Biochemistry

Article

pubs.acs.org/biochemistry

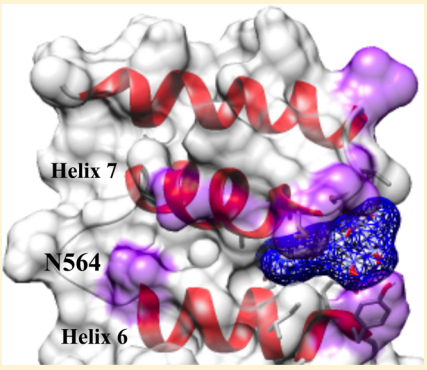
Identification of a Ligand-Binding Site on the *Staphylococcus aureus* DnaG Primase C-Terminal Domain

Jonathan Catazaro, Jessica Periago, Matthew D. Shortridge, Bradley Worley, Andrew Kirchner, Robert Powers,* and Mark A. Griep*

Department of Chemistry, University of Nebraska—Lincoln, Lincoln, Nebraska 68588-0304, United States

Supporting Information

ABSTRACT: The interface between the DnaG primase C-terminal domain (CTD) and the N-terminal domain of DnaB helicase is essential for bacterial DNA replication because it allows coordinated priming of DNA synthesis at the replication fork while the DNA is being unwound. Because these two proteins are conserved in all bacteria and distinct from those in eukaryotes, their interface is an attractive antibiotic target. To learn more about this interface, we determined the solution structure and dynamics of the DnaG primase CTD from *Staphylococcus aureus*, a medically important bacterial species. Comparison with the known primase CTD structures shows there are two biologically relevant conformations, an open conformation that likely binds to DnaB helicase and a closed conformation that does not. The *S. aureus* primase CTD is in the closed conformation, but nuclear magnetic resonance (NMR) dynamic studies indicate there is considerable movement in the linker between the two subdomains and that N564 is the most dynamic residue within the linker. A high-throughput NMR ligand affinity screen identified potential binding compounds, among which were acycloguanosine and myricetin. Although the affinity for these compounds and adenosine was in the millimolar range, all three bind to a common pocket that is present only on the closed conformation of the CTD. This binding pocket is at the opposite end of helices 6 and 7 from N564, the key hinge residue. The identification of this binding pocket should allow the development of stronger-binding ligands that can prevent formation of the CTD open conformation that binds to DnaB helicase.



The initiation of DNA synthesis in living organisms and many large viruses results from the coordinated interaction of two enzymes, primase and helicase.[1,2] In bacteria, DnaG primase is the specialized DNA-dependent RNA polymerase that synthesizes short oligoribonucleotide polymers called primers. DNA polymerase, an enzyme that lacks the ability to initiate chain synthesis, elongates the primers. DnaG primase is stimulated by the homohexameric DnaB helicase (note that the name DnaB helicase is used here for the protein that is actually named DnaC helicase in *Staphylococcus aureus* and *Bacillus subtilis* because, in all other bacteria, DnaB is the name of the replicative helicase and DnaC is the name of the helicase loading enzyme). DnaB helicase dissociates the two strands of duplex DNA during DNA replication while hydrolyzing ATP and travels processively in the 5′−3′ direction along the single-stranded lagging template toward the replication fork, communicating allosterically with the multisubunit replicative DNA polymerase. This action keeps it in the proximity of the replication fork and ensures the primers are synthesized on the exposed single-stranded DNA closest to the replication fork. Because primase activity is weak, the stimulation by DnaB causes primase to synthesize primers only at the replication fork when and where they are needed.[3,4] On the other hand, there are phylogenetic differences concerning the interaction of primase and helicase. For instance, the helicases from *Escherichia coli* and *Clostridium difficile* broaden the initiation specificity of their cognate primases,[5,6] whereas this is not observed in the *S. aureus* system.[7] *S. aureus* helicase is a much weaker stimulator of its primase than in the *E. coli* system,[7] and *C. difficile* helicase is active only in the presence of its primase.[6] Nevertheless, the DnaG−DnaB interaction is an attractive antibiotic target because it is conserved in bacteria, essential for DNA replication, and distinctly different from that of viruses, archaea, and eukaryotes.[8−10]

DnaG primase is composed of three functional domains: the N-terminal zinc-binding domain (ZBD, Pfam entry PF08275) that is responsible for DNA binding specificity, the RNA polymerase domain (RPD, Pfam entry PF01751) that is responsible for enzymatic activity, and the C-terminal domain (CTD, Pfam entry PF10410) that is responsible for the interaction with the DnaB helicase.[11−13] The 110-residue ZBD contains the highest percentage of conserved residues and structural conservation, whereas the 140-residue CTD has the lowest level of sequence conservation and the most structural diversity (Table S1). Of these domains, the ZBD is unique to

Received: December 16, 2016
Revised: January 11, 2017
Published: January 26, 2017

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

bacterial primases, the magnesium-binding residues of the RPD are somewhat similar to the magnesium-binding residues of topoisomerases and some DNA-hydrolyzing enzymes, while the CTD is similar only to the N-terminal domain of DnaB helicases, the domain to which it binds.[14−17] Currently, no high-resolution structure of a full, intact bacterial primase has been determined, but structures of all three individual domains are available. The primase CTD is the domain with the most experimental structures.[9,18−21] There are X-ray crystal structures and nuclear magnetic resonance (NMR) solution structures (Table 1) of *S. aureus* [Protein Data Bank (PDB) entry 2LZN, this work], *E. coli* (PDB entries 2HAJ and 1T3W), *Geobacillus stearothermophilus* (PDB entries 1Z8S and 2R6A), and *Helicobacter pylori* (PDB entry 4EHS).

Table 1. Primase CTD Structures Used for Analysis

species	PDB entry, chain	ref
S. aureus	2LZN, 14	this work
G. stearothermophilus	1Z8S, 5	9
E. coli	2HAJ, 19	20
G. stearothermophilus	2R6A, C	19
H. pylori	4EHS, A; 4EHS, B	22
E. coli	1T3W, A; 1T3W, B	21

In addition, the ZBD−RPD domain pair from *Aquifex aeolicus* has been experimentally determined.[14] It has been proposed that the conformations of the two *E. coli* structures in 1T3W may not be biologically relevant because of crystallization conditions and packing effects.[20]

To develop the DnaG−DnaB interface as an antibiotic target, we determined the solution structure of the DnaG CTD from *S. aureus*. This organism was chosen because there is an urgent need for new antibiotic targets due to the rapid rise in antibiotic resistance.[23] A comparison to known structures from other bacteria indicated that the *S. aureus* CTD was composed of two subdomains. The CTD from Firmicutes and Proteobacteria may adopt similar conformations when bound to helicase but may adopt different free conformations. A study of the structural dynamics of the free *S. aureus* CTD confirmed that the junction between the two subdomains is the focal point for the greatest movement. Also, one of the subdomains is significantly more ordered than the other. The structure was used to identify three small compounds that bound to the close conformation, which is likely to weaken the primase−helicase interaction.

EXPERIMENTAL PROCEDURES

Chemicals. Dimethyl sulfoxide-d_6 (99.9% D) and deuterium oxide (99.9% D) were obtained from Aldrich (Milwaukee, WI). 3-(Trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt (98% D) was purchased from Cambridge Isotopes (Andover, MA). Potassium phosphate dibasic salt (anhydrous, 99.1% pure) and monobasic salt (crystal, 99.8% pure) were purchased from Mallinckrodt (Phillipsburg, NJ). All compounds used for screening were obtained as described previously.[24] Briefly, the compound library is composed of 437 known biologically active compounds distributed across 113 mixtures with three or four compounds in each mixture.

Proteins. The uniformly ^{13}C- and ^{15}N-labeled *S. aureus* DnaG primase CTD and the uniformly ^{15}N-labeled *S. aureus* DnaG primase CTD were designed and purified by Nature Technologies (Lincoln, NE). The details of protein expression and purification were previously described.[25] The vector added 20 N-terminal residues: an HN metal affinity tag (MGHNHNHNHNHNGG) followed by a protease-sensitive DDDD sequence.[25]

NMR Data Analysis, Structure Calculations, and Refinement. The NMR spectra for obtaining the protein backbone assignments were recorded at 298 K on a five-channel 600 MHz Bruker Avance spectrometer equipped with a 5 mm TXI probe. The NMR spectra for the protein side chain assignments were recorded at the Rocky Mountain Regional 900 MHz NMR Facility on a four-channel 900 MHz Varian INOVA spectrometer equipped with a 5 mm HCN probe. The nearly complete *S. aureus* DnaG primase CTD NMR resonance assignments have been previously reported.[25] Distance constraints were obtained from three-dimensional (3D) ^{15}N-edited NOESY and 3D ^{13}C-edited NOESY spectra that were recorded at 900 MHz.[26] Hydrogen bond constraints were determined using the (CLEANEX-PM)-FHSQC experiment.[27] All torsion angle constraints were obtained by chemical shift analysis using the TALOS[28] software program and measured coupling constants from an HNHA experiment.[29]

For the backbone NMR assignment experiments, ^{13}C- and ^{15}N-labeled protein was concentrated to 1.2 mM in a 95% H_2O/5% D_2O buffered solution of 100 mM NaCl and 25 mM potassium phosphate at pH 6.64 (uncorrected) using an Amicon ultra centricon [molecular weight (MW) cutoff of 10000 Da]. Long-term protein stability was enhanced by adding 50 mM arginine and 50 mM glutamine. For the side chain experiments, ^{13}C- and ^{15}N-labeled protein was concentrated to 1.4 mM in the same buffer. All multidimensional experiments were processed using NMRPipe[30] and analyzed using PIPP[31] and CCPNMR.[32]

Nuclear Overhauser effect (NOE) assignments were obtained by using 3D ^{15}N-edited NOESY and 3D ^{13}C-edited NOESY. NOE intensities were sorted visually into four classes: strong (1.8−2.5 Å), medium (1.8−3.0 Å), weak (1.8−4.0 Å), and very weak (3.0−5.0 Å). Upper limits for distances involving methyl protons and nonstereospecifically assigned methylene protons were corrected appropriately for center averaging. Initial NOE assignments were completed with Autostructure.[33] Despite the high magnetic field of 900 MHz, the extent of cross-peak overlap was great enough to warrant manual refinement of the NOE assignments.

Hydrogen bond constraints were determined using the (CLEANEX-PM)-FHSQC experiment, which identifies amide residues with fast water exchange rates.[27] Hydrogen bond constraints were assigned to amides in secondary structure regions for any two-dimensional (2D) ^{1}H−^{15}N HSQC peaks lacking a corresponding peak in the (CLEANEX-PM)-FHSQC spectrum. The carboxyl oxygen−amide nitrogen hydrogen bond distances were set at 2.8 Å, while the carbonyl oxygen−amide proton distances were set at 1.8 Å. All carboxyl groups within 2.5 Å of slowly exchanging amide groups were constrained to be involved in a hydrogen bond.

The structures were refined using the hybrid distance geometry dynamic-simulated annealing method[34] using XPLOR-NIH[35,36] adapted to incorporate pseudopotentials for $^{3}J_{HN-H\alpha}$ coupling constants, secondary $^{13}C\alpha/^{13}C\beta$ chemical shift constraints, and a conformational database potential.[37−41] The force constant for the conformational database was kept relatively low (0.5−1.0 kcal/mol) throughout the simulation to allow the experimental distance and torsion angle constraints to predominately influence the resulting structures. The force

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

constants for the NOE and dihedral constraints were 30 and 10 times stronger, respectively, than the force constants used for the conformational database.[42] All peptide bonds were constrained to be planar and trans. There were no hydrogen bonding or electrostatic empirical potential energy terms in the target function. A total of 1000 structures were calculated, and the 20 lowest-energy structures were selected to become the ensemble of *S. aureus* DnaG primase CTD structures. Each of these 20 structures was subjected to further energy minimization using explicit water with Crystallography and NMR system (CNS) version 1.2.[43] Lennard-Jones and electrostatic potentials accounted for explicit water solvation using a modification of the procedure and force field of Nilges.[44,45]

Protein Backbone Dynamics. The NMR experiments for protein dynamics analysis were performed on a Bruker (Billerica, MA) 500 MHz Avance spectrometer equipped with a triple-resonance, *Z*-axis gradient cryoprobe. The sample was the uniformly ^{15}N-labeled *S. aureus* DnaG primase CTD concentrated to 1.2 mM in a 95% H_2O/5% D_2O buffered solution of 100 mM NaCl, 50 mM arginine, 50 mM glutamine, and 25 mM potassium phosphate at pH 6.64 (uncorrected) using an Amicon ultra centricon (MW cutoff of 10000 Da).

The previously described[46−48] experiments included a 2D ^{15}N−^{1}H HSQC experiment (*hsqct1etf3gpsi*) designed to measure T_1 relaxation rates with delay times of 0.0, 5.39, 53.92, 134.80, 269.60, 404.40, 539.20, 674.00, and 1078.40 ms, a 2D ^{15}N−^{1}H HSQC experiment (*hsqct2etf3gpsi*) designed to measure T_2 relaxation rates with delay times of 0.0, 17.6, 35.2, 52.8, 70.4, 105.6, 123.2, 140.8, 158.4, and 176.0 ms, and a 2D ^{15}N−^{1}H HSQC experiment (*hsqcnoef3gpsi*) designed to measure NOE changes.

The relaxation rates (T_1 and T_2) for each amino acid within the structure were calculated by fitting the intensity of each peak to the intensity decay curve (eq 1)

$$I_t = I_0 \exp\left(-\frac{t}{T_{1,2}}\right) \quad (1)$$

where I_t is the intensity of each peak at delay time t and I_0 is the initial steady state intensity.

The NOE values were determined by the ratio of peak intensity between the saturated (I_{sat}) and unsaturated (I_{unsat}) spectra.

$$\text{NOE} = I_{sat}/I_{unsat} \quad (2)$$

The T_1 rates, T_2 rates, and NOE ratios were determined from their fits to their respective equations using Kaleidagraph (Synergy Software, Reading, PA) and exported to the Fast-Model Free[49] to calculate an overall correlation time (τ_m) and per residue order parameters (S^2), internal motion (τ_e), and chemical exchange (R_{ex}) using the Lipari−Szabo model-free method.[50] Some of the NOE ratios were above the theoretical maximum of 1.0 and were set to 1.0 to run Fast-Model Free.

Primase CTD and C1 Subdomain Similarity Phylograms. An all-versus-all root-mean-square deviation (RMSD) matrix table (Table S2) was generated using the MatchMaker tool in UCSF Chimera.[51] All chains from X-ray structures were included in the structural comparison. Only the best representative structure from the NMR ensemble, identified by XPLOR as having the lowest RMSD relative to the mean, was included in the comparison (see Table S3).

The computed RMSD matrices were then inserted into the Splitstree4[52] program to generate the unrooted phylogenic trees depicted in Figure 3. The trees were generated using the Neighbor-join algorithm. The pairwise RMSDs were calculated between the complete C-terminal domains (Figure 3A) and between the aligned C1 subdomains (Figure 3C).

Primase CTD Sequence Similarity. The sequences for the *E. coli*, *G. stearothermophilus*, *H. pylori*, and *S. aureus* primase C-terminal domains were acquired from the PDB.[53] Three different sequence alignment programs, Muscle, T-Coffee, and Clustal Omega,[54−56] were used to determine a sequence similarity among the four primase CTD proteins. The results obtained from the three sequence alignment programs were identical and were converted into the Phylip format for analysis using Splitstree4.[52] The unrooted phylogenic tree (Figure 3B) was created using the Neighbor-join algorithm and equal angle method.

NMR Ligand Affinity Assays. Sample preparation and experimental parameters for the NMR ligand affinity screen were as described previously.[57] Briefly, each ligand mixture (113 total) was screened using a one-dimensional (1D) ^{1}H NMR spectrum with excitation sculpting.[58] Each NMR sample contained 100 μM ligand and 25 μM protein in a 99.99% D_2O-buffered solution of 20 mM d_{19}-bis-Tris at pH 7.0 (uncorrected) with 2% DMSO-d_6 to maintain ligand solubility and 11.1 μM 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt as a chemical shift reference. All 1D ^{1}H NMR spectra were processed with ACD/1D NMR manager version 12.0 (Advanced Chemistry Development, Inc., Toronto, ON). Each 1D ^{1}H NMR spectrum was compared to the corresponding free ligand mixture reference spectrum and visually analyzed to identify binding ligands based on a decrease in the intensity of ligand NMR resonances. A 2D ^{1}H−^{15}N HSQC spectrum was collected for each ligand with a positive response from the 1D ^{1}H NMR line broadening screen using 500 μM ligand and 100 μM protein under the same buffer conditions as the 1D ^{1}H NMR screen except for the 95% H_2O/5% D_2O solvent. A single ligand-free 2D ^{1}H−^{15}N HSQC spectrum was collected as a reference, where a binding event was confirmed on the basis of a clustering of surface residues that incurred a chemical shift change one standard deviation above the average.

Protein−Ligand Titration and Docking. The DnaG primase CTD HSQC spectrum was recorded at a protein concentration of 1 mM and titrated with adenosine (0.9, 1.8, 2.6, 3.3, and 4.0 mM), acycloguanosine (0.7, 1.9, 3.5, and 5.2 mM), and myricetin (0.9, 1.8, 2.6, 3.3, and 4.0 mM). All NMR data were collected at 25 °C on a 700 MHz Bruker Avance III spectrometer equipped with a 5 mm QCI-P probe with cryogenically cooled carbon and proton channels. The 2D ^{1}H−^{15}N HSQC protein titration data were processed and analyzed in NMRPipe.[30] Briefly, the spectral titration data for each ligand were automatically peak picked and adjusted manually within NMRDraw. The ^{1}H−^{15}N and ^{15}N chemical shifts from the selected peaks were then fit to a K_a ($1/K_d$) according to a previously published method.[59] Subsequently, the peaks that showed significant binding were manually assigned to residues. AutoDock was then used to individually dock the three ligands into a primase CTD-binding site defined by the observed chemical shift perturbations. Chimera was used to visualize the lowest-energy docked complexes.[51]

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

RESULTS

To fight antibiotic-resistant infections, there is a need to identify new antibiotic targets. The bacterial primase CTD is an especially promising target because eukaryotic primases do not have a cognate sequence or structure and because its critical function is to interact with DnaB helicase to limit DNA replication to the replication fork. Among a number of factors, evaluating the "drugability" of a protein target requires a fundamental understanding of its structure−function relationship.[60] As a first step toward this goal, we determined the three-dimensional structure of the primase CTD from the pathogen *S. aureus*.

***S. aureus* Primase CTD Structure Determination.** Backbone and side chain resonance assignments were completed using standard triple-resonance NMR experiments.[25] In summary, the backbone resonance assignment was 93% complete with 133 amino acids of the 143 primase residues assigned unambiguously in the 2D $^{1}H-^{15}N$ HSQC spectrum. Nearly all the peaks were uniformly shaped and separated from the others, which indicates that each residue is in a unique environment and that the *S. aureus* primase CTD exists in a single conformation. Unassigned residues include M1−H13, D19, E470, H479, L480, M481, T500, R536, E537, E543, P551, and Y552. Primase CTD residues are numbered relative to the complete DnaG primase sequence, while the N-terminal purification tag is simply numbered from residue 1 to 20. Most of the unassigned residues were located in the purification tag or within highly solvent exposed regions. The structure consists of eight α-helices. Residues H479, L480, and M481 were in a turn region between helices 1 and 2. Residue T500 was in an unstructured loop region between helices 2 and 3, whereas residues P551 and Y552 were in an unstructured loop region between helices 5 and 6. The only unassigned residues that appear to be localized within a secondary structure were E470, R536, E537, and E543. E470 is the second residue of helix 1; R536 and E537 are in the middle of helix 5, and E543 is near the end of helix 5. An exhaustive analysis of the NMR data set could not yield assignments for these residues, suggesting the end of helices 1 and 5 may undergo partial unfolding leading to chemical shift exchange broadening for these residues.

The solution structure of the *S. aureus* primase CTD was calculated using 1823 distance constraints, 280 dihedral constraints, 256 $^{13}C\alpha$ and $^{13}C\beta$ chemical shift constraints, and 82 $^{3}J_{NH\alpha}$ coupling constant constraints, among others (Table 2).

In the initial phase, 1000 structures were calculated from 10 individual sets of 100 structures each using XPLOR-NIH[35,36] as described previously.[64] Each set of structures was started from a randomly generated seed. The lowest-energy structures were consolidated to generate a set of 20 low-energy structures that were further refined within a virtual water environment using the RECOORD recalculated coordinates[45] as implemented in CNS.[43]

The resulting *S. aureus* primase CTD structures (Figure 1) form a self-consistent set as determined by a range of statistical analyses (Table 2).

The root-mean-square deviation (RMSD) for all 1823 experimental distance constraints is 0.040 ± 0.023 Å, which implies a good agreement between the structural ensemble and the constraints. None of the distance constraints have a violation that exceeds 0.247 Å. There are also small deviations from the constraints for the $^{13}C\alpha$ chemical shifts, $^{13}C\beta$ chemical shifts, and the $^{3}J_{HN-H\alpha}$ coupling constants, indicating consistent conformations for both the backbone and the side chains. Similarly, the deviations from idealized covalent geometries are consistent with good quality structures.

Table 2. Structural Statistics[a]

	⟨SA⟩	σ	$(SA)_r$
distance restraints (Å)			
all (1823)	0.040	0.023	0.022
inter-residue sequential ($\|i-j\|=1$) (460)	0.056	0.021	0.030
inter-residue short-range ($1<\|i-j\|\leq 5$) (444)	0.062	0.018	0.041
inter-residue long-range ($\|i-j\|>5$) (140)	0.247	0.093	0.374
intraresidue (663)	0.001	0.001	0.003
backbone hydrogen bonds (116)	0.014	0.005	0.008
dihedral angle restraints (deg) (280)	1.21	0.23	1.12
$C\alpha$ chemical shift restraints (ppm) (130)	1.05	0.05	1.08
$C\beta$ chemical shift restraints (ppm) (126)	1.06	0.04	1.09
$^{3}J_{NH\alpha}$ coupling restraints (Hz) (82)	0.011	0.0005	0.011
F_{NOE} (kcal mol^{-1})	87	36	46
$F_{torsion}$ (kcal mol^{-1})	28	20	21
F_{VDW} (kcal mol^{-1})	−509	31	−510
idealized covalent geometry			
bond lengths (Å) (2400)	0.0114	0.0005	0.0108
all angles (deg) (4302)	1.44	0.06	1.43
improper torsion angles (deg) (1287)	1.95	0.23	1.84

[a]⟨SA⟩ is the average value from an all-versus-all comparison of the set of 20 annealed structures. σ is the standard deviation for the all-versus-all comparison. $(SA)_r$ is the value for the restrained minimized mean structure. The number of restraints for each parameter is given in parentheses. For backbone NH−CO hydrogen bonds, the two restraints are as follows: r(NH−O) = 1.5−2.3 Å, and r(N−O) = 2.5−3.3 Å. The values of the square-well NOE (F_{NOE}) and torsion angle ($F_{torsional}$) potentials[61] are calculated with force constants of 50 kcal mol^{-1} Å^{-2} and 200 kcal mol^{-1} rad^{-2}, respectively. The value of the Lennard-Jones van der Waals term (F_{VDW})[62] is calculated with a force constant of 4 kcal mol^{-1} Å^{-4} with the CHARMM[63] empirical energy function. The improper torsion angle restraints serve to maintain planarity and chirality.

The quality of the *S. aureus* primase CTD NMR structure was assessed using two suites of programs (Table 3).

PROCHECK[65] indicates the average minimized structure has an overall G factor of −0.12 ± 0.04 with no bad contacts, consistent with a good quality structure. The Protein Structure Validation Software suite of programs (PSVS)[66] gave a Verify3D score of −4.58 ± 0.71, which is within the typical range. The Molprobity module indicates that the *S. aureus* primase CTD structure has a very good Z score (−2.50) compared to the average Z score (−10.74) for all NMR structures in the PDB. Finally, a Ramachandran plot of backbone dihedral angles for all non-glycine residues indicates that 83.4 ± 3.2% lie within the "most favored" region and only 0.7 ± 0.6% lie within "disallowed" regions. That is, one residue in 14 of the 20 structures is found in the disallowed region but never in a consistent location within the structure (Table S4). These residues tend to have a smaller number of constraints, which results in a small penalty for unusual conformations during energy minimization.

The final step in the analysis was to compare the 20 energy-minimized structures with the mean structure (Table 4).

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

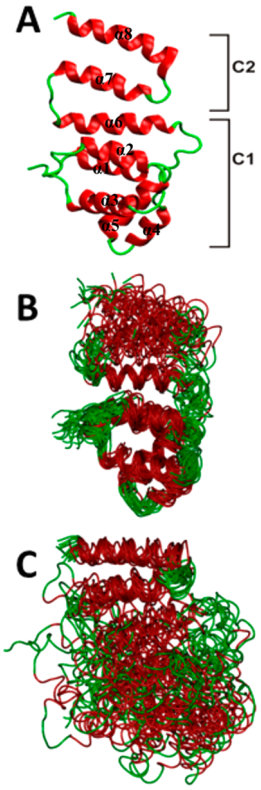


Figure 1. Primase CTD ensemble overlay. (A) Ribbon diagram of the average water-refined structure. The C1 subdomain is composed of helices 1−6, and the C2 subdomain is composed of helices 7 and 8. All structures were generated with Visual Molecular Dynamics (VMD-XPLOR) and are colored according to the secondary structure: red for α-helices and green for loops. (B) Overlay of the backbone trace of the 20 lowest-energy, water-refined structures aligned with residues 468−566 from the C1 subdomain. (C) Overlay of the backbone trace of the 20 lowest-energy structures aligned with residues 570−603 from the C-terminal C2 subdomain.

Within the C1 subdomain comprising the first two-thirds of the residues, the average RMSD of the 20 lowest-energy structures about the mean coordinate position is 1.54 ± 0.19 Å for backbone atoms and 2.37 ± 0.17 Å for all heavy atoms. When only secondary structure elements are used for alignment, the mean coordinate position deviation is 1.18 ± 0.14 Å for backbone atoms and 1.93 ± 0.16 Å for all heavy atoms. These deviations indicate the backbone structure is well-defined. The side chain residues have fewer constraints and a correspondingly higher degree of structural disorder.

***S. aureus* Primase CTD Structure.** The 20 energy-minimized *S. aureus* primase CTD structures (Figure 1) are deposited in the PDB as entry 2LZN. The structure is composed of eight helices arranged into two subdomains as is the primase CTD NMR structure from *G. stearothermophilus*.[9] In our *S. aureus* primase CTD structure, the first six helices create the C1 subdomain and encompass residues 467−565 (Figure 1A). The last two helices, 7 and 8, create the C2 subdomain encompassing residues 572−603. When the 20 structures are overlaid using the alignment of the C1 subdomain backbone residues as a guide (Figure 1B), the C2 subdomain structures do not superimpose well. The converse overlay based on the alignment of the C2 subdomain shows the same effect (Figure 1C). The poor superimposition arises because of the very small number of structural restraints between the subdomains. However, the linker between the subdomains is constrained by 11 NOEs from the C1 subdomain, 77 NOEs within the linker, and 17 NOEs from the C2 subdomain. On the other hand, the loop region between the subdomains lacks sequential NH−NH NOEs in the $^{1}H-^{15}N$ HSQC-edited NOESY, and loop residues G567, Q568, and E569 exhibit exchange peaks in the CLEANEX experiment.[27] These residues are likely undergoing rapid exchange with the solvent, a feature indicative of exposed residues that lack protection from hydrogen bonds found in

Table 3. Ensemble Self-Consistency[a]

	⟨SA⟩	σ	$(SA)_r$
PROCHECK[b]			
overall *G* factor	−0.12	0.04	−0.09
H-bond energy	0.82	0.06	0.8
no. of bad contacts per 100 residues	0.0		0.0
PSVS *Z* scores[c]			
Verify3D	−4.58	0.71	−4.82
ProsaII (-ve)	−1.14	0.34	−1.20
Procheck (ϕ-ψ)	−0.56	0.42	−0.28
Procheck (all)	−1.92	0.32	−2.07
MolProbity clash score	−2.50	0.97	−2.46
Ramachandran space[d]			
most favored regions	83.4%	3.2	85.5%
additional allowed regions	14.5%	3.0	13.0%
generously allowed regions	1.43%	1.2	0.7%
disallowed regions	0.7%	0.56	0.7%

[a]⟨SA⟩ is the average value from an all-versus-all comparison of the set of 20 annealed structures. σ is the standard deviation for the all-versus-all comparison. $(SA)_r$ is the value for the restrained minimized mean structure. [b]PROCHECK values are in comparison to the likelihood of finding each residue within the derived structure as compared to a database of residues found in similar environments within structures with a resolution below 1.8 Å. [c]PSVS (Protein Structure Validation Suite) provides an analysis of five separate structural analysis programs, each of which measures a different parameter. The PSVS reports *Z* scores for each analysis, where a *Z* score is the number of standard deviations from the mean. The authors of PSVS suggest that *Z* scores of less than −5 should be reevaluated. [d]Ramachandran space analysis from the PDB sum Web site uses a database of structures with a resolution below 1.8 Å to determine how a region should be classified.

Table 4. Ensemble Self-Consistency of the C1 Subdomain[a]

comparison	backbone atoms	all heavy atoms
all residues		
⟨SA⟩ vs SA_{mean} (Å)	1.54 ± 0.19	2.37 ± 0.17
⟨SA⟩ vs $(SA)_r$	1.98 ± 0.28	3.05 ± 0.32
$(SA)_r$ vs SA_{mean}	1.26	1.94
secondary structures		
⟨SA⟩ vs SA_{mean}	1.18 ± 0.14	1.93 ± 0.16
⟨SA⟩ vs $(SA)_r$	1.56 ± 0.22	2.55 ± 0.26
$(SA)_r$ vs SA_{mean}	1.03	1.66

[a]⟨SA⟩ is the average value from an all-versus-all comparison of the set of 20 annealed structures. $(SA)_r$ is the value for the restrained minimized mean structure. SA_{mean} is the Cartesian mean structure created from the geometric mean for each atom of all 20 structures. The SA_{mean} structure does not adhere to any of the restraints for bond lengths, angles, etc. The secondary structure residues are 469−478 (α1), 484−493 (α2), 502−515 (α3), 524−528 (α4), 533−544 (α5), and 552−564 (α6). The secondary structure RMSD values were measured by aligning all secondary structure elements and calculating an average and standard deviation.

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

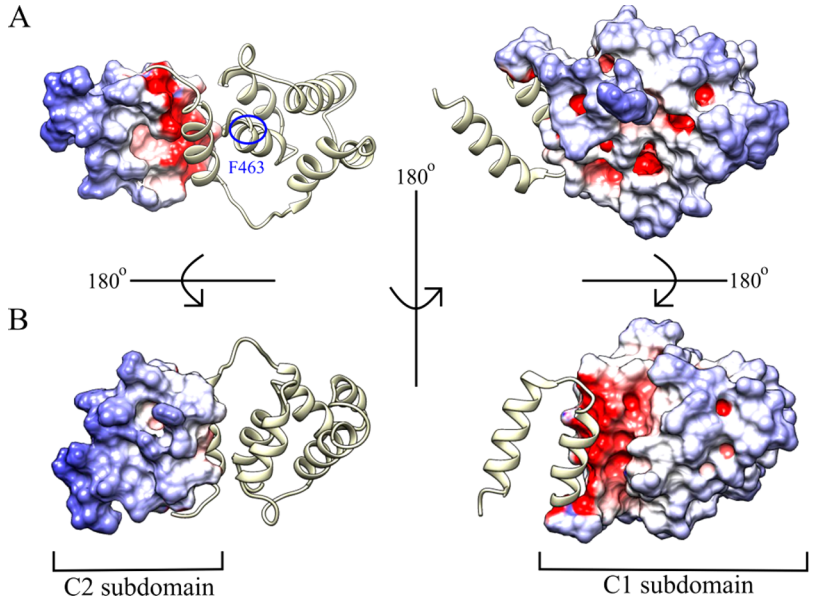


Figure 2. Electrostatic surface potential. Surface charge distribution of the C-terminal domain showing (A) the face toward the RNA polymerase domain (F463 is the N-terminal residue of this domain) with the space-filled C2 subdomain on the left and the ribbon-diagrammed C1 subdomain on the right and (B) the solvent-exposed face (180° rotation around the *x*-axis relative to the structures in panel A) with the space-filled C1 subdomain and ribbon-diagrammed C2 subdomain.

secondary structures. As we shall see below, the dynamics analysis indicates that the subdomains behave as two independent domains.

The surface charge distribution shows the exposed surface is uniformly but weakly electrically positive (Figure 2).

The buried interface between the subdomains is electrically negative, as reflected in the moderately large Verify 3D *Z* score. This repulsion may destabilize the interaction between the C1 and C2 subdomains, causing them to prefer a more extended conformation. In addition, there are some partially buried electrically negative residues in the C1 subdomain. The overall charge distribution is unlike the uniformly negative CTD from *G. stearothermophilus*,[9] suggesting that it is not a universal feature of primase CTDs.

Primase CTD Conformations. Every enzyme family has a distinct sequence−structure relationship linked to the evolutionary pressures on the sequence and structure within each organism.[67] To determine how the *S. aureus* primase CTD structure relates to other known primase CTD structures (Table 1), the MatchMaker tool in UCSF Chimera[51] was used to calculate pairwise RMSDs for the eight known primase CTD structures (Table S2). The resulting unrooted phylogram (Figure 3A) shows three main branches that correlate with distinct conformations (Figure 4).

The first branch (conformation A) includes our NMR structure from *S. aureus* (2LZN, chain 14) and the *G. stearothermophilus* NMR structure (1Z8S, chain 5). The second branch (conformation B) includes the *E. coli* NMR structure (2HAJ, chain 19), the *G. stearothermophilus* crystal structure (2R6A, chain C), and both *H. pylori* crystal structures (4EHS, chains A and B). The third branch (conformation C) includes both *E. coli* crystal structures (1T3W, chains A and B). Conformations A and B are more similar to each other than they are to conformation C. Furthermore, it has been proposed[20] that conformation C may not be biologically relevant because the crystals were formed at a low pH (4.6), which resulted in a packing effect of the swapped dimer conformation.

Because the unrooted phylogram for the full length structure did not split according to sequence phylogeny, we created unrooted phylograms based on CTD sequence similarity (Figure 3B) and the primase CTD C1 subdomain structure (Figure 3C). From the sequence similarity phylogram, we observed two distinct branches corresponding to the expected phylogenetic split between the Firmicutes (*G. stearothermophilus* and *S. aureus*) and the Proteobacteria (*E. coli* and *H. pylori*). The same phylogenetic split was observed for the C1 subdomain structures, showing that there is a structure−sequence phylogenetic split for this protein.

Primase CTD Flexibility. An NMR dynamics analysis of the *S. aureus* primase CTD structure was conducted to complement the structural analysis. NMR relaxation parameters T_1 and T_2 and the relative ratio of NOE enhancement were measured on a per residue basis and analyzed using Fast-Model Free[49] to calculate order parameters (S^2) and chemical exchange (R_{ex}) values (Figure 5).

In general, the T_1 values for the C1 subdomain residues were higher than the T_1 values for the C2 subdomain residues, indicating that the two subdomains exhibit different dynamics. A similar observation was made for the T_2 values, except that the T_2 values for the C1 subdomain were lower than the T_2 values for C2.

It is important to note that Fast-Model Free did not converge on a result when using the entire primase CTD structure. This is consistent with the observation that the C1 and C2 subdomains behave as two separate domains and are dynamically independent. A convergent result was obtained only when the NMR dynamics data of the C1 and C2 subdomains were modeled separately. The C1 subdomain included residues 463−572, and the C2 subdomain included residues 565−605. The linker region, corresponding to residues 565−572, was included in the dynamics analysis for both subdomains.

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

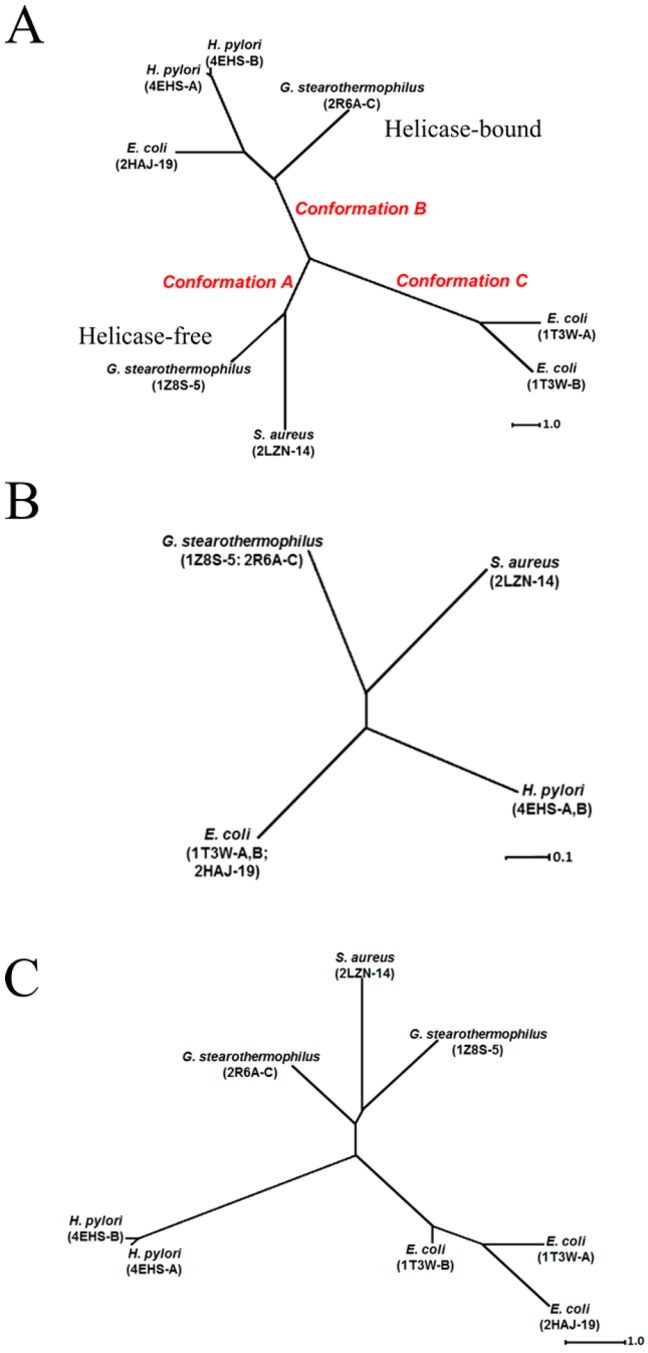


Figure 3. Unrooted phylograms. Representation of the (A) primase CTD structure, (B) sequence similarity, and (C) primase CTD C1 subdomain structure. All trees were generated with SplitsTree4. Each tree contains the following species: *S. aureus* primase CTD (this study), *G. stearothermophilus* primase CTD, *H. pylori* primase CTD, and *E. coli* primase CTD.

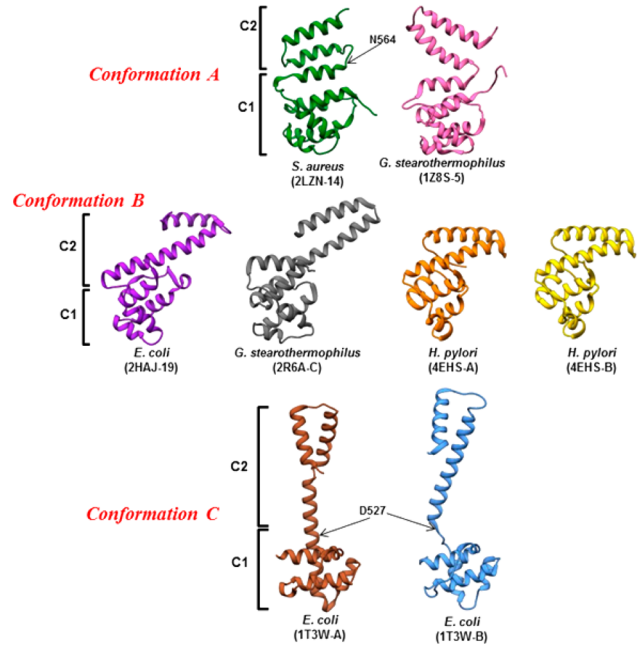


Figure 4. Primase CTD conformation classes. The species structures were categorized in three conformations from Figure 3A on the basis of compactness: A, compact; B, intermediate; C, extended. Conformation C may be biologically irrelevant.[20]

The order parameters (S^2) for the C1 subdomain residues covered a large range (0.240−1.000), whereas those of the C2 subdomain residues were close to 1 (Figure 5D). The average S^2 values were 0.79 ± 0.02 and 0.97 ± 0.03 for the C1 and C2 domains, respectively. In general, an S^2 value approaching 1.0 indicates a highly ordered residue with limited flexibility, while a lower S^2 value indicates greater flexibility. This implies that the C1 subdomain has an overall flexibility greater than that of the C2 subdomain. Within the primase CTD structure, there are two stretches of three or more residues with S^2 values of <0.75: residues 488−498, residues 555−557, and residues 561−563. Residues 488−498 lie in helix 2 and the linker between helices 2 and 3. Residues 555−557 and 561−563 lie in helix 6 (Table S5). The flexibility of the linker between helices 2 and 3 is consistent with the variable lengths for helix 3 observed in the structures from other organisms. It is possible that the decrease in the flexibility of helix 3 is part of an allosteric system that provides an entropic energy change when helices 6 and 8 bind to the helicase N-terminal domain (NTD).[68] The flexibility of the second half of helix 6 shows how the unstructured linker between helices 6 and 7 may act as a nucleus for unraveling helix 6. Surprisingly, the linker region corresponding to residues 565−572 was observed to be dynamic, indicating that the point of flexibility between the two domains is located at the end of helix 6. This finding further suggests that the dynamic, second half of helix 6 (residues 561−564) compensates by either unraveling or bending for the increased level of motion between the two subdomains at the linker region.

To identify the hinge between the two subdomains, we searched for a residue with a low order parameter and fast chemical exchange. N564 had one of the lowest order parameters (0.658) and the fastest chemical exchange (17.7 Hz), indicating a high degree of motion on the picosecond to nanosecond and millisecond time scales. It is the final residue of helix 6 and is located between a very rigid C2 subdomain and a flexible C1 domain, potentially implicating N564 as a key residue for a conformational change.

In addition to per residue order parameters and chemical exchange rates, an overall correlation time (τ_m) was calculated for each subdomain. The correlation times for the C1 and C2 subdomains were 7.2 and 6.4 ns, respectively. The predicted value for a hydrated protein with a molecular weight of 17.2 kDa, corresponding to the intact primase CTD structure, is 7.2 ns, where $\tau_m \approx$ MW/2400.[69] The predicted tumbling times

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

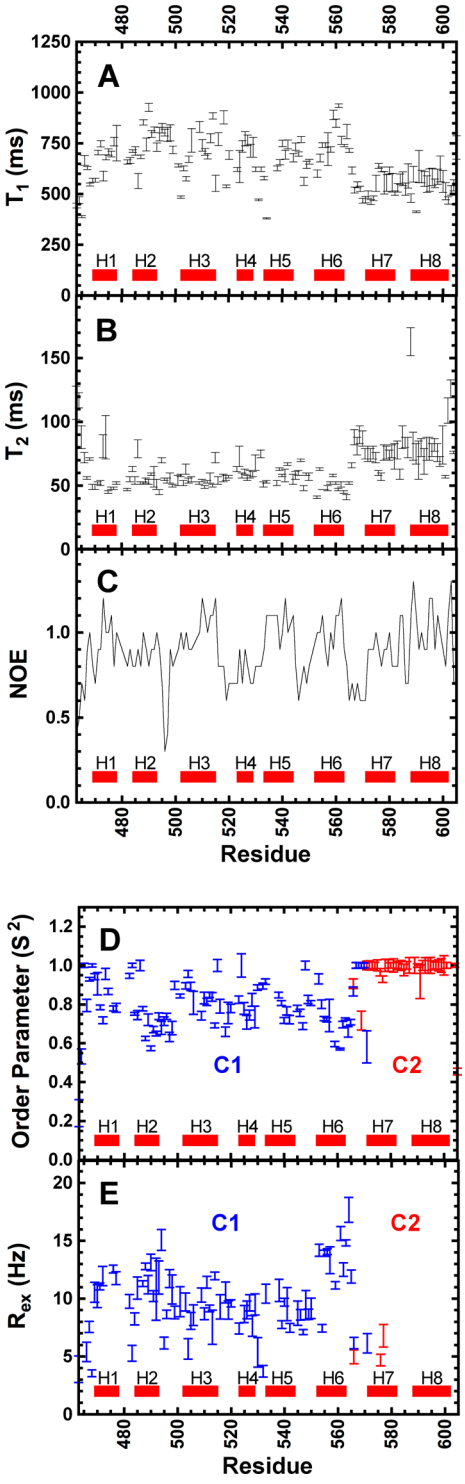


Figure 5. Dynamics of the *S. aureus* primase CTD. NMR relaxation parameters (A) T_1 and (B) T_2, (C) NOE enhancements, (D) S^2 order parameter, and (E) R_{ex} chemical exchange rate all plotted per residue. The C1 subdomain included residues 463−572, and the C2 subdomain included residues 565−605. The linker region corresponding to residues 565−572 was included in the dynamics analysis for both subdomains.

were 5.45 and 2.1 ns for the C1 (13.1 kDa) and C2 (5.0 kDa) subdomains, respectively. The observed τ_m values for both subdomains are substantially higher than predicted and share a greater degree of similarity than the predicted values. This indicates the C1 and C2 subdomains are both tumbling at a rate approximately the same as that of the intact protein. The model that emerges is one in which the two subdomains can adopt a wide range of relative orientations, with N564 as a pivot point, but share an overall tumbling rate.

Identification of Ligands That Bind to the *S. aureus* Primase CTD. A high-throughput NMR ligand affinity screen of the *S. aureus* primase CTD was undertaken to identify potential binding compounds. Of the 423 compounds tested, a total of 13 compounds were shown to bind the *S. aureus* primase CTD using a 1D ^{1}H NMR line broadening screen: acycloguanosine, 3-aminopropionitrile fumarate, chelerythrine chloride, didecyldimethylammonium bromide, 5,5-diphenylhydantoin, L-histidine, (±)-α-lipoamide, 1-methylimidazole, mitoxantrone dihydrochloride, myricetin, (±)-propranolol hydrochloride, sodium creatine phosphate, and sodium DL-lactate. Of these compounds, acycloguanosine and myricetin were further analyzed for binding with the primase CTD from *S. aureus*. Furthermore, adenosine was selected for binding studies because of its structural similarity to acycloguanosine. Quercetin, luteolin, and kaempferol were also selected for their structural similarity to myricetin, but their low solubility prevented further study.

A 2D ^{15}N−^{1}H HSQC spectrum was collected for free primase and for the primase−ligand complex using acycloguanosine, adenosine, or myricetin. All three compounds were shown to bind the primase CTD on the basis of protein chemical shift perturbations (CSPs) of Y552, N564, T570, L574, E580, I584, G585, L589, and Q590 (Figure 6A).

Although initially unassigned, Y552 was discovered during the titration experiment and assigned on the basis of its spatial proximity to other perturbed residues. None of the three compounds reached CSP saturation at the highest concentration tested (≥4.0 mM), indicating that their K_D's were >4 mM. Acycloguanosine, adenosine, and myricetin perturbed CSPs more than one standard deviation from the average of all nonbinding site residues. Specifically, acycloguanosine, adenosine, and myricetin caused CSPs of 0.35 ± 0.09, 0.16 ± 0.29, and 0.07 ± 0.02 ppm, respectively, for the binding site residues.

When the nine residues incurring CSPs were mapped onto the structure of primase CTD, eight of the residues were located adjacent to one another on helix 6 or 7 (Figure 6B). The residue farthest from the other binding site residues was N564, which was previously identified as a key residue for a C1−C2 conformation change (Figure 5). When adenosine was docked into the groove between helices 6 and 7 (Figure 6B), the adenine moiety filled a hydrophobic pocket while its ribose was exposed to solvent and made specific interactions with three residues: Y552, I584, and G585 (Figure 6C). All three compounds docked into the same primase CTD pocket and adopted a uniform binding conformation. The hydrophobic fused rings were buried deep in the pocket, and the hydrophilic moieties interacted with the same three residues at the opening of the pocket (Figure 6D). Because these compounds bind to the closed conformation of the primase CTD, they are likely to act as an inhibitor and weaken the helicase−primase interaction.

DISCUSSION

We report the first primase CTD structure from a medically relevant, mesophilic Firmicutes bacterium. Its structure shares many features with the known thermophilic Firmicutes and mesophilic proteobacterial primase CTD. It has an N-terminal

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

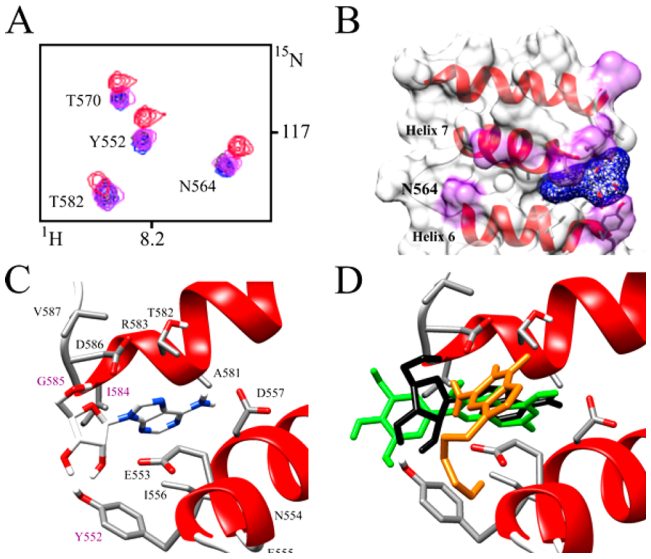


Figure 6. Adenosine-binding site identification. When the *S. aureus* primase CTD (1 mM) was titrated with up to 4.0 mM adenosine, nine residues exhibited chemical shift perturbations in the 2D ^{1}H−^{13}C HSQC spectrum of the primase CTD. (A) A detail from the overlay of the 2D ^{1}H−^{13}C HSQC spectra is colored according to the addition of 0 mM (blue), 2.6 mM (purple), and 4.0 mM (red) adenosine. (B) An expanded view of a surface rendition of the primase CTD NMR structure is superimposed onto a ribbon structure to highlight secondary structure elements. Adenosine (blue tessellated structure) was docked into the primase CTD region defined by the nine residues with significant chemical shift perturbations (colored purple), which are located between helices 6 and 7. The side chains for all the residues that interact with adenosine are drawn with licorice bonds and are the same side chains displayed in panel C. Residue N564 and α-helices 6 and 7 are labeled. (C) Expanded view of the ligand-binding pocket (flipped 180° from panel B) represented as a ribbon structure with side chains interacting with adenosine shown as licorice bonds. The residues interacting with adenosine are labeled. The labels for residues with significant CSPs are colored purple. (D) Superposition of the docked conformations of myricetin (green), adenosine (black), and acycloguanosine (orange) into the same expanded view of the primase CTD-binding site from panel C.

six-helix bundle called the C1 subdomain connected by linker residues to a C-terminal helix−turn−helix segment called the C2 subdomain. The C1 subdomains from different species share the same overall fold, although there are differences in the lengths of several helices such that the subdomains from *S. aureus* and *G. stearothermophilus* are more similar to each other than to those of *E. coli* or *H. pylori*. The overall resolution of the *S. aureus* structure was low because of significant peak overlap in the NMR spectra, even at a field of 900 MHz. The high leucine and α-helical content of the CTD, coupled with a high degree of dynamic motion, led to peak broadening such that only 56% of side chain Hγ and 76% Hδ could be unambiguously assigned. The number of long-range (>5 Å) NOEs was also lower than anticipated (140) for a 19.6 kDa protein. Nevertheless, the overall RMSD of the 20 lowest-energy structures was reasonable at 1.54 ± 0.19 Å, and the ensemble of structures lies in acceptable ranges of common structure validation programs (all *Z* scores above −5.0).

When a representative structure for *S. aureus* primase CTD was compared to the X-ray and NMR CTD structures from other organisms, three conformational classes were observed (Figures 3A and 4). Two of the conformations are likely related to the function of primase. The conformation of the *E. coli* 1T3W structure is probably not biologically relevant given the low pH of its crystallization and some strong interface packing effects.[20] Of all the structures, the *G. stearothermophilus* crystal structure 2R6A is the only one bound to helicase.[19] This suggests that the structures of *H. pylori* (4EHS) and *E. coli* (2HAJ) are also helicase-bound forms. Because conformation B contains members from both Firmicutes and Proteobacteria, it indicates that the helicase-bound form of the primase CTD is not phylum-dependent. Because the solution structures of the *G. stearothermophilus* and *S. aureus* primase CTD are in conformation A, it appears to be the conformation adopted when primase dissociates from helicase. In the case of *S. aureus*, it has also been possible to use NMR dynamics to identify N564 as the key hinge residue that allows helix 6 to bend into the proximity of helix 7.

Three small molecules, acycloguanosine, adenosine, and myricetin, were discovered to bind a common site on the closed conformation of the primase CTD. The ligand-binding pocket is created only when helices 6 and 7 are adjacent to one another. Among the nine residues with CSPs that define this ligand-binding site, N564 exhibited the largest CSP and was also identified as a key residue in a conformational change between the open and closed form of the primase CTD. As a result, this ligand-binding site is a potential target for the further development of small molecule inhibitors that may "lock" the primase CTD in its closed form and prevent its interaction with the NTD of the helicase. In fact, two of the compounds shown to bind the primase CTD have known inhibitory effects on DNA replication proteins, although their known modes of action are different from that identified here. Acycloguanosine is a known inhibitor of the primase−helicase interaction in herpes simplex virus and acts as a chain terminator.[70,71] Additionally, myricetin inhibits bacterial helicases with an IC_{50} of 10 μM[72] by competing with the ATPase active sites.

Because the *S. aureus* CTD structure shares many features with the *G. stearothermophilus* solution structure, we used the *G. stearothermophilus* DnaB/primase CTD cocrystal structure[19] to identify potential interface residues. The cocrystal structure of the *G. stearothermophilus* primase CTD C2 subdomain (Figure 7A,C) contains five nonpolar residues (F577, L578, A581, A584, and I588) in the final helix that could make contact with the DnaB NTD1 domain (L67, A72, A75, L80, and V86) (Table S6).

Following this idea, we created a homology model of the *S. aureus* primase CTD and DnaB NTD (Figure 7B,D) and found there were conserved but not identical residues in the equivalent positions. The *S. aureus* CTD C2 subdomain residues were V598, L594, K591, E588, and V587, and the *S. aureus* DnaB NTD1 residues were V67, D72, S75, L80, and P86. It has been previously shown that a C-terminal fragment containing the last three helices of the *G. stearothermophilus* C2 subdomain[9] exhibited subnanomolar binding affinity for DnaB but was not capable of stimulating the ATPase activity of DnaB. It could be that DnaG and DnaB from these two organisms interact using spatially similar, hydrophobic, but differently sized residues. The DnaG and DnaB structures may have co-evolved to preserve the primase CTD interaction in a manner that led to species specificity.

The observation that the primase CTD−helicase structural interface is species-specific is consistent with previous studies showing that *S. aureus* helicase will stimulate primer synthesis only upon being incubated with the cognate primase.[9]

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

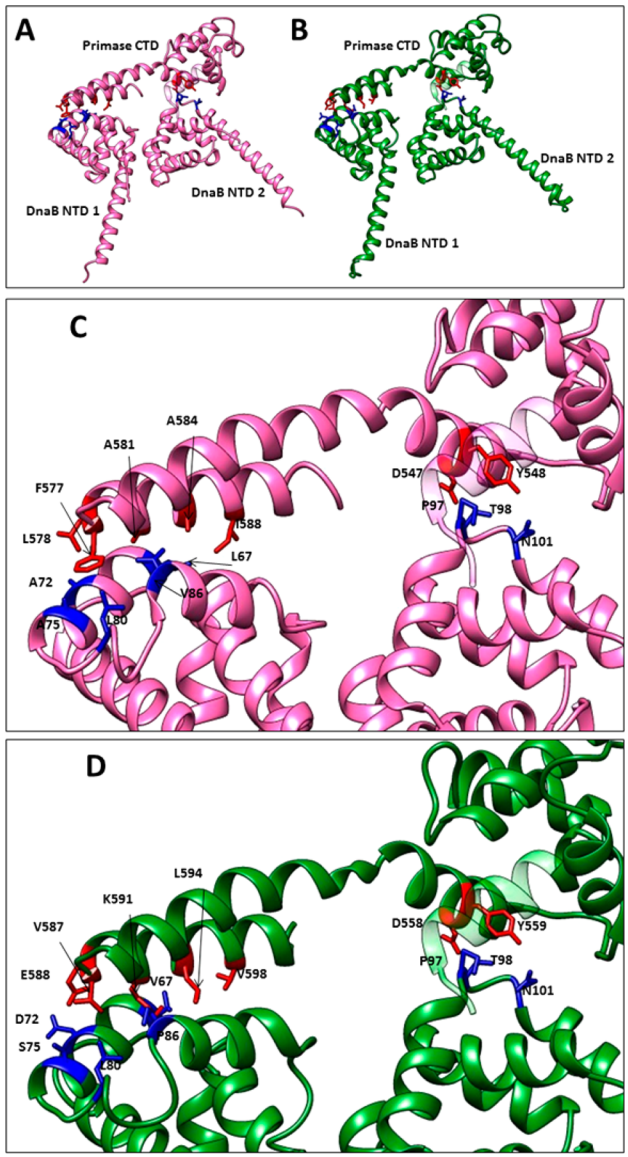


Figure 7. Critical interface residues between the primase and helicase. (A) The cocrystal structure of the complex from *G. stearothermophilus* involves two DnaB NTDs, one of which interacts with the C1 subdomain of the primase and the other with the C2 subdomain. (C) The specific interface residues are shown in an expanded view. (B) This was the template for the homology model of the *S. aureus* primase CTD and DnaB NTDs, which is also shown in an expanded view in panel D.

Stimulation of DnaB ATPase activity by primase binding has been shown to involve the C1 subdomain Y548 in *G. stearothermophilus*[73] and F534 in the *H. pylori*[22] C1 subdomain. Mutation of these residues to alanines reduced binding affinity and the level of ATPase stimulation. An examination of the *G. stearothermophilus* primase CTD structure shows that Y548 and the adjacent residue D547 in helix 6 of the C1 subdomain make contact with P97, T98, and N101 in the DnaB NTD2 domain (Figure 7C and Table S5). The structurally equivalent residues in the *S. aureus* C1 subdomain are D558 and Y559, and within the *S. aureus* DnaB NTD2 domain are also P97, T98, and N101 (Figure 7D). Because the residues are identical between *G. stearothermophilus* and *S. aureus*, it indicates the mechanism of ATPase stimulation is highly conserved.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.6b01273.

Sequence alignment for primase from eight bacteria (Table S1), pairwise structural comparison RMSD matrices for the eight primase C-terminal domain structures (Table S2), comparison of each solution structure with its respective ensemble average (Table S3), residues of the 2LZN ensemble structures in the Ramachandran plot disallowed regions (Table S4), and mean order for secondary structure elements (Table S5), and equivalent interface residues in the DnaB N-terminal domain of *G. stearothermophilus* and *S. aureus* (Table S6) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*Department of Chemistry, University of Nebraska—Lincoln, 722 Hamilton Hall, Lincoln, NE 68588-0304. Phone: 402-472-3039. E-mail: rpowers3@unl.edu.

*Department of Chemistry, University of Nebraska—Lincoln, 736 Hamilton Hall, Lincoln, NE 68588-0304. Phone: 402-472-3429. E-mail: mgriep1@unl.edu.

ORCID

Robert Powers: 0000-0001-9948-6837

Mark A. Griep: 0000-0002-7967-5822

Author Contributions

J.C. and J.P. contributed equally to this work.

Funding

This work was supported by grants from the UNL/UNMC Collaborative Research Fund and the Layman Foundation Fund to M.A.G. and by the National Institute of Allergy and Infectious Diseases (R21AI081154), Nebraska Tobacco Settlement Biomedical Research Development Funds, Nebraska EPSCoR, the Maude Hammond Fling Faculty Research Fellowship, and a Nebraska Research Council Interdisciplinary Research Grant to R.P.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was performed in facilities renovated with support from the National Institutes of Health (RR015468-01). The authors thank the Rocky Mountain Regional 900 MHz NMR Facility for contributing NMR data to this project.

■ REFERENCES

(1) Griep, M. A., and Periago, J. (2017) Primase. In *Reference Module in Life Sciences*, pp 1–6, Elsevier, New York.

(2) Frick, D. N., and Richardson, C. C. (2001) DNA Primases. *Annu. Rev. Biochem. 70*, 39–80.

(3) Johnson, S. K., Bhattacharyya, S., and Griep, M. A. (2000) DnaB Helicase Stimulates Primer Synthesis Activity on Short Oligonucleotide Templates. *Biochemistry 39*, 736–744.

(4) Chintakayala, K., Larson, M. A., Grainger, W. H., Scott, D. J., Griep, M. A., Hinrichs, S. H., and Soultanas, P. (2007) Domain Swapping Reveals that the C- and N-Terminal Domains of DnaG and DnaB, Respectively, are Functional Homologues. *Mol. Microbiol. 63*, 1629–1639.

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932–943

(5) Bhattacharyya, S., and Griep, M. A. (2000) DnaB Helicase Affects the Initiation Specifity of *Escherichia coli* Primase on Single-Stranded DNA Templates. *Biochemistry 39*, 745−752.
(6) van Eijk, E., Paschalis, V., Green, M., Friggen, A. H., Larson, M. A., Spriggs, K., Briggs, G. S., Soultanas, P., and Smits, W. K. (2016) Primase Is Required for Helicase Activity and Helicase Alters the Specificity of Primase in the Enteropathogen Clostridium difficile. *Open Biol. 6*, 160272.
(7) Koepsell, S. A., Larson, M. A., Frey, C. A., Hinrichs, S. H., and Griep, M. A. (2008) *Staphylococcus aureus* Primase Has Higher Initiation Specificity, Interacts with Single-Stranded DNA Stronger, but Is Less Stimulated by Its Helicase than *Escherichia coli* Primase. *Mol. Microbiol. 68*, 1570−1582.
(8) Keck, J. L., and Berger, J. M. (2001) Primus Inter Pares (First Among Equals). *Nat. Struct. Biol. 8*, 2−4.
(9) Syson, K., Thirlway, J., Hounslow, A. M., Soultanas, P., and Waltho, J. P. (2005) Solution Structure of the Helicase-Interaction Domain of the Primase DnaG: A Model for Helicase Activation. *Structure 13*, 609−616.
(10) Koepsell, S. A., Larson, M. A., Griep, M. A., and Hinrichs, S. H. (2006) *Staphylococcus aureus* Helicase but not *Escherichia coli* Helicase Stimulates *S. aureus* Primase Activity and Maintains Initiation Specificity. *J. Bacteriol. 188*, 4673−4680.
(11) Urlacher, T. M., and Griep, M. A. (1995) Magnesium Acetate Induces a Conformational Change in Escherichia coli Primase. *Biochemistry 34*, 16708−16714.
(12) Tougu, K., and Marians, K. J. (1996) The Extreme C Terminus of Primase is Required for Interaction with DnaB at the Replication Fork. *J. Biol. Chem. 271*, 21391−21397.
(13) Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., and Bateman, A. (2016) The Pfam Protein Families Database: Towards a More Sustainable Future. *Nucleic Acids Res. 44*, D279−285.
(14) Corn, J. E., Pease, P. J., Hura, G. L., and Berger, J. M. (2005) Crosstalk Between Primase Subunits Can Act to Regulate Primer Synthesis in Trans. *Mol. Cell 20*, 391−401.
(15) Podobnik, M., McInerney, P., O'Donnell, M., and Kuriyan, J. (2000) A TOPRIM Domain in the Crystal Structure of the Catalytic Core of *Escherichia coli* Primase Confirms a Structural Link to DNA Topoisomerases. *J. Mol. Biol. 300*, 353−362.
(16) Aravind, L., Leipe, D. D., and Koonin, E. V. (1998) Toprim–A Conserved Catalytic Domain in Type IA and II Topoisomerases, DnaG-type Primases, OLD Family Nucleases and RecR Proteins. *Nucleic Acids Res. 26*, 4205−4213.
(17) Soultanas, P. (2005) The Bacterial Helicase-Primase Interaction: A Common Structural/Functional Module. *Structure 13*, 839−844.
(18) Pan, H., and Wigley, D. B. (2000) Structure of the Zinc-Binding Domain of *Bacillus stearothermophilus* DNA Primase. *Structure 8*, 231−239.
(19) Bailey, S., Eliason, W. K., and Steitz, T. A. (2007) Structure of Hexameric DnaB Helicase and its Complex with a Domain of DnaG Primase. *Science 318*, 459−463.
(20) Su, X. C., Schaeffer, P. M., Loscha, K. V., Gan, P. H., Dixon, N. E., and Otting, G. (2006) Monomeric Solution Structure of the Helicase-Binding Domain of *Escherichia coli* DnaG Primase. *FEBS J. 273*, 4997−5009.
(21) Oakley, A. J., Loscha, K. V., Schaeffer, P. M., Liepinsh, E., Pintacuda, G., Wilce, M. C., Otting, G., and Dixon, N. E. (2005) Crystal and Solution Structures of the Helicase-Binding Domain of *Escherichia coli* Primase. *J. Biol. Chem. 280*, 11495−11504.
(22) Abdul Rehman, S. A., Verma, V., Mazumder, M., Dhar, S. K., and Gourinath, S. (2013) Crystal Structure and Mode of Helicase Binding of the C-Terminal Domain of Primase from *Helicobacter pylori*. *J. Bacteriol. 195*, 2826−2838.
(23) Larson, E. (2007) Community Factors in the Development of Antibiotic Resistance. *Annu. Rev. Public Health 28*, 435−447.
(24) Mercier, K. A., Germer, K., and Powers, R. (2006) Design and Characterization of a Functional Library for NMR Screening Against Novel Protein Targets. *Comb. Chem. High Throughput Screening 9*, 515−534.
(25) Shortridge, M. D., Griep, M. A., and Powers, R. (2012) 1)H, (1)(3)C, and (1)(5)N NMR Assignments for the Helicase Interaction Domain of *Staphylococcus aureus* DnaG Primase. *Biomol. NMR Assignments 6*, 35−38.
(26) Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear Multidimensional NMR Experiments for the Structure Determination of Proteins in Solution Employing Pulsed Field Gradients. *Prog. Nucl. Magn. Reson. Spectrosc. 34*, 93−158.
(27) Hwang, T. L., Mori, S., Shaka, A. J., and vanZijl, P. C. M. (1997) Application of Phase-Modulated CLEAN Chemical EXchange Spectroscopy (CLEANEX-PM) to Detect Water-Protein Proton Exchange and Intermolecular NOEs. *J. Am. Chem. Soc. 119*, 6203−6204.
(28) Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: A Hybrid Method for Predicting Protein Backbone Torsion Angles from NMR Chemical Shifts. *J. Biomol. NMR 44*, 213−223.
(29) Vuister, G. W., Wang, A. C., and Bax, A. (1993) Measurement of 3-Bond Nitrogen Carbon-J Couplings in Proteins Uniformly Enriched in N-15 and C-13. *J. Am. Chem. Soc. 115*, 5334−5335.
(30) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR 6*, 277−293.
(31) Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) A Common-Sense Approach to Peak Picking in 2-Dimensional, 3-Dimensional, and 4-Dimensional Spectra Using Automatic Computer-Analysis of Contour Diagrams. *J. Magn. Reson. 95*, 214−220.
(32) Fogh, R., Ionides, J., Ulrich, E., Boucher, W., Vranken, W., Linge, J. P., Habeck, M., Rieping, W., Bhat, T. N., Westbrook, J., Henrick, K., Gilliland, G., Berman, H., Thornton, J., Nilges, M., Markley, J., and Laue, E. (2002) The CCPN Project: An Interim Report on a Data Model for the NMR Community. *Nat. Struct. Biol. 9*, 416−418.
(33) Huang, Y. J., Tejero, R., Powers, R., and Montelione, G. T. (2006) A Topology-Constrained Distance Network Algorithm for Protein Structure Determination from NOESY Data. *Proteins: Struct., Funct., Genet. 62*, 587−603.
(34) Clore, G. M., Appella, E., Yamada, M., Matsushima, K., and Gronenborn, A. M. (1990) Three-Dimensional Structure of Interleukin 8 in Solution. *Biochemistry 29*, 1689−1696.
(35) Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR Molecular Structure Determination Package. *J. Magn. Reson. 160*, 65−73.
(36) Schwieters, C. D., Kuszewski, J. J., and Clore, G. M. (2006) Using Xplor-NIH for NMR Molecular Structure Determination. *Prog. Nucl. Magn. Reson. Spectrosc. 48*, 47−62.
(37) Garrett, D. S., Kuszewski, J., Hancock, T. J., Lodi, P. J., Vuister, G. W., Gronenborn, A. M., and Clore, G. M. (1994) The Impact of Direct Refinement Against Three-Bond HN-CalphaH Coupling Constants on Protein Structure Determination by NMR. *J. Magn. Reson., Ser. B 104*, 99−103.
(38) Kuszewski, J., Qin, J., Gronenborn, A. M., and Clore, G. M. (1995) The Impact of Direct Refinement Against 13Calpha and 13Cbeta Chemical Shifts on Protein Structure Determination by NMR. *J. Magn. Reson., Ser. B 106*, 92−96.
(39) Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1996) Improving the Quality of NMR and Crystallographic Protein Structures by Means of a Conformational Database Potential Derived from Structure Databases. *Protein Sci. 5*, 1067−1080.
(40) Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1997) Improvements and Extensions in the Conformational Database Potential for the Refinement of NMR and X-ray Structures of Proteins and Nucleic Acids. *J. Magn. Reson. 125*, 171−177.
(41) Kuszewski, J., and Clore, G. M. (2000) Sources of and Solutions to Problems in the Refinement of Protein NMR Structures Against Torsion Angle Potentials of Mean Force. *J. Magn. Reson. 146*, 249−254.

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943

(42) Powers, R., Mirkovic, N., Goldsmith-Fischman, S., Acton, T. B., Chiang, Y., Huang, Y. J., Ma, L., Rajan, P. K., Cort, J. R., Kennedy, M. A., Liu, J., Rost, B., Honig, B., Murray, D., and Montelione, G. T. (2005) Solution Structure of *Archaeglobus fulgidis* Peptidyl-tRNA Hydrolase (Pth2) Provides Evidence for an Extensive Conserved Family of Pth2 Enzymes in Archea, Bacteria, and Eukaryotes. *Protein Sci. 14*, 2849−2861.
(43) Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR System. *Nat. Protoc. 2*, 2728−2733.
(44) Linge, J. P., and Nilges, M. (1999) Influence of Non-Bonded Parameters on the Quality of NMR Structures: A New Force Field for NMR Structure Calculation. *J. Biomol. NMR 13*, 51−59.
(45) Nederveen, A. J., Doreleijers, J. F., Vranken, W., Miller, Z., Spronk, C. A., Nabuurs, S. B., Guntert, P., Livny, M., Markley, J. L., Nilges, M., Ulrich, E. L., Kaptein, R., and Bonvin, A. M. (2005) RECOORD: A Recalculated Coordinate Database of 500+ Proteins from the PDB Using Restraints from the BioMagResBank. *Proteins: Struct., Funct., Genet. 59*, 662−672.
(46) Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone Dynamics of Proteins as Studied by 15N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry 28*, 8972−8979.
(47) Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone Dynamics of a Free and Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15N NMR Relaxation. *Biochemistry 33*, 5984−6003.
(48) Mandel, A. M., Akke, M., and Palmer, A. G., III (1995) Backbone Dynamics of *Escherichia coli* Ribonuclease HI: Correlations with Structure and Function in an Active Enzyme. *J. Mol. Biol. 246*, 144−163.
(49) Cole, R., and Loria, J. P. (2003) FAST-Modelfree: A Program for Rapid Automated Analysis of Solution NMR Spin-Relaxation Data. *J. Biomol. NMR 26*, 203−213.
(50) Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules 0.1. Theory and Range of Validity. *J. Am. Chem. Soc. 104*, 4546−4559.
(51) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera–A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem. 25*, 1605−1612.
(52) Huson, D. H., and Bryant, D. (2006) Application of Phylogenetic Networks in Evolutionary Studies. *Mol. Biol. Evol. 23*, 254−267.
(53) Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235−242.
(54) Edgar, R. C. (2004) MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res. 32*, 1792−1797.
(55) Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A Novel Method for Fast and Accurate Multiple Sequence Alignment. *J. Mol. Biol. 302*, 205−217.
(56) Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol. 7*, 539.
(57) Mercier, K. A., Baran, M., Ramanathan, V., Revesz, P., Xiao, R., Montelione, G. T., and Powers, R. (2006) FAST-NMR: Functional Annotation Screening Technology Using NMR Spectroscopy. *J. Am. Chem. Soc. 128*, 15292−15299.
(58) Hwang, T. L., and Shaka, A. J. (1995) Water Suppression That Works - Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J. Magn. Reson., Ser. A 112*, 275−279.
(59) Johnson, P. E., Tomme, P., Joshi, M. D., and McIntosh, L. P. (1996) Interaction of Soluble Cellooligosaccharides with the N-Terminal Cellulose-Binding Domain of Cellulomonas fimi CenC. 2. NMR and Ultraviolet Absorption Spectroscopy. *Biochemistry 35*, 13895−13906.
(60) Barril, X. (2013) Druggability Predictions: Methods, Limitations, and Applications. *WIREs Comput. Mol. Sci. 3*, 327−338.
(61) Clore, G. M., Nilges, M., Sukumaran, D. K., Brunger, A. T., Karplus, M., and Gronenborn, A. M. (1986) The Three-Dimensional Structure of Alpha1-purothionin in Solution: Combined Use of Nuclear Magnetic Resonance, Distance Geometry and Restrained Molecular Dynamics. *EMBO J. 5*, 2729−2735.
(62) Nilges, M., Gronenborn, A. M., Brunger, A. T., and Clore, G. M. (1988) Determination of Three-Dimensional Structures of Proteins by Simulated Annealing with Interproton Distance Restraints. Application to Crambin, Potato Carboxypeptidase Inhibitor and Barley Serine Proteinase Inhibitor 2. *Protein Eng., Des. Sel. 2*, 27−38.
(63) Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) Charmm - A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem. 4*, 187−217.
(64) Halouska, S., Zhou, Y., Becker, D. F., and Powers, R. (2009) Solution Structure of the *Pseudomonas putida* Protein PpPutA45 and its DNA Complex. *Proteins: Struct., Funct., Genet. 75*, 12−27.
(65) Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) Procheck - A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr. 26*, 283−291.
(66) Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) Evaluating Protein Structures Determined by Structural Genomics Consortia. *Proteins: Struct., Funct., Genet. 66*, 778−795.
(67) Shortridge, M. D., Triplet, T., Revesz, P., Griep, M. A., and Powers, R. (2011) Bacterial Protein Structures Reveal Phylum Dependent Divergence. *Comput. Biol. Chem. 35*, 24−33.
(68) Motlagh, H. N., Wrabl, J. O., Li, J., and Hilser, V. J. (2014) The Ensemble Nature of Allostery. *Nature 508*, 331−339.
(69) Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry Part II: Techniques for the Study of Biological Strcuture and Function*, pp 461, W. H. Freeman and Co., San Francisco.
(70) Crute, J. J., Grygon, C. A., Hargrave, K. D., Simoneau, B., Faucher, A. M., Bolger, G., Kibler, P., Liuzzi, M., and Cordingley, M. G. (2002) Herpes Simplex Virus Helicase-Primase Inhibitors are Active in Animal Models of Human Disease. *Nat. Med. 8*, 386−391.
(71) Kleymann, G., Fischer, R., Betz, U. A., Hendrix, M., Bender, W., Schneider, U., Handke, G., Eckenberg, P., Hewlett, G., Pevzner, V., Baumeister, J., Weber, O., Henninger, K., Keldenich, J., Jensen, A., Kolb, J., Bach, U., Popp, A., Maben, J., Frappa, I., Haebich, D., Lockhoff, O., and Rubsamen-Waigmann, H. (2002) New Helicase-Primase Inhibitors as Drug Candidates for the Treatment of Herpes Simplex Disease. *Nat. Med. 8*, 392−398.
(72) Griep, M. A., Blood, S., Larson, M. A., Koepsell, S. A., and Hinrichs, S. H. (2007) Myricetin Inhibits *Escherichia coli* DnaB Helicase but not Primase. *Bioorg. Med. Chem. 15*, 7203−7208.
(73) Chintakayala, K., Larson, M. A., Griep, M. A., Hinrichs, S. H., and Soultanas, P. (2008) Conserved Residues of the C-terminal p16 Domain of Primase are Involved in Modulating the Activity of the Bacterial Primosome. *Mol. Microbiol. 68*, 360−371.

DOI: 10.1021/acs.biochem.6b01273
Biochemistry 2017, 56, 932−943