Crystal structure of the high-affinity (light blue) and low-affinity (rust) conformations of the CcdA residue 37−72 peptide bound to CcdB (yellow).[15] (B) Model of CcdB bound to two C-terminal arms from two different CcdA molecules. (C) Average backbone RMSDs for residues CcdA 40−61 and residues 64−72 to each of the two structures within the high-affinity bound conformation (blue, panel A) are shown for the C-terminal of monomer 1 (C1, top panel) and monomer 2 (C2, bottom panel). Error bars denote standard deviation, and points are colored according to their free energy, with darker colors indicating lower free energy. (D) CcdA conformation with lowest RMSD to S2 for an umbrella window with free energy less than 3 kcal/mol. This structure has an RMSD of 1.6 Å to S2 and is from the simulation window with R_g 1.80 Å, corresponding to a free energy of 2.36 kcal/mol. CcdA monomer 1 is shown in green, CcdA monomer 2 is shown in purple, and the S2 conformation is shown superimposed on the full molecule in blue.

region of the extended-NMR structures, the corresponding C-terminal residues were modeled as being extended and solvent exposed, a common assumption when modeling random coil structures. However, our data argue that such extended structures are rarely sampled.

Insights into the Structure of CcdA using spFRET. To assess the range of conformations accessible to CcdA in solution, we used spFRET to quantify intermonomer distances sampled by CcdA.[29] CcdA was expressed and purified as described in the Supporting Information, and native mass spectrometry experiments verified that the protein is primarily dimeric.[30,31] The molecular weight of the CcdA dimer measured by native mass spectrometry was 16745.5 Da, which agrees well with its theoretical molecular weight of 16744.7 Da (Figure S2 and Table S1).

For the spFRET experiments, we added donor and acceptor fluorophores to the C-terminal arms of the protein. As there is no cysteine naturally present in the C-terminus of CcdA to which a fluorophore could be attached, we first created a mutant protein in which one residue from each monomer was mutated to cysteine. We selected residue F58 for this mutation due to its central location in the disordered C-terminal domain,

DOI: 10.1021/jacs.6b11450
J. Am. Chem. Soc. 2017, 139, 2693−2701

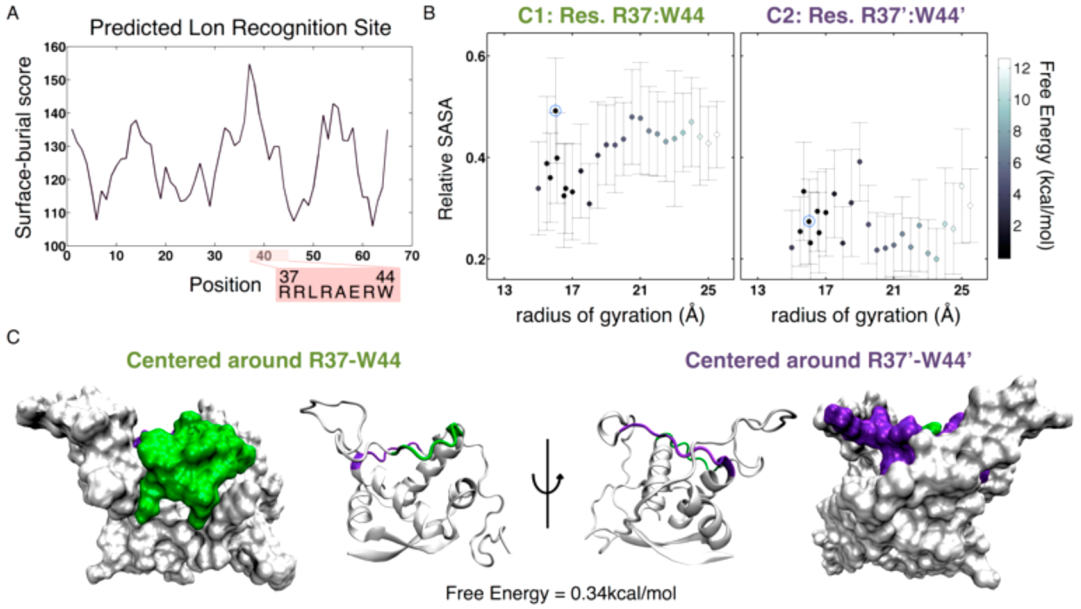


Figure 7. Predicted Lon recognition site is transiently exposed in CcdA. (A) Surface burial score (a sequence-based predictor for Lon recognition sites) computed for each consecutive-residue window in the CcdA monomer. The peak at position 37 corresponds to the highlighted 8-residue segment that begins at residue 37. (B) Mean relative SASA in residues R37−W44 (left) and R37′−W44′ (right) for each simulation window, with error bars denoting standard deviation and colors corresponding to the free energy of each window. The lowest free energy window for which the predicted Lon recognition site is exposed is circled in blue. (C) Two views of a representative conformation of CcdA from the circled window in B. The potential Lon recognition site is colored green on monomer 1 and purple on monomer 2, and the remainder of CcdA is drawn in white.

in addition to the assumption that placing a fluorophore in a position that already accommodates a residue with a large side chain would minimize any perturbation to the structure (Figure 5A). The measured FRET efficiencies for the mutated CcdA protein were then transformed to distances using parameters as described in the Supporting Information. Low-FRET states correspond to relatively long intermonomer distances, and high FRET states correspond to relatively short intermonomer distances.

spFRET analysis of F58C CcdA, with donor and acceptor fluorophores attached to residue C58, shows a bimodal distance distribution (Figure 5B, middle panel). For direct comparison with the free energy profile of CcdA, we transformed the PMF to an axis describing the intermonomer distances between the C_α atoms of residues F58−F58′ (Figures 5B, top panel, and S3). The low-FRET state samples distances near 45 Å, around which we also see a broad local energy minimum in the PMF (I_D, Figure 5B). Representative structures from this local energy minimum correspond to closed conformations (Figure 5C). By contrast, the high-FRET state samples distances near 30 Å, which corresponds to another local energy minimum in the transformed PMF of CcdA (II_D, Figure 5B), and this state is populated by both partially open and closed conformations (Figure 5C). Altogether, the spFRET data are wholly explained by the free energy simulations that demonstrate that CcdA preferentially adopts closed and partially open states. For comparison, we also computed the intermonomer F58−F58′ distances within the previously constructed NMR structures (Figures 5B, bottom panel, and 1A,B). While both the closed-NMR and extended-NMR structure models sample a range of intermonomer distances, the bimodality apparent from spFRET distance distribution is not apparent in the NMR models.

CcdB-Binding-Competent Structures Are Enriched in Apo CcdA's Partially Open State. The CcdB toxin has two partially overlapping bindings sites for CcdA that can simultaneously bind two C-terminal arms from distinct CcdA molecules (Figure 6A,B).[15] The binding sites have different affinities for CcdA.[15] In the low-affinity binding site, a C-terminal arm from CcdA binds CcdB with residues R40−M61 and forms an α helix, which we refer to as S1, while residues N62−W72 remain disordered. In the high-affinity binding site, a C-terminal arm from a different CcdA dimer binds CcdB through both an extended S1 (R40−G63) and a second short structure with a turn involving residues S64−W72, which we call S2.[15]

We explored whether either C-terminal arm of apo CcdA adopts conformations similar to S1 or S2. Average backbone root-mean-square deviation (RMSD) from S1 remains above 4 Å for all radii of gyration (Figure 6C), indicating that neither C-terminal arm adopts conformations that are similar to S1. However, C2 (the arm that more frequently adopted open conformations during the free energy simulations) frequently samples structures within 3 Å of S2 (Figure 6C). A representative structure with low RMSD to S2 is shown in Figure 6D. This representative structure has a R_g of approximately 18 Å. Conformers within this simulation window have a mean RMSD to S2 of 3.1 ± 0.5 Å and an associated relative free energy of 2.36 kcal/mol, indicating that 0.4% of CcdA molecules in solution adopt conformations with this R_g (and the corresponding range of RMSDs to S2) at any given time.

The fact that the open C-terminal arm of CcdA samples states that are similar to S2 raises interesting questions about the mechanism of CcdA binding to CcdB. Since it is known that a peptide composed of only residues F65−W72 is sufficient to partially restore gyrase activity after inhibition of gyrase by CcdB, it has been postulated that binding of these residues, which form S2 when bound to CcdA, triggers an allosteric mechanism that releases CcdB from gyrase.[15] Our data suggest that residues S64−W72 in CcdA's open arm

DOI: 10.1021/jacs.6b11450
J. Am. Chem. Soc. 2017, 139, 2693−2701

sample a structure similar to S2 and, consequently, that such structures are preformed to bind CcdB.

Specific CcdA Conformations Are Recognized by Lon Protease. CcdA is cleaved by ATP-dependent Lon protease at specific known sites.[20,32] However, Lon-mediated cleavage occurs via a number of steps, beginning with substrate recognition, and CcdA's Lon recognition sites are not known. It is known, however, that Lon protease recognizes clusters of hydrophobic residues within sequences as short as 7−20 residues and that these recognition sites have certain hallmarks, such as aromatic residues and high surface burial scores (e.g., greater than 140).[33] We thus screened the CcdA sequence for the 7−20 residue long subsequence with the highest surface burial score (Figures 7A and S4A).[34,35] Residues R37−W44 (RRLRAERW) had a surface burial score of 154.8, suggesting that Lon may recognize this sequence.

Our data suggest that this potential Lon recognition site (residues R37−W44) is significantly more solvent exposed in monomer 1 than in monomer 2 (the monomer whose C-terminal more frequently adopted open states during the free energy simulations) (Figure 7B). Closure of a C-terminal arm involves formation of contacts between residues N62−S64 of that arm with residues L39′−E42′ from the other monomer (Figure 2F). Therefore, opening of one C-terminal arm necessarily exposes the predicted Lon recognition region on the other monomer. This is illustrated in Figure 7C, which shows the solvent-accessible surface of a representative conformation of CcdA from the simulation window with free energy of 0.34 kcal/mol. C1 partially blocks the predicted recognition site on monomer 2, whereas the position of C2 exposes the predicted recognition site on monomer 1 to solvent.

Since CcdB is known to protect CcdA from cleavage, we also computed the relative SASA for the predicted recognition site within the CcdB-bound structure (shown in Figure 6A) to determine whether these residues are buried when bound to CcdB. The mean relative SASA of residues R40−W44 (for which structural information was available in both the low-affinity and high-affinity bound conformers) was 0.27 ± 0.13. Notably, aromatic residues are key components of Lon recognition sites, and W44 had a relative SASA of 0.11 in the low-affinity bound conformer and 0.22 in the high-affinity bound conformer. In contrast, for the window with a free energy of 0.34 kcal/mol (for which a representative conformation is shown in Figure 7C), W44 has a relative SASA of 0.55 ± 0.08 in monomer 1 (exposed site) and 0.30 ± 0.12 in monomer 2. The low relative SASA in the predicted Lon recognition region of the bound conformation is consistent with the significant reduction in cleavage of CcdA when bound to CcdB.

CONCLUSIONS

Results from the free energy simulations coupled with data from spFRET demonstrate that unbound CcdA exists as an equilibrium distribution of several states, where the dominant state is a closed conformation. In the closed state, both C-terminal arms fold back against the folded NTD, forming a relatively compact state. In the partially open state, only one of the two C-terminal arms contacts the NTD, and in the open state, which is infrequently sampled, neither C-terminal arm contacts the NTD. Our data also reveal that the partially open state contains conformations in which residues S64−W72 are preformed for binding the CcdB toxin. Additionally, some partially open conformations expose predicted Lon recognition sites to solvent. Hence, the ability to adopt partially open states enables CcdA to bind its cognate toxin CcdB and thereby disrupt complexes between CcdB and DNA gyrase, as well as to be recognized by Lon protease, which is responsible for CcdA degradation.

Previous NMR studies of unbound CcdA suggest that the C-terminal arms sample more than one state on the NMR time scale. Although $^{1}H-^{15}N$ HSQC spectra indicated that CcdA can adopt three distinct conformations in solution, only two could be reliably modeled from the NMR data: a closed structure similar to the dominant state in our free energy landscape and an extended structure where the C-terminal arms adopt extended conformations that do not contact the folded NTD. Since CcdA is a homodimer, it is difficult to distinguish between chemical shifts of the two monomers in a HSQC experiment. Indeed, a single resonance for a given nucleus is an average of the resonances for the corresponding nuclei in each monomer of the homodimer. Even when multiple resonances for a given nucleus are observed, it is difficult to know whether this reflects the existence of different symmetric homodimer structures or the presence of an asymmetric homodimer conformation. Therefore, to create structural models from such data, simplifying assumptions are needed, e.g., that the homodimer itself has a symmetric structure. The result is that by using the set of folded-like chemical shifts and NOEs observed by NMR for CcdA, it is not possible to distinguish between the closed state and the partially open state (Figures 2E,F and 3). The existence of multiple conformational states adds to the challenge: the interproton NOEs of proteins that sample multiple conformations will be biased toward conformations in which the protons of interest are closest together.[36] Apart from the challenge of interpreting the NMR data, the NMR experiments on CcdA used a CcdAR70K mutant in place of wild-type CcdA to reduce proteolytic cleavage of the protein. Since residue R70 is close to residues in the C-terminal arm (residues A66 and D67) that form NOEs with residues in the N-terminal region, this mutation, in addition to the conditions of the NMR experiment, may affect the structure in the region where contacts are formed between the folded N-terminal domain and the disordered C-terminal arms.

Our data demonstrate that the previously measured NMR observables are consistent with a model where the protein can adopt a closed state, an open state, and a partially open state. We used spFRET to study the structure of the protein in solution and compare these results to insights obtained via our free energy simulations.[29] Agreement between the distance distributions obtained from spFRET and the corresponding ensemble-averaged distributions from the free energy landscape demonstrate that the free energy landscape better models the native ensemble of CcdA than the structures derived from the NMR data. These observations highlight that care needs to be taken when interpreting NMR experiments on disordered homodimeric proteins. In such cases we believe that detailed calculations of the underlying free energy surface should play an important role in the interpretation of experimental results.

The equilibrium distribution between CcdA's three states may provide an additional mechanism for regulating cell death and quiescence in that each state has properties that influence their interactions with CcdB and Lon protease. For example, the partially open state is enriched in structures that can bind CcdB (Figure 6). At the same time, opening one arm exposes a

DOI: 10.1021/jacs.6b11450
J. Am. Chem. Soc. 2017, 139, 2693−2701

potential recognition site for Lon protease on the other monomer; i.e., our data argue that both the partially open and rarely sampled open states expose sites that can be recognized by Lon protease. By contrast, as the closed state contains few structures capable of binding CcdB or exposing Lon recognition sites, sampling closed conformations provides a mechanism for limiting CcdA's interaction with partner proteins.

The existence of a dominant closed state explains the slow degradation rate of CcdA by Lon compared to that of many other Lon substrates.[32] For example, the rate of CcdA degradation by Lon protease can be described by $\nu = V_{max}[S_o]/K_M$, where V_{max} is the maximal reaction rate, K_M is the Michaelis−Menten constant, and $[S_o]$ is the total concentration of conformations that have exposed Lon protease recognition sites. This simplified version of the Michaelis−Menten relation is valid when the substrate concentration is low.[37] If $[S]$ denotes the total concentration of CcdA and $[S_c]$ denotes the concentration of closed conformers, then the reaction rate, as a function of closed states, is $\nu = V_{max}([S] - [S_c])/K_M$. The existence of a closed state may therefore slow cleavage of CcdA by Lon, thereby increasing the resilience of cells to small fluctuations in Lon protease activity (e.g., in response to environmental changes), as speculated by Van Melderen et al.[32]

Conditions that cause a shift in the equilibrium distribution would affect the proportion of CcdA conformations competent to bind CcdB or be cleaved by Lon, thereby promoting either normal cell growth or cell death/quiescence. For example, if the equilibrium distribution were shifted to more strongly favor the closed state, the number of binding-competent conformations for CcdB would decrease. This would in turn increase the number of CcdB molecules free to bind gyrase, resulting in halted cell growth. By contrast, if the equilibrium distribution were shifted to favor the partially open state, the number of binding-competent conformations for CcdB would increase, promoting the normal cell cycle. However, under conditions that favor increased Lon activity, such as heat stress, the increased percentage of partially open states would also result in more rapid degradation of CcdA by Lon protease. Overall, the balance between promoting cell viability and cell death/quiescence will depend on which conformational states are favored under a given set of cellular conditions. The equilibrium distribution between CcdA's distinct conformational states therefore adds a level of complexity that can be used to fine tune bacterial response to stress. Approaches that strive to modify the relative amounts of these conformational states form a platform for the design of new antimicrobial agents.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b11450.

Experimental methods; distance cutoff to define the closed state; native mass spectrometry of CcdA; distribution of F58−F58′ distances within each umbrella window; solvent exposure of the R37:44−Y14′:Y20′ region; experimental and theoretical molecular weights for oligomeric states of CcdA; and ensemble-averaged C_α distances (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*cmstultz@mit.edu

ORCID

Collin M. Stultz: 0000-0002-3415-242X

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation Postdoctoral Research Fellowship in Biology under grant no. 1309247 to V.M.B. and an Early Postdoctoral Mobility Fellowship from the Swiss National Science Foundation (SNSF) to V.M.B. A.V. received an FWO predoctoral fellowship (grant no. FWOTM63/11E5715N7). J.H. was funded by the Research Foundation Flanders (FWO Vlaanderen), project no. G0B4915N. We thank Sarah Haesaerts for her assistance in expression and purification of CcdA, Prof. Don C. Lamb (Ludwig-Maximilians Universität, Munich, Germany) for sharing the PAM software, Dipl. Chem., Anders Barth (LMU Munich) for helping in the programming of the burst analysis software, Niels Vandenberk for quantum yield measurements, and Johan Hofkens for the use of his microscopy facilities and financial support. R.L. received funding from the Fonds voor Wetenschappelijk Onderzoek Vlaanderen (FWO research grant no. G.0135.15N). The Synapt G2 instrument is funded by a grant from the Hercules Foundation−Flanders to Prof. Frank Sobott.

■ REFERENCES

(1) Hayes, F. *Science* **2003**, *301*, 1496.
(2) Pandey, D. P.; Gerdes, K. *Nucleic Acids Res.* **2005**, *33*, 966.
(3) Gerdes, K.; Rasmussen, P. B.; Molin, S. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 3116.
(4) Jaffe, A.; Ogura, T.; Hiraga, S. *J. Bacteriol.* **1985**, *163*, 841.
(5) Page, R.; Peti, W. *Nat. Chem. Biol.* **2016**, *12*, 208.
(6) Chan, W. T.; Espinosa, M.; Yeo, C. C. *Front Mol. Biosci* **2016**, *3*, 9.
(7) Sala, A.; Bordes, P.; Genevaux, P. *Toxins* **2014**, *6*, 1002.
(8) Yamaguchi, Y.; Park, J. H.; Inouye, M. *Annu. Rev. Genet.* **2011**, *45*, 61.
(9) Gelens, L.; Hill, L.; Vandervelde, A.; Danckaert, J.; Loris, R. *PLoS Comput. Biol.* **2013**, *9*, e1003190.
(10) Ogura, T.; Hiraga, S. *Cell* **1983**, *32*, 351.
(11) Bernard, P.; Couturier, M. *J. Mol. Biol.* **1992**, *226*, 735.
(12) Bahassi, E. M.; O'Dea, M. H.; Allali, N.; Messens, J.; Gellert, M.; Couturier, M. *J. Biol. Chem.* **1999**, *274*, 10936.
(13) Maki, S.; Takiguchi, S.; Horiuchi, T.; Sekimizu, K.; Miki, T. *J. Mol. Biol.* **1996**, *256*, 473.
(14) Bernard, P.; Kezdy, K. E.; Van Melderen, L.; Steyaert, J.; Wyns, L.; Pato, M. L.; Higgins, P. N.; Couturier, M. *J. Mol. Biol.* **1993**, *234*, 534.
(15) De Jonge, N.; Garcia-Pino, A.; Buts, L.; Haesaerts, S.; Charlier, D.; Zangger, K.; Wyns, L.; De Greve, H.; Loris, R. *Mol. Cell* **2009**, *35*, 154.
(16) Tam, J. E.; Kline, B. C. *Mol. Gen. Genet.* **1989**, *219*, 26.
(17) Afif, H.; Allali, N.; Couturier, M.; Van Melderen, L. *Mol. Microbiol.* **2001**, *41*, 73.
(18) Dao-Thi, M. H.; Charlier, D.; Loris, R.; Maes, D.; Messens, J.; Wyns, L.; Backmann, J. *J. Biol. Chem.* **2002**, *277*, 3733.
(19) Madl, T.; Van Melderen, L.; Mine, N.; Respondek, M.; Oberer, M.; Keller, W.; Khatai, L.; Zangger, K. *J. Mol. Biol.* **2006**, *364*, 170.
(20) Van Melderen, L.; Thi, M. H. D.; Lecchi, P.; Gottesman, S.; Couturier, M.; Maurizi, M. R. *J. Biol. Chem.* **1996**, *271*, 27730.

DOI: 10.1021/jacs.6b11450
J. Am. Chem. Soc. 2017, 139, 2693−2701

(21) D'Urzo, A.; Konijnenberg, A.; Rossetti, G.; Habchi, J.; Li, J.; Carloni, P.; Sobott, F.; Longhi, S.; Grandori, R. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 472.
(22) Ferreon, A. C.; Moran, C. R.; Gambin, Y.; Deniz, A. A. *Methods Enzymol.* **2010**, *472*, 179.
(23) Milles, S.; Lemke, E. A. *Biophys. J.* **2011**, *101*, 1710.
(24) Ullman, O.; Fisher, C. K.; Stultz, C. M. *J. Am. Chem. Soc.* **2011**, *133*, 19536.
(25) Burger, V. M.; Gurry, T.; Stultz, C. M. *Polymers* **2014**, *6*, 2684.
(26) Rauscher, S.; Pomes, R. *Biochem. Cell Biol.* **2010**, *88*, 269.
(27) Sokalingam, S.; Raghunathan, G.; Soundrarajan, N.; Lee, S.-G. *PLoS One* **2012**, *7*, e40410.
(28) Mrabet, N. T.; Van den Broeck, A.; Van den Brande, I.; Stanssens, P.; Laroche, Y.; Lambeir, A. M.; Matthijssens, G.; Jenkins, J.; Chiadmi, M. *Biochemistry* **1992**, *31*, 2239.
(29) Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 6264.
(30) Konijnenberg, A.; Butterer, A.; Sobott, F. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 1239.
(31) Vahidi, S.; Stocks, B. B.; Konermann, L. *Anal. Chem.* **2013**, *85*, 10471.
(32) Van Melderen, L.; Bernard, P.; Couturier, M. *Mol. Microbiol.* **1994**, *11*, 1151.
(33) Gur, E.; Sauer, R. T. *Genes Dev.* **2008**, *22*, 2267.
(34) Rose, G. D.; Geselowitz, A. R.; Lesser, G. J.; Lee, R. H.; Zehfus, M. H. *Science* **1985**, *229*, 834.
(35) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. In *The Proteomics Protocols Handbook*; Walker, J. M., Ed.; Humana Press: Totowa, NJ, 2005.
(36) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, 2nd ed.; Wiley: New York, 2000.
(37) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*; W.H. Freeman: Basingstoke, 2012.

DOI: 10.1021/jacs.6b11450
J. Am. Chem. Soc. 2017, 139, 2693−2701