Biomol NMR Assign (2012) 6:35–38
DOI 10.1007/s12104-011-9320-7

ARTICLE

^{1}H, ^{13}C, and ^{15}N NMR assignments for the helicase interaction domain of *Staphylococcus aureus* DnaG primase

Matthew D. Shortridge · Mark A. Griep · Robert Powers

Received: 17 March 2011 / Accepted: 23 May 2011 / Published online: 7 June 2011

Abstract The interaction between DnaG primase and DnaB helicase is essential for stimulating primer synthesis during bacterial DNA replication. The interaction occurs between the N-terminal domain of helicase and the C-terminal domain of primase. Here we present the ^{1}H, ^{13}C, and ^{15}N backbone and side-chain resonance assignments for the C-terminal helicase interaction domain of *Staphylococcus aureus* primase.

Keywords *Staphylococcus aureus* · DnaG · Primase · Helicase · NMR · Resonance assignments

Biological context

Bacterial primase (DnaG) is a conserved and essential enzyme responsible for the synthesis of short RNA polymers during DNA replication (Frick and Richardson 2001). The protein is composed of three domains: N-terminal domain responsible for DNA sequence-specific binding; the catalytic core responsible for RNA synthesis; and the C-terminal domain (CTD) responsible for its interaction with DnaB helicase (Bird et al. 2000; Tougu and Marians 1996). The bacterial primase sequence is conserved along its entire length of approximately 570 residues (Frick and Richardson 2001; Pan and Wigley 2000; Syson et al. 2005). Conservation is highest in the N-terminal domain, intermediate in the catalytic core, and lowest in the CTD.

The solution structures of DnaG primase CTD show significant variability between *Geobacillus stearothermophilus* (PDB 1Z8S) and *Escherichia coli* (PDB 2HAJ) (Su et al. 2006; Syson et al. 2005). Generally, the DnaG primase CTD structure is composed of two sub-domains, an N-terminal six-helix bundle (sub-domain C1) essential for DnaB binding activity and a helical hairpin (sub-domain C2) that modulates the interaction with DnaB (Oakley et al. 2005; Syson et al. 2005). The two CTD solution structures share significant structure similarity in their C1 sub-domains, but there are sharp differences in their C2 sub-domains (Oakley et al. 2005; Su et al. 2006; Syson et al. 2005). In *E. coli* primase, the CTD structure is composed of 7 helices, where the long, rigid helix 6 connects the two sub-domains (Su et al. 2006). In contrast, the *G. stearothermophilus* structure has two helices connecting the two sub-domains, where a kink occurs at Pro556 (Syson et al. 2005).

The biological significance of these structural differences may provide an explanation for the observed divergence in primase at the level of bacterial phyla. For instance, *Firmicutes* DnaB helicases have been shown to stimulate a range of *Firmicutes* primases, but not *Proteobacterial* primases, and vice versa (Koepsell et al. 2006; Larson et al. 2010). This appears to be a general principle since we recently demonstrated that a wide range of bacterial enzymes show structural and functional divergence at the level of phyla (Shortridge et al. 2011). We undertook the determination of the solution structure of *Staphylococcus aureus* primase CTD because it belongs to the same

M. D. Shortridge
Department of Chemistry, University of Nebraska-Lincoln, 721 Hamilton Hall, Lincoln, NE 68588-0304, USA
e-mail: mds8575@huskers.unl.edu

M. A. Griep
Department of Chemistry, University of Nebraska-Lincoln, 736 Hamilton Hall, Lincoln, NE 68588-0304, USA
e-mail: mgriep1@unl.edu

R. Powers (✉)
Department of Chemistry, University of Nebraska-Lincoln, 722 Hamilton Hall, Lincoln, NE 68588-0304, USA
e-mail: rpowers3@unl.edu

phylum as *G. stearothermophilus*. We expect the *S. aureus* primase CTD structure to determine whether the bacterial primase CTD structure is also conserved within a phylum. Here we present the complete backbone and side chain assignments for the DnaG primase CTD of *S. aureus* to provide the basis for determining its solution structure.

Methods and experiments

Expression and purification of *S. aureus* DnaG primase CTD

The amplicon for *S. aureus* primase CTD gene (143 aa) was generated by polymerase chain reaction from the genome reference strain in the American Tissue Culture Collection by Nature Technology, Inc., Lincoln, Nebraska. It was inserted into a vector that added 20 N-terminal residues (MGHNHNHNHNHNGGDDDD) comprising an HN metal affinity tag and a protease-sensitive DDDD sequence. After transformation, the resulting *E. coli* overexpression strain had a doubling time of 15 min. Since this was a "leaky" expression system, this fast growth rate indicated the protein was not toxic and actually beneficial.

The protein used in the NMR structure determination experiments was uniformly labeled with ^{13}C and ^{15}N by growing the overexpression strain in minimal media containing ^{13}C-glucose and ^{15}N-ammonia. The expression level was extremely high with a final yield of 229 mg from a 2-l preparation. After the cells were collected by centrifugation, resuspended, treated with lysozyme, and then recentrifuged, the target protein was found entirely in the supernatant. The protein was affinity purified on a nickel-NTA column and eluted with 100 mM imidazole. A denaturing polyacrylamide gel showed the purified protein was primarily a single band of the expected size (19.6 kDa). Its concentration was determined by absorbance using the extinction coefficient, ε_{280}, of 14,080 M^{-1} cm^{-1}, derived from its 11 tyrosines.

Nuclear magnetic resonance (NMR) spectroscopy

Samples of [U-^{13}C,^{15}N] DnaG primase CTD for backbone resonance assignment by NMR spectroscopy were prepared at a protein concentration of 1.2 mM in 95% H_2O/5% D_2O solution containing 25 mM potassium phosphate, 100 mM NaCl, 50 mM glutamine and 50 mM arginine at pH 6.6 (uncorrected) in a sealed Shigemi tube (Shigemi Inc., Allison Park, PA). Samples of [U-^{13}C,^{15}N] DnaG primase CTD for side-chain resonance assignment by NMR spectroscopy were prepared at a protein concentration of 1.5 mM in 99% D_2O solution containing 25 mM potassium phosphate, 100 mM NaCl, 50 mM glutamine and 50 mM arginine at pH 6.6 (uncorrected) in a sealed Shigemi tube (Shigemi Inc., Allison Park, PA). All NMR experiments used for the protein backbone assignments of DnaG primase CTD were collected at 298 K on a five-channel 600 MHz Bruker Avance spectrometer equipped with a 5 mm TXI probe. NMR experiments used for the protein side chain resonances were collected 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. Assignments of the backbone and side chain resonances were obtained from the following spectra: 2D ^{1}H-^{15}N-HSQC, 2D ^{1}H-^{13}C-HSQC, HNCO, HNCA, CBCACONH, CBCANH, HNHA, HBHACONH, CCCONH, HCCCONH, H(CCH)-COSY, 3D ^{13}C-edited NOESY and 3D ^{15}N-edited NOESY (Ferentz and Wagner 2000). The NMR data was processed using NMRpipe (Delaglio et al. 1995) and analyzed using PIPP (Garrett et al. 1991) and CcpNmr v.2.1.5 (Vranken et al. 2005).

Assignments of primase CTD and data deposition

The backbone resonance assignment was 93% complete with 133 amino acids of the 143 unambiguously assigned in the 2D ^{1}H-^{15}N HSQC spectrum (Fig. 1). 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, but the N-terminal purification tag is simply numbered from residue 1–20. Residues M1–H13 and D19 comprise the majority of the unassigned residues and are all found within the N-terminal purification tag. The remaining unassigned residues were located within highly solvent-exposed regions. Residues H479, L480, and M481 were in a turn region between helix 1 and 2. Residue T500 was in an unstructured loop region between helix 2 and helix 3, whereas residues E543, P551 and Y552 were in an unstructured loop region between helix 5 and helix 6. The only residues that appear to be localized within secondary structure were R536 and E537. They are the second and third residues of helix 5. An exhaustive analysis of the NMR data set was unable to yield an assignment for these two residues, suggesting the end of the helix may undergo partial unfolding that leads to chemical shift exchange broadening.

Aliphatic side chain carbon chemical shift assignments were completed using the CCCONH experiment that correlated the preceding $(i-1)$ residue to the following (i) backbone amide chemical shift. Aliphatic side chain proton chemical shifts were completed with the HCCH-COSY and HCCCONH experiments. Aromatic side chain assignments were completed using the 3D ^{13}C-edited NOESY experiment. The statistics for resonance assignment include, 139/163 HN, 139/201 N, 139/163 Cα,

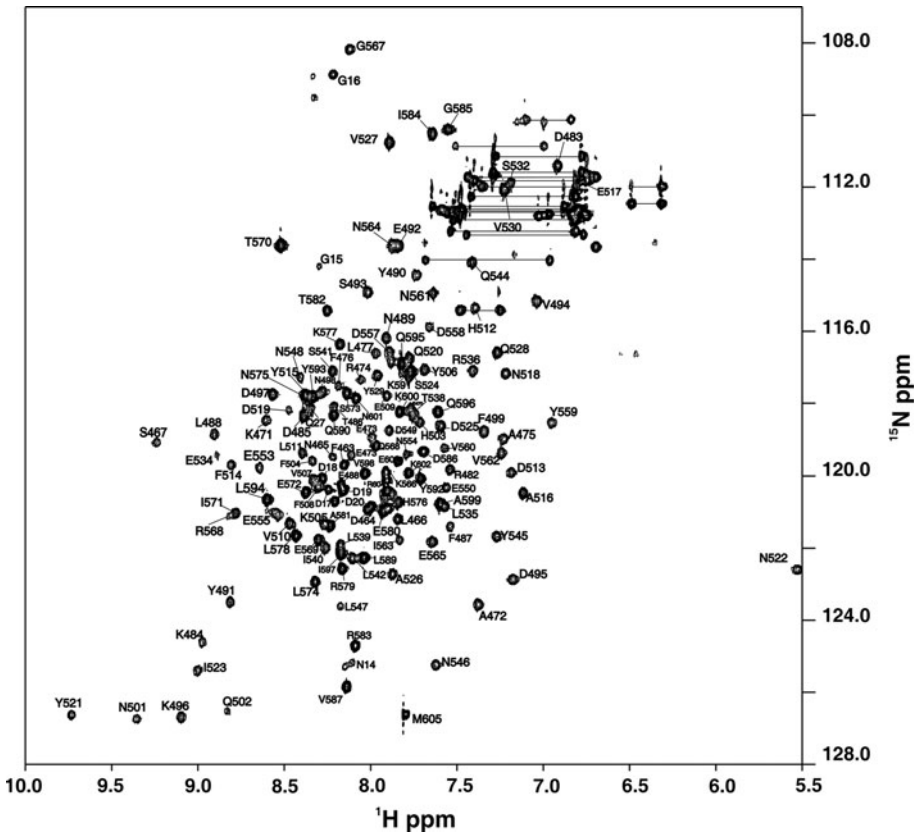

Fig. 1 The 2D ^{1}H–^{15}N HSQC spectrum of uniformly ^{13}C, ^{15}N labeled *S. aureus* primase CTD at 1.2 mM in 95% H_2O/5% D_2O at pH 6.6 (uncorrected). The spectrum was collected on a Bruker Avance 600 MHz spectrometer equipped with a 5 mm TXI probe. The backbone amide resonances are assigned. The side-chain amide groups of Asn and Gln are connected by *horizontal lines*

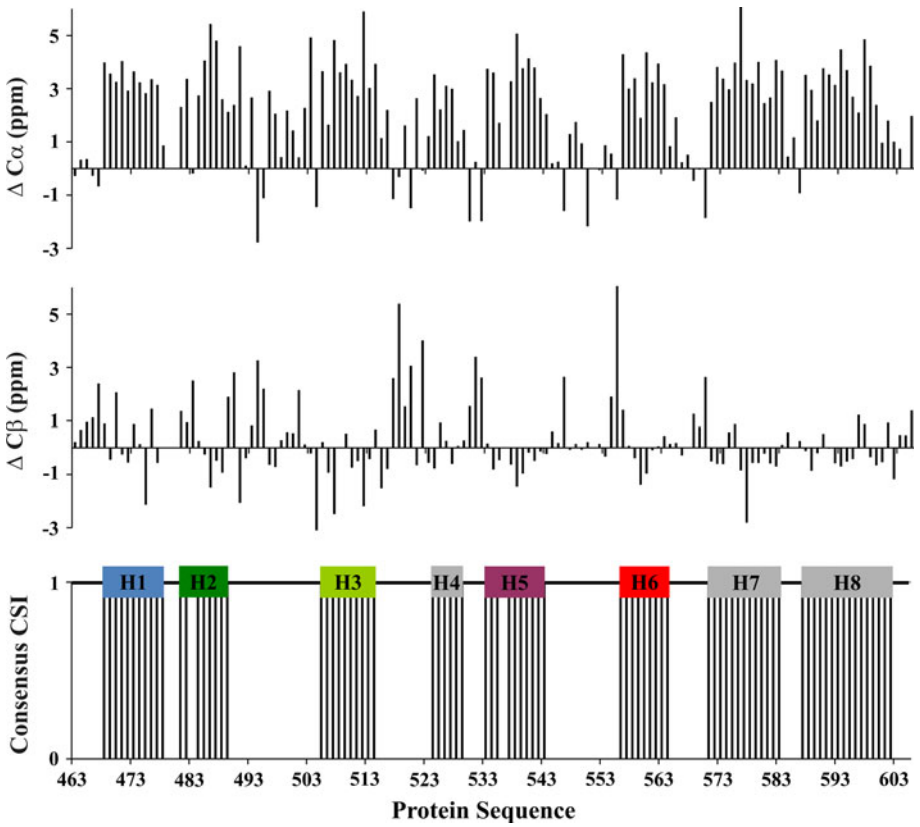

Fig. 2 The $^{13}C\alpha$ and $^{13}C\beta$ secondary structure chemical shifts are plotted as a function of sequence. The *S. aureus* primase CTD secondary structure is predicted based on the consensus chemical shift index (CSI) using the ^{13}C chemical shift data (Wishart and Sykes 1994). The secondary structure is shown above the chemical shift index where α-helices (*bars*) are numbered

134/168 Hα, 128/141 Cβ, 148/181 Hβ 85/92 Cγ, 89/160 Hγ, 49/64, Cδ, 49/64 Hδ, 19/31 Cε, 19/31 Hε, 4/7 C_ζ, 4/7 H_ζ, and 132/143 CO. Figure 1 shows an annotated 2D ^{1}H–^{15}N HSQC spectrum of the *S. aureus* helicase interaction domain of primase.

The *S. aureus* primase CTD secondary structure was predicted using the ^{13}C chemical shifts and a consensus chemical shift index (CSI) (Wishart and Sykes 1994). CSI predicts an all α-helical protein with 8 helices (Fig. 2). Helical structures in primase CTD include Helix 1 Arg468-Lys478, Helix 2 Met481-Asn489, Helix 3 Lys505-Phe514, Helix 4 Ser524-Tyr529, Helix 5 Asn533-Glu543, Helix 6 Ile556-Asn564, Helix 7 Ile571-Arg583, Helix 8 Val587-Lys602. The C1 sub-domain of primase CTD includes Helix 1–5 and the C2 sub-domain includes Helix 6–8. This is consistent with the *S. aureus* DnaG primase CTD homology modeling predicted from the *Geobacillus stearothermophilus* structure (Syson et al. 2005).

Of particular interest are the residues between the predicted helices 6 and 7 (residues K566–T570). This region is significantly different in *G. stearothermophilus* primase CTD compared to the *E. coli* primase CTD solution structures (PDB 1Z8S and 2HAJ, respectively) (Su et al. 2006). In *G. stearothermophilus,* this region is a loop forming two distinct sub-domains (C1, C2) of primase CTD. In *E. coli,* the region is a long and rigid α helix. For *S. aureus*, the consensus chemical shift index suggests this region is similar to the *G. stearothermophilus* structure with an extended loop starting at residue Gly567 (Fig. 2). The $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts for residues in this region have chemical shift values that are similar to those found in random coil structures.

The sequence specific ^{1}H, ^{13}C, and ^{15}N resonance assignments for *S. aureus* primase CTD were deposited in the BioMagResBank under accession number 17512.

Acknowledgments This work was supported in part by a grant from the UNL/UNMC Collaborative Research Fund to M.G. and grants from 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. The research was performed in facilities renovated with support from the NIH (RR015468-01). The authors would like to thank the Rocky Mountain Regional 900 MHz NMR Facility for contributing NMR data to this project.

Ethical standards The experiments comply with the current laws of the United States of America.

Conflict of interest The authors declare that they have no conflict of interest.

References

Bird LE, Pan H, Soultanas P, Wigley DB (2000) Mapping protein-protein interactions within a stable complex of DNA primase and DnaB helicase from bacillus stearothermophilus. Biochemistry 39:171–182

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293

Ferentz AE, Wagner G (2000) NMR spectroscopy: a multifaceted approach to macromolecular structure. Q Rev Biophys 33:29–65

Frick DN, Richardson CC (2001) DNA primases. Annu Rev Biochem 70:39–80

Garrett DS, Powers R, Gronenborn AM, Clore GM (1991) A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. J Magn Reson (1969–1992) 95:214–220

Koepsell SA, Larson MA, Griep MA, Hinrichs SH (2006) Staphylococcus aureus helicase but not Escherichia coli helicase stimulates *S. aureus* primase activity and maintains initiation specificity. J Bacteriol 188:4673–4680

Larson MA, Griep MA, Bressani R, Chintakayala K, Soultanas P, Hinrichs SH (2010) Class-specific restrictions define primase interactions with DNA template and replicative helicase. Nucleic Acids Res 38:7167–7178

Oakley AJ, Loscha KV, Schaeffer PM, Liepinsh E, Pintacuda G, Wilce MCJ, Otting G, Dixon NE (2005) Crystal and solution structures of the Helicase-binding domain of *Escherichia coli* primase. J Biol Chem 280:11495–11504

Pan H, Wigley DB (2000) Structure of the zinc-binding domain of *Bacillus stearothermophilus* DNA primase. Structure 8:231–239

Shortridge MD, Triplet T, Revesz P, Griep M, Powers R (2011) Bacterial protein structures reveal phylum dependent divergence. Comput Biol Chem 35:24–33

Su X-C, Schaeffer PM, Loscha KV, Gan PHP, Dixon NE, Otting G (2006) Monomeric solution structure of the helicase-binding domain of *Escherichia coli* DnaG primase. FEBS J 273:4997–5009

Syson K, Thirlway J, Hounslow AM, Soultanas P, Waltho JP (2005) Solution structure of the helicase-interaction domain of the primase DnaG: a model for helicase activation. Structure 13:609–616

Tougu K, Marians KJ (1996) The interaction between helicase and primase sets the replication fork clock. J Biol Chem 271: 21398–21405

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins Struct Funct Bioinf 59:687–696

Wishart DS, Sykes BD (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. J Biomol NMR 4:171–180