Chapter 3

STRUCTURAL INSIGHT INTO SUTA'S MECHANISM OF TRANSCRIPTIONAL REGULATION

3.1 Summary of Contributions

This work is a continuation of the work described in Chapter 2 and my collaboration with Dr. Megan Bergkessel continued accordingly. Contributions that were primarily my own include chemical cross-linking and foot-printing experiments, preparation and analysis of LC-MS/MS experiments, and analytical chromatography. We shared the following work equally: generation of plasmids and mutant strains, experimental planning, purification of SutA and RNA polymerase, data interpretation, and writing of the manuscript. We expect that follow-up experiments described at the end of this chapter will lead to a published manuscript.

3.2 Abstract

We recently reported the discovery of SutA, a small, acidic transcription factor expressed during slow growth in *Pseudomonas aeruginosa*. The primary sequence of SutA does not match any characterized protein domains and nothing is known about SutA's mechanism of transcriptional regulation. We performed a series of *in vitro* cross-linking and protein foot-printing experiments to investigate the structural interaction between SutA and the RNAP complex. We find that SutA interacts with the β and β' RNAP subunits in the presence and absence of nucleic acids, and through this interaction, is involved in a conformational change of RNAP. We propose a model for SutA's mechanism of transcriptional regulation that incorporates these new structural data and the previously observed affects of SutA on the transcription of ribosomal RNA and ribosomal protein genes.

3.3 Introduction

We previously described the discovery of SutA, a new type of transcriptional regulator in *Pseudomonas aeruginosa* [1]. SutA is a small, acidic protein that binds RNA polymerase (RNAP). The physiological role of SutA is broad; its expression is upregulated in slow growth conditions, and it enhances biofilm formation, regulates the production of phenazines, and contributes to survival during fluctuating conditions. SutA's effects on transcription are also broad; SutA colocalizes with RNA polymerase throughout much of the chromosome, and exhibits particularly strong association with the genes encoding ribosomal components [ribosomal RNA (rRNA) and ribosomal proteins (rProteins)]. Based on chromatin immunoprecipitation sequencing (ChIP-Seq) and RNA-Seq measurements, SutA generally has a positive influence on the transcription of genes with which it associates (e.g., rRNA are two-fold less abundant in the absence of SutA). However, nothing is known about SutA's mechanism of transcriptional regulation.

The transcription of ribosomal components is the primary function of bacterial RNAP. Under growth conditions, rRNA makes up an estimated 85% of cellular RNA, and rProteins 10-20% of cellular protein [2]. Accordingly, the regulation of this transcription in response to changing environmental conditions is critical and complex. rRNA transcription in particular is subject to a variety of regulatory mechanisms, including the binding of protein factors to DNA upstream of promoter regions (e.g., Fis and H-NS), the direction of RNAP to rRNA promoters by σ-factors, regulation of RNAP initiation and elongation steps by small proteins (e.g., DksA) and small molecules [e.g., guanosine tetraphosphate (ppGpp)] involved in the stringent response, and by "antitermination," which prevents premature release of RNAP during elongation of rRNA operons (reviewed in [3, 4]). rProtein transcription is also controlled by the stringent response regulators DksA and ppGpp [5], but the consequent translation of rProtein transcripts adds an additional layer of regulation to those genes, including recent evidence for translation-coupled transcriptional regulation of elongation [6].

The set of small proteins that bind RNAP and modulate its activity is varied. These include proteins that bind the RNAP secondary channel like the well-studied stringent response regulator DksA and its homologs (the *P. aeruginosa* genome encodes four DksA-like proteins), GreA and GreB, and Rnk [7]; the elongation and antitermination Nus factors; and a variety of other factors like the nonessential δ subunit from *Bacillus subtilis* [8], the small acidic protein AtfA from *Acinetobacter*

spp. [9], and the widely conserved transcription factor CarD [10].

To better understand the mechanism by which SutA regulates transcription and to place SutA in the context of known transcriptional regulators, we sought to establish a structural model of its interaction with RNAP. We performed a set of *in vitro* protein cross-linking and foot-printing experiments to find the domains of the RNAP complex to which SutA binds. We use these and other *in vitro* experiments to provide evidence for the role of SutA as a modulator of RNAP conformation, and suggest a model by which SutA enhances transcription during periods of slow growth. We finish with recommendations for future work.

3.4 Results

SutA Interacts with the β and β' Subunits of RNAP.

SutA lacks sequence similarity to any characterized proteins or domains, so the molecular mechanism by which it affects transcription is difficult to predict. It is a small (11.2 kDa) protein with a striking number of acidic residues, particularly in the N-terminal third of its sequence. Of its 105 amino acids, 32 are negatively charged and only 13 positively charged, giving a predicted pI of 3.87 and predicted charge of -19 at pH 7. Much of the protein is predicted to be disordered [11] and sequence-based structural analysis with Phyre [12] predicts a short alpha helical domain between residues Ala^{59} and Ser^{76} (Figure 3.S1A). Circular dichroism (CD) measurements of purified SutA are consistent with these predictions, with strong local minima observed near 200 nm and at 222 nm, corresponding to unstructured and α-helical conformations, respectively (Figures 3.S1B). To determine the regions of the RNAP complex with which SutA interacts, we undertook two *in vitro* cross-linking approaches: an unbiased cross-linking approach using bis(sulfosuccinimidyl) suberate (BS^3), and a directed approach that relies on the site-selective incorporation of the UV-activated amino acid cross-linker L-benzoylphenylalanine (Bpa) (Figure 3.1A).

BS^3 is a homobifunctional cross-linker, with two amine-reactive NHS-esters bridged by a hexamethylene linker. BS^3 reacts specifically with terminal amines on lysine side chains or protein N-termini that are in close proximity within a complex. This approach has previously been applied to study the structure of protein complexes, including eukaryotic RNA Pol II [13]. We added BS^3 to a mixture of purified SutA and RNAP in a 10:1 molar ratio. SDS-PAGE analysis showed reduction in intensity of the SutA and RNAP subunit bands and the appearance of high molecular weight complexes upon cross-linker addition (Figure 3.1B). With or without SutA, two

large cross-linked products were observed: one that did not migrate through the gel, and a large, but mobile product that likely corresponds to a single RNAP complex. Gel intensity analysis revealed that the addition of SutA to the cross-linking reaction resulted in preferential formation of the smaller product (70% with SutA vs. 35% without) (Figure 3.S2A).

Cross-linked complexes were digested and analyzed via liquid chromatography-tandem mass spectrometry (LC-MS/MS) and spectra were searched for matches to cross-linked peptides. Following stringent filtering of spectra from two independent experiments (see Materials and Methods, Figure 3.S2B-C), we identified 16 cross-links: nine between RNAP residues, three between SutA residues, and three between SutA and RNAP (Table 3.1). The three cross-links between SutA and RNAP were between Lys^{62} or Lys^{69} of SutA and Lys^{116} of the β subunit and between Lys^{95} of SutA and Lys^{40} of the β' subunit. Each was identified via high-quality matches of MS/MS spectra to fragment ion masses from each peptide, including fragment ions that contained the cross-linked sites (Figure 3.S3). β Lys^{116} and β' Lys^{40} are located in the "β lobe 1" and the "β' clamp" domains, respectively, which lie on opposite sides of the clamp that closes around downstream template DNA during transcription elongation [14].

Due to its reaction with amines and its relatively long linker (11.4 Å), cross-linking with BS^3 can be performed with wild-type proteins and allows for the discovery of protein domains that are in close proximity but not necessarily in direct contact. However, the reaction relies on lysine residues which are conspicuously absent from much of SutA's sequence. We therefore took a complementary, targeted cross-linking approach using the cross-linking amino acid Bpa. Bpa can be site-selectively incorporated into a protein by stop-codon reassignment using a mutant aminoacyl-tRNA synthetase/tRNA pair [15]. Upon UV activation, the diphenyl ketone can react nonspecifically with a nearby carbon. Because Bpa is a linker-free cross-linker and due to the short-lived radical intermediate, we expect Bpa cross-links to reflect short-range interactions between SutA and RNAP.

We purified nine SutA mutants with Bpa incorporated at positions throughout the protein sequence (Leu^6, Leu^{11}, Leu^{22}, Leu^{54}, Gln^{61}, Phe^{74}, Ile^{84}, Val^{89}, or Tyr^{100}). We mixed each SutA mutant with RNAP under the same conditions used for the BS^3 cross-linking (10:1 molar ratio). We exposed mixtures to 365 nm light, separated resulting species via SDS-PAGE, and looked for RNAP subunits with increased mass corresponding to cross-linking to SutA. We observed cross-linking between SutA

Table 3.1: **BS^3 cross-linked peptides.** Cross-links detected following BS^3 cross-linking of RNAP-SutA complexes. The first four columns indicate the proteins and amino acid positions involved in each cross-link. Evidence columns represent the number of spectra matched to each cross-linked amino acid pair in each replicate. Maximum Score Difference (SD) is reported as calculated by Protein Prospector. Distance refers to the inter α-carbon distance. Cross-links between SutA and RNAP are bolded.

Protein 1	**AA 1**	**Protein 2**	**AA 2**	**Ev. Rep. 1**	**Ev. Rep. 2**	**Max. SD**	**Dist. (Å)**
RpoB	116	RpoB	481	NA	8	32.8	18.4
SutA	95	SutA	99	2	NA	30.9	NA
RpoB	**116**	**SutA**	**69**	**5**	**9**	**26.1**	NA
RpoB	265	RpoB	284	5	1	23.9	17.2
SutA	69	SutA	80	31	6	22.8	NA
RpoB	**116**	**SutA**	**62**	**NA**	**18**	**17.3**	NA
RpoB	1144	RpoB	1160	9	20	15.3	13.3
RpoB	284	RpoC	1047	NA	3	14.6	40.1
RpoB	650	RpoC	678	NA	2	14.5	27.6
RpoB	1144	RpoB	1215	5	NA	13.1	33.3
RpoC	50	RpoC	87	4	6	12.8	18.1
RpoC	**40**	**SutA**	**95**	**2**	**1**	**12.4**	NA
RpoB	600	RpoB	631	6	NA	10.9	17.7
RpoB	1154	RpoB	1160	5	20	9.7	8.3
SutA	60	SutA	69	1	4	8.4	NA
SutA	1	SutA	80	4	NA	8.3	NA

Leu^{54} and the RNAP β subunit in a UV dose-dependent manner (Figure 3.S4A) and did not observe cross-linking following incubation with wild-type SutA (Figure 3.1C, left lane) or in the absence of UV treatment for any SutA mutant (Figure 3.S4C). We surveyed the panel of SutA mutants and observed cross-linking between positions 54 and 84 to the β subunit with nearly complete yield, while positions 6, 11, and 22 cross-linked to both β and β' subunits with lower yield (Figure 3.1C). No mutants cross-linked to the α subunit (Figure 3.S4B), and the other SutA mutants tested showed minimal evidence of cross-linking to any RNAP subunit. Because we suspect SutA to interact with RNAP during transcription, we performed the experiment on a pre-formed complex of RNAP and a ssRNA-ssDNA dimer similar to one previously used to constrain RNAP in a transcription elongation complex [16]. Based on gel analysis, cross-linked products were qualitatively the same in the presence (Figure 3.1C) or absence (Figure 3.S4C) of the nucleic acids.

We cut cross-linked products from each lane, digested the proteins, and analyzed

peptides via LC-MS/MS. We searched for masses uniquely identified in each cross-linked sample and not found from a control sample of uncross-linked β and β' subunits. Because the radical-mediated cross-linking of Bpa is nonspecific, searching for particular modifications is challenging, but we were able to identify three cross-linked products (Table 3.2, Figure 3.S5). We detected cross-links between SutA Leu^{6}, Leu^{11}, and Leu^{22} and peptides from the β' subunit clamp domain.

Table 3.2: **Bpa cross-links** Modified peptides detected by MS following cross-linking of SutA Bpa mutants to RNAP.

SutA position	**RNAP subunit**	**Peptide**	**AA positions**
L6	RpoC	KGQLLNDE	156-163
L11	RpoC	QYFEALE	165-170
L22	RpoC	KRM(ox)LQE	296-301

To complement these cross-linking results, we performed a protein foot-printing experiment [17]. Samples containing RNAP with or without SutA were reacted with *N*-hydroxysuccinimidobiotin (NHS-biotin) (Figure 3.1A), which covalently modifies solvent-exposed lysines, leading to a biotin mass modification readily detectable by MS. Proteins were separated via SDS-PAGE and the β and β' subunits were digested and analyzed via LC-MS/MS. More than 80% of the lysines in each protein were identified and approximately half of those were also found in their biotinylated state. To identify regions of RNAP whose solvent accessibility changes in the presence of SutA, we compared the intensity of peptides containing modified lysines between the two experimental conditions (i.e., with or without SutA). We identified eight residues whose modification was reduced at least 1.5-fold in the presence of SutA (Figure 3.S6). We grouped these residues into three categories: residues found near SutA cross-links: Lys^{116} and Lys^{119} of the β subunit; residues along the main channel: Lys^{1257} and Lys^{1277} of the β subunit and Lys^{332} and Lys^{1231} of the β' subunit; and residues on the secondary channel face: Lys^{1215} of the β subunit and Lys^{996} of the β' subunit. We also found two residues with higher modified intensities in the presence of SutA: Lys^{207} of the β subunit along the main channel and Lys^{603} of the β' subunit on the secondary channel face. We note that a change in biotinylation of a given residue can indicate either SutA binding at that location or a change in RNAP conformation that alters the accessibility of that residue.

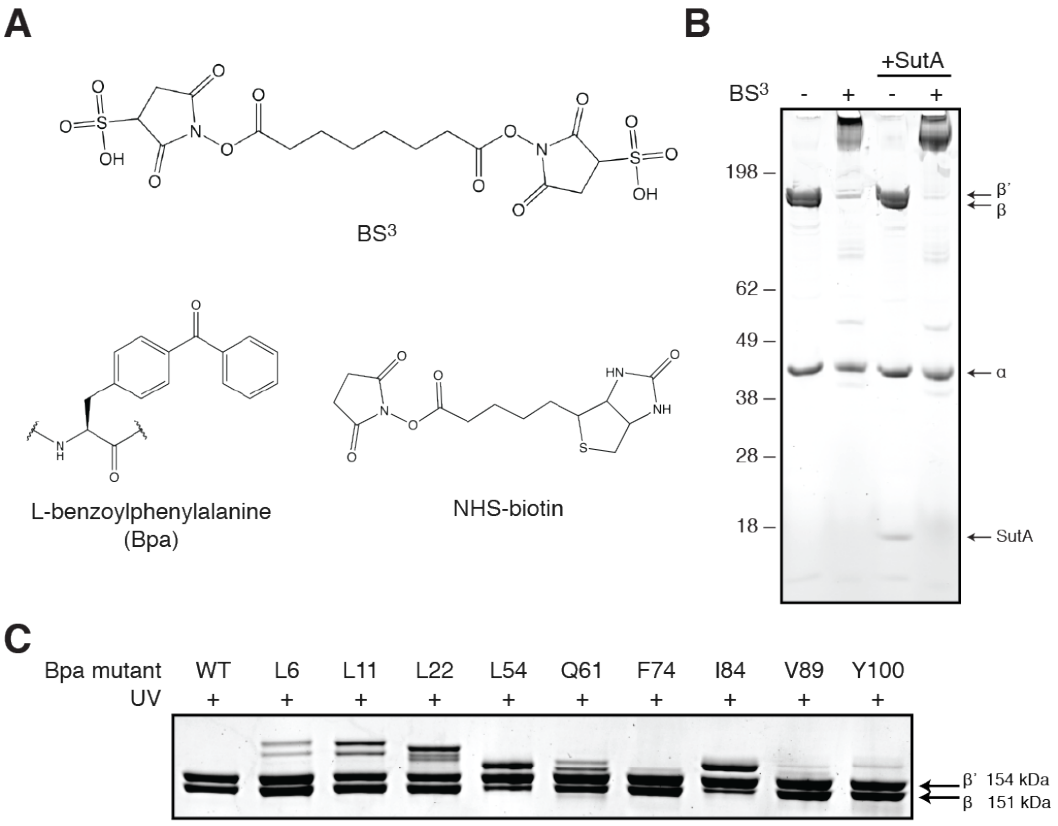

Figure 3.1: **SutA cross-links to RNAP *in vitro*.** (A) Chemicals used for cross-linking and foot-printing expriments. (B) SDS-PAGE analysis of RNAP and SutA before and after cross-linking with BS3. The gel is annotated with positions of molecular weight markers (kDa, left) and RNAP subunits (right). (C) Cross-linked RNAP subunits following incubation with SutA Bpa mutants and UV irradiation. The position of the SutA residue replaced by Bpa is indicated above.

SutA Binds Across the RNAP Clamp.

The structure of *P. aeruginosa* RNAP has not been solved, so to visualize cross-linking and foot-printing results, we mapped cross-linked residues to the SutA sequence (Figure 3.2A) and to homologous residues in the *Escherichia coli* RNAP structure (PDB: 3LU0, Figure 3.2B) [18]. The primary sequences of *E. coli* RNAP subunits are more than 84% similar to their *P. aeruginosa* homologs. Distances between α-carbons of intra-RNAP BS^3 cross-linked residues ranged from 8-40Å, a range similar to that observed for cross-links between RNA Pol II lysines in a previous BS^3 cross-linking experiment (Table 3.1, Figure 3.S7) [13]. Distances longer than expected based on the length of the cross-linker can be explained by (i) differences between the structures of the *E. coli* and *P. aeruginosa* complexes, (ii) dynamics of the complex in solution, or (iii) a conformational change of RNAP in

the presence of SutA.

The three BS^3 cross-links between SutA and RNAP (Figure 3.2, red) point to SutA's interaction with both sides of the RNAP clamp (β Lys^{116} and β' Lys^{40}) that closes around downstream DNA during elongation (forming the "closed clamp" or DNA "open complex"). The two BS^3 cross-links to the β lobe 1 coincide with two obscured residues detected in the foot-printing experiment: β Lys^{116} and β Lys^{116} (Figure 3.2B, blue). Taken together, these provide strong evidence for a direct interaction between the central region of SutA and the β lobe 1 domain. Near complete cross-linking of SutA Bpa mutants Leu^{54} and Ile^{84} to the β subunit further corroborate these results.

Additional obscured residues on the β and β' subunits lie along the RNAP main channel that accommodates downstream template DNA. If the C-terminal half of SutA bridges the clamp, the acidic, and likely unstructured N-terminus could be in position to obscure these residues. However, Bpa cross-links between positions in the N-terminal portion of SutA (Leu^6, Leu^{11}, and Leu^{22}) place these residues against the β' clamp, just outside of the main channel (Figure 3.2, green). Additionally, our previous observations that SutA associates with the chromosome across the gene coding regions [1] suggests that its interaction with RNAP can occur during elongation while DNA is present in the main channel. While these observations do not rule out the occupation of the main channel by the SutA N-terminus, they suggest the alternative hypothesis that main-channel lysines are instead obscured by a conformational change of RNAP that reduces their solvent accessibility.

SutA is Involved with a Conformational Change of RNAP.

To explore the possibility that SutA induces a conformational change of RNAP, we explored the intra-RNAP cross-links captured by our BS^3 experiment. A majority of the detected cross-links spanned residues separated by a distance equal to or less than expected by the length of the cross-linker and the lysine side chains (approx. 25 Å) (Table 3.1). Of the three cross-links that spanned larger distances, β Lys^{284}-β' Lys^{1047} (40 Å) and β Lys^{650}-β' Lys^{678} (28 Å), span the RNAP clamp (Figure 3.S7).

Following our initial purification of SutA and RNAP, to confirm that their interaction was preserved in the *in vitro* conditions, we evaluated binding using analytical size exclusion chromatography (SEC) of RNAP alone, SutA alone, or RNAP and SutA mixed together (Figure 3.3A). SDS-PAGE analysis of RNAP elution fractions from the mixed sample provided evidence for SutA binding under these conditions (Figure

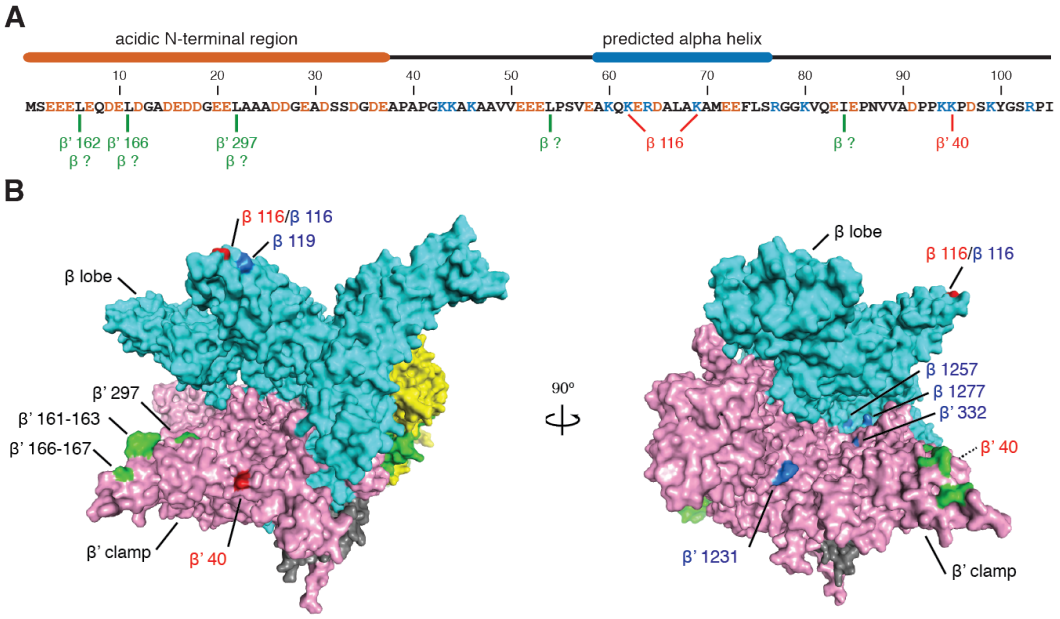

Figure 3.2: **SutA interacts with both sides of the RNAP clamp.** (A) SutA sequence, predicted structural domains, and residues found to cross-link to RNAP. Acidic residues are orange and basic residues are blue. Domains predicted by Phyre are indicated above. Residues found to cross-link to RNAP via BS^3 (red) or Bpa (green) are indicated below. Particular positions of cross-linking are indicated by RNAP subunit and residue position. Bpa cross-links for which interacting residues were not identified are shown with question marks. (B) Views of the RNAP core enzyme showing mapped locations of lysines involved in BS^3 cross-links to SutA (red), regions involved in Bpa cross-links (green), and lysines with reduced modification by NHS-biotin in the presence of SutA (blue). Subunits are colored as follows: α_I (yellow), α_{II} (green), β (cyan), β' (pink), and ω (gray). The structure was adapted from PDB:3LU0 [18].

3.3B); SutA was present in early eluent fractions only when RNAP was included. Interestingly, we noticed that RNAP in the presence of SutA exhibited later elution compared to the core enzyme alone (1.33 mL vs. 1.28 mL), suggestive of a possible conformational change of RNAP upon SutA binding. This change is consistent with a decrease in the Stokes radius (i.e., a more compact complex conformation). Similarly, we observed a difference in the preferred cross-linked product following BS^3 cross-linking of RNAP with and without SutA. We cannot unambiguously assign these gel bands to molecular species, but note that the inclusion of SutA led to a majority of the smaller product.

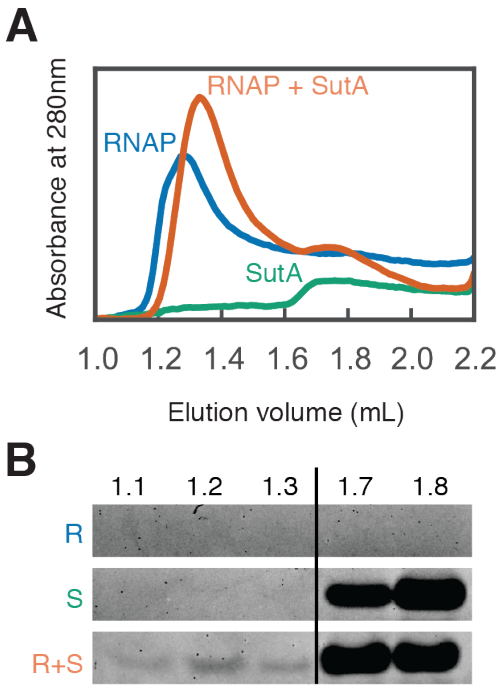

Figure 3.3: **SEC analysis of the SutA-RNAP complex.** (A) UV absorbance traces of size exclusion chromatography of RNAP (blue), SutA (green), and the mixed complex (orange). (B) Fractions (0.1 mL total volume, with start volumes listed above each lane) from each chromatographic separation in (A) were concentrated, separated via SDS-PAGE, and stained with Coomassie. Images show the SutA region of each gel.

The C-terminal, RNAP-binding Domain is Required for SutA Function.

To test whether the sites of interaction identified in the cross-linking experiments are relevant *in vivo*, we expressed SutA mutants composed of either the N-terminal or the C-terminal portion of SutA in a SutA deletion strain (Δ*sutA*) and evaluated each mutant's ability to phenocopy a strong phenotype of the deletion strain: overproduction of the small, colored phenazine, pyocyanin (PYO). We grew each strain in minimal medium and measured absorbance of the culture supernatant at 312 nm as a proxy for PYO abundance (Figure 3.4). The C-terminal fragment complemented the PYO overproduction phenotype of the Δ*sutA* strain while the acidic N-terminal fragment did not, suggesting that SutA's C-terminal portion, which contains all residues captured by BS^3 and the Bpa cross-links to the β subunit, is more important for SutA's function than the N-terminal portion.

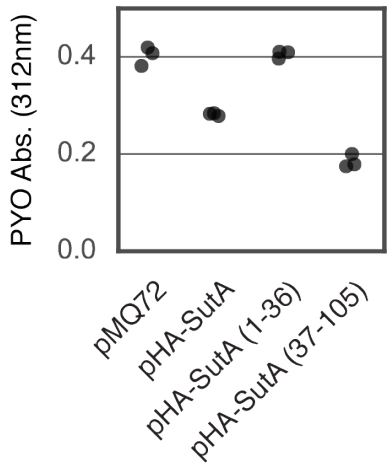

Figure 3.4: **Effects of SutA truncation on pyocyanin production.** The *sutA* deletion mutant was transformed with the indicated plasmids. Absorbance of each culture supernatant serves as a proxy for pyocyanin concentration.

3.5 Discussion

Here we provide structural context for the interaction between RNAP and SutA, a small, acidic, previously uncharacterized transcription factor. Previous work exploring SutA's physiological effects showed strong association of SutA with the loci encoding ribosomal genes, where it enhances their transcription. We also previously noted striking contrasts with the function of the small protein DksA that binds in the RNAP secondary channel. As part of the stringent response, DksA is involved in downregulating transcription of ribosomal components during nutrient downshifts. Mutation of *dksA* in *Pseudomonas spp.* has opposite effects on biofilm formation and pyocyanin production than we have observed in the *sutA* deletion strain.

DksA has complex allosteric effects on the activities of RNAP, so anticipating how SutA might cause opposing effects is difficult [19]. However, size exclusion chromatography indicates that SutA may affect RNAP conformation, and cross-linking and foot-printing results allow us to speculate about what these conformational changes might entail. The coincidence of the cross-linked (Lys^{116}) and obscured residues (Lys^{116} and Lys^{119}) on the β lobe 1 domain suggest that this is a site of direct interaction between the predicted alpha helical portion of SutA and RNAP. The cross-link between the C-terminal region of SutA and Lys^{40} in the β' clamp raises the possibility that SutA might bridge the cleft between the β lobe 1 and β' clamp domains. Though the distance between the cross-linked residues in these domains

is approximately 70 Å in the open clamp conformations of RNAP observed in most crystal structures [18, 20], the mobility of the β' clamp is well documented [21] and the main channel contracts upon DNA binding and transition to the transcription elongation complex [22]. Without atomic-level structural information about SutA and *P. aeruginosa* RNAP, we cannot determine whether SutA requires (or forces) a closed-clamp conformation to bridge this gap. However, our observation that lysine residues within the RNAP main channel are less accessible to chemical modification in the presence of SutA is consistent with clamp closure. Identification of Bpa cross-links between residues within the N-terminal domain of SutA to the β' clamp near, but outside the main channel suggest that the N-terminus itself does not obscure those residues. We interpret the partial cross-linking of Leu^{6}, Leu^{11}, Leu^{22} positions to both β and β' subunits to reflect the mobility of the presumably unstructured N-terminal acidic domain.

In addition to rotation of the clamp domain, other movements of large, mobile RNAP domains in the *E. coli* enzyme, such as the βi9 and β'i6 domains, have been observed or inferred [18, 23]. Several of the lysines for which the accessibility to chemical modification changes in the presence of SutA (β' Lys996 and Lys1231, which exhibit decreased modification in the presence of SutA; and β Lys207 and β' Lys603, which exhibit increased modification) are near or part of the β'i6 domain, suggesting that this domain could occupy a different position in the presence of SutA. This domain does not appear in any crystal structures, likely due to its mobility, so alternate conformations are not well characterized. However, it has been suggested that the position of the β'i6 domain can affect the ability of DksA to act on RNAP [23]. Likewise, an open clamp conformation is thought to facilitate DksA's interaction with polymerase [19, 24]. If SutA interacts preferentially with RNAP in a closed clamp conformation, this could contribute to the opposite effects on gene expression and phenotype observed between the *dksA* and *sutA* mutants. Future efforts to obtain additional SutA structural information and to characterize its effects on RNAP *in vitro* will help to distinguish among these possibilities.

A model of SutA as a protein that can alter the clamp position of RNAP is reminiscent of what has been observed for the elongation factor NusG [25], though their exact points of interaction are likely different. There are several other clear differences between the roles and activities of NusG and the potential roles for SutA. NusG is essential, and is critical for facilitating interactions between RNAP and other complexes [26], while SutA is nonessential, even under the conditions in

which it is upregulated, and appears to interact only with the polymerase. NusG is thought to be broadly recruited to most or all genes in *E. coli*, and accumulates as transcription along the gene proceeds [27], while SutA appears to be preferentially recruited to some genes more than others, and its peak of ChIP signal appears just slightly downstream of the peak of RNAP ChIP signal associated with the promoter. Nevertheless, our data suggest that SutA enhances expression of the genes to which it is recruited and one possibility is that like NusG, it does so primarily by enhancing transcription elongation at protein-coding genes (e.g., rProtein genes). At rRNA genes, in contrast, SutA associates primarily to the promoter region, suggesting that its activity may be different for these genes, and as with DksA, may have a primary role in regulating initiation. A key question for future studies will be whether and how SutA may interact or compete with a variety of other RNAP-binding regulators, including the rRNA antitermination complex or σ-factors, for example. More detailed structural resolution and measurements of the direct effects of SutA on RNAP activity will also be required to adequately test the proposed model.

3.6 Future Work

An atomic-level structure of SutA alone or in complex with RNAP would greatly enhance our understanding of its mechanism of transcriptional regulation. Thus far, attempts to obtain a co-crystal structure of SutA and RNAP have been unsuccessful. However, preliminary NMR measurements of SutA alone have been promising. Heteronuclear single quantum coherence (HSQC) measurements of ^{15}N-labeled SutA show well-dispersed ^{1}H $-^{15}$ N peaks (Figure 3.S8). Three-dimensional NMR measurements with ^{15}N, ^{13}C-labeled SutA should be useful in determining its solution structure. Additionally, experiments on labeled SutA mixed with RNAP may provide information about any SutA rearrangements that occur upon binding.

Our model of SutA as a modulator of the RNAP clamp brings with it predictions about specific effects on transcription. For example, at rRNA promoters, the RNAP-DNA open complex (clamp closed) is notoriously unstable; under rich nutrient conditions, this instability allows RNAP to move quickly off of promoter regions and proceed to elongation, while in nutrient-limited conditions, factors like DksA further destabilize the RNAP-DNA open complex at rRNA promoters and reduce transcription. The effect of SutA on open complex stability at the promoters of rRNA and other genes can be tested explicitly with *in vitro* transcription experiments. Furthermore, results from our ChIP experiments provide a list of other genomic loci to which SutA associates strongly (e.g., rProtein and other protein-coding genes).

Effects of SutA on RNAP behaviors such as elongation rate or premature termination can also be examined at these loci.

3.7 Experimental Procedures.

Strain construction. See Table 3.3 for a complete list of strains used. Cloning was performed using standard methods. Enzymes and supplies were purchased from New England Biolabs. For overexpression and purification, SutA was cloned from *P. aeruginosa* genomic DNA, appended with an N-terminal 6x Histidine tag followed by a TEV cleavage site (MRGSHHHHHHENLYFQS), and cloned into pQE80L (Qiagen) to generate DKN1697. Plasmids for overexpression of mutants for Bpa incorporation were generated from DKN1697 by replacing the codon at each indicated position with an amber stop codon (TAG) via PCR amplification and blunt-end ligation to create BMB14X, where X indicates the position in the amino acid sequence that encodes for Bpa. Plasmids for overexpression of SutA truncation mutants were generated from DKN1640 [1] by removing the coding region for the N-terminal (positions 1-36; DKN1688) or C-terminal (positions 37-105; DKN1687) amino acids via PCR amplification and blunt-end ligation.

Media and growth conditions. All cultures were grown at 37 °C with shaking unless otherwise noted. Liquid media were LB (5 g yeast extract, 10 g tryptone, 10 g NaCl per liter), 2xYT (10 g yeast extract, 16 g tryptone, and 5 g NaCl per liter), or M9 (12.8 g $Na_2HPO_4 \cdot 7\,H_2O$, 3 g KH_2PO_4, 0.5 g NaCl, 1 g NH_4Cl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$ per liter).

SutA purification. DKN1697 cells were grown in the presence of 200 µg ampicillin. A 20 mL culture grown overnight in LB was distributed between two flasks each containing one liter of 2xYT and grown to OD_{600} = 0.6. Protein expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and expression was allowed to continue for 4 h.

For Bpa incorporation, BMB14X cells were co-transformed with pEVOL-pBpF and approx. 20 colonies were scraped from the agar plate and grown at 33 °C in LB to a OD_{600} = 0.6. Cultures were treated with 1 mM Bpa (Iris-Biotech) and 1 mM IPTG and incubated in the dark for 20 h.

For ^{15}N labeled SutA, DKN1697 cells were grown for 8 h in 5 ml LB then added to 50 ml M9 medium containing 1 g/L $^{15}NH_4Cl$. These starter cultures were grown overnight, then added to 4x 1 L of the same M9 medium. Cells were grown to OD_{600} = 0.8 and protein expression was induced with 1 mM IPTG and allowed to

continue for 6 h.

For all, cells were pelleted and frozen at -80 °C. Pellets were resuspended in lysis buffer (40 mM NaH_2PO_4, 300 mM NaCl, pH 8) containing 5 mM imidazole, 1 mg/mL lysozyme, and cOmplete mini protease inhibitor, EDTA free and lysed by probe sonication. The lysate was treated with Benzonase Nuclease on ice for 30 min and centrifuged. Soluble protein was applied to His-Pur Ni-NTA (Thermo Scientific) washed three times with lysis buffer containing 20 mM imidazole and eluted three times with lysis buffer containing 250 mM imidazole. Eluents were combined, loaded onto an Amicon 10 kDa centrifugal filter (EMD Millipore), and buffer exchanged to TEV-digestion buffer (50 mM Tris pH 8.0, 0.5 mM EDTA, and 1 mM DTT). The 6xHis-tag was cleaved by addition of His-tagged TEV protease in a 1:50 mass ratio and incubation overnight at 4 °C. The digested sample was reapplied to His-Pur Ni-NTA, and washed with lysis buffer containing 20 mM imidazole; SutA eluted in this wash step, while the cleaved peptide tag and His-tagged TEV protease remained bound to the resin. The cleaved protein product includes the native SutA sequence with an additional N-terminal serine. SutA fractions were pooled and concentrated on an Amicon 10 kDa centrifugal filter, applied to a Superdex 75 10/300 column (GE Healthcare), buffer exchanged to SutA storage buffer (25 mM Tris pH 8, 100 mM NaCl, 20% glycerol, and 2 mM β-mercaptoethanol), and stored at -80 °C.

Due to the lower yield and additional contaminant proteins following expression in M9 medium, an additional ion exchange step was used to further purify SutA for NMR studies. Following TEV cleavage and removal of the tag, SutA was buffer exchanged into IEX buffer (20 mM *N*-methylpiperazine) with 100 mM NaCl and loaded onto a 5 mL Q FF column (GE Healthcare). Contaminants were washed with IEX buffer with 100 mM NaCl, and SutA was eluted via a linear gradient of IEX buffer from 100 to 600 mM NaCl. Purified protein was buffer exchanged into 10 mM Tris pH 7.0 with 100 mM NaCl via SEC as described above. ^{15}N incorporation was greater than 97% as verified by whole-protein MS.

***P. aeruginosa* RNAP purification.** RNAP was purified from the *P. aeruginosa* Δ*sutA* strain essentially as previously described ([28] and references therein). Cells were grown in 6 L of Terrific Broth medium to an OD_{600} of approximately 1.0. Cells were washed with TBS and pellets were frozen at -80 °C. Cell pellets were resuspended in 90 mL RNAP lysis buffer (50 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA, and cOmplete Ultra EDTA-free protease inhibitor tablets (Roche)) contain-

ing 40 Kunitz units DNAseI and cells were lysed by passage through an EmulsiFlex-C3 (Avestin). Lysates were clarified by centrifugation at 12,000g, and nucleic acids and acidic proteins were precipitated by addition of a 10% polyethyleneimine (polymin P; Sigma-Aldrich) solution at pH 7.9 to a final concentration of 0.5%. Precipitated protein was pelleted, washed with TGEB (10 mM Tris pH 8.0, 5% glycerol, 0.1 mM EDTA, 10 mM β-mercaptoethanol) plus 0.3 M NaCl, and the RNAP fraction was eluted with TGEB plus 1 M NaCl. Residual polymin P was removed by ammonium sulfate precipitation (2M). The ammonium sulfate pellet was resuspended in TGEB and loaded onto a 50 mL Heparin Sepharose 6 Fast Flow column (GE Healthcare). The column was washed with 2 column volumes of TGEB plus 0.3 M NaCl, and RNAP was eluted with a step to TGEB plus 0.6 M NaCl. The elution fraction was precipitated with 2 M ammonium sulfate, and resuspended into approximately 1 mL of TGEB plus 0.5 M NaCl. Low molecular weight contaminants were removed via size exclusion chromatography on a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare). Fractions containing RNAP were diluted in TGEB to a final NaCl concentration of 0.3 M and loaded onto a HiTrap Q FF 5 mL column (GE Healthcare). RNAP was eluted into TGEB with a gradient between 0.3 M and 0.5 M NaCl over 20 column volumes. RNAP was dialyzed into RNAP storage buffer (20 mM Tris-HCl pH 8.0, 0.1 mM EDTA, 10 mM β-mercaptoethanol, 100 mM NaCl, 20% glycerol), concentrated to 1.4 mg/mL and frozen at -80 °C. The total yield was approximately 2.9 mg of high purity core enzyme.

BS^3 cross-linking. Bis(sulfosuccinimidyl) suberate (BS^3) d_0 and d_4 isotopologs were purchased from Thermo Scientific. RNAP and SutA were mixed in a 1:10 molar ratio (0.5 µM RNAP, SutA 5.0 µM) in 10 mM HEPES pH 8, 100 mM potassium acetate and incubated on ice for 1.5 h. Cross-linking was initiated by addition of 5 mM of a 4:1 molar ratio of BS^3 d_0:d_4 and the reaction was incubated on ice for 2 h. Cross-linking was quenched by addition of ammonium bicarbonate to a final concentration of 50 mM. In-solution digestion, HPLC desalting, and LC-MS/MS were performed as described below. The experiment was performed with two replicates.

Bpa cross-linking. RNAP and SutA mutants were mixed in a 1:10 molar ratio (0.5 µM RNAP, SutA 5.0 µM) in 25 mM Tris, pH 8, 100 mM NaCl and incubated on ice for 1 h. Samples were irradiated with 500 mW/cm^2 of 365 nm light for 120 s (or as otherwise indicated). Samples were separated via SDS-PAGE, stained with

Coomassie, and imaged.

Protein foot-printing. RNAP (control sample) or RNAP and SutA mixed in a 1:10 molar ratio (0.5 µM RNAP, SutA 5.0 µM) (SutA sample) were incubated in 10 mM HEPES pH 8, 100 mM potassium acetate on ice for 1.5 h. Modification of lysines was initiated by addition of 100 µM NHS-biotin, and samples were incubated at room temperature for 30 min. The reaction was quenched by addition of ammonium bicarbonate to a final concentration of 50 mM. Samples were separated via SDS-PAGE and GeLCMS was performed on the bands corresponding to the β and β' subunits of RNAP. HPLC desalting, and LC-MS/MS were performed as described below. The experiment was performed with two replicates.

Mass spectrometry. Liquid chromatography-mass spectrometry were essentially carried out as previously described [29]. Protein foot-printing experiments were performed on a nanoflow LC system, EASY-nLC 1000 coupled to a hybrid linear ion trap Orbitrap Classic mass spectrometer (Thermo Scientific) equipped with a nanoelectrospray ion source (Thermo Scientific) with the following modifications: For the EASY-nLC II system, solvent A consisted of 97.8% H2O, 2% ACN, and 0.2% formic acid and solvent B consisted of 19.8% H2O, 80% ACN, and 0.2% formic acid. For the LC-MS/MS experiments, digested peptides were directly loaded at a flow rate of 500 nL/min onto a 16-cm analytical HPLC column (75 µm ID) packed in-house with ReproSil-Pur C_{18}AQ 3 µm resin (120 Å pore size, Dr. Maisch, Ammerbuch, Germany). The column was enclosed in a column heater operating at 30 °C. After 30 min of loading time, the peptides were separated with a 50 min gradient at a flow rate of 350 nL/min. The gradient was as follows: 0–30% B (50 min), and 100% B (10 min). The Orbitrap was operated in data-dependent acquisition mode to automatically alternate between a full scan (m/z=400–1600) in the Orbitrap and subsequent 10 CID MS/MS scans in the linear ion trap. CID was performed with helium as collision gas at a normalized collision energy of 35% and 30 ms of activation time.

BS^3 and Bpa cross-linking experiments were run on the Orbitrap Elite, equipped with a nanoUPLC. Solvent A and B, and column were the same as described above. The gradient was as follows: 2% B for five min, 2-40% B (60 min), and 100% B (10 min). The Orbitrap was operated in data-dependent acquisition mode to automatically alternate between a full scan (m/z=300-1600) in the Orbitrap and subsequent 5 HCD MS/MS scans in the Orbitrap. Normalized collision energy was 40% and activation time was 100 ms. Resolution on MS was set to 120,000 and

MS/MS was 15,000.

Analysis of BS^3 cross-links. Raw files were converted to peak lists with ProteoWizard [27] and analyzed with Protein Prospector online, version 5.12.4 following reported protocols with modifications below [30]. The protein database contained the sequences for SutA, RpoA, RpoB, RpoC, and RpoZ. 80 peaks from each spectrum were searched using a tolerance of 10 ppm for precursor ions and 25 ppm for product ions. Enzyme specificity was GluC, and up to two missed cleavages per peptide were allowed. Carbamidomethylation of cysteines was specified as a constant modification, and protein N-terminal acetylation, oxidation of methionine, and dead-end modification with the cross-linker at lysine positions and protein N-termini were set as variable modifications. Additionally, incorrect monoisotopic peak assignments were considered as variable modifications. The analysis was run twice for each set of peak lists to search for both cross-linker isotopologs. Raw files were independently searched using MaxQuant for precursor mass pairs, differing by 4.02 Da, that represent cross-links made by both linker isotopologs.

Cross-links detected by Protein Prospector were matched against the mass pair list to remove cross-links not present as 4.02 Da offset mass pairs. For cross-links detected between RNAP proteins, we used a reported structural model of the *E. coli* RNAP complex (PDB: 3LU0) to calculate the inter α-carbon distance between amino acids [18]. We used this calculated distance as a metric to distinguish "quality" cross-links from all others. Based on the length of the linker, the maximum inter α-carbon distance between lysines cross-linked by BS^3 is 24.6 Å, so we considered cross-links with distances near or below this value to be reasonable. Like the study by Trnka et al., we found Score Difference to be the best discriminant for making this distinction. A Score Difference cutoff of 5.6 (similar to the value of 8.5 found by Trnka et al.) separated high-distance and low-distance cross-links (Figure 3.S2C) yielding an FDR of < 0.05 (see Figure 3.S2B for the ROC curve for this classification model). The final criteria for assigning quality cross-links were: (i) found as a precursor mass pair, (ii) Score Difference greater than 5.6, and (iii) matched by at least two spectra. These cross-links were aggregated to determine the number of spectra from each replicate and the maximum Score Difference for each amino acid linkage (Table 3.1). The best spectra used to identify each cross-link between SutA and RNAP are shown in Figure 3.S3.

Analysis of Bpa cross-links. Raw files were searched using MaxQuant against a protein database containing the sequences for SutA, RpoB, and RpoC and a

contaminant database (246 sequences). GluC was specified as the digestion enzyme with up to two missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification and protein N-terminal acetylation and methionine oxidation were variable modifications. We used the "matchedFeatures" output file to search for unique precursor masses in each run (e.g., found in a particular Bpa mutant run, but not in any other, including the uncross-linked RpoB and RpoC sample). We also searched raw files for variable mass modifications at any position corresponding to the expected SutA cross-linked peptide (Table S2) using MS-GF+ [31]. We cross-referenced the list of modified peptides (MSGF Score > 0) with the list of unique masses to find cross-linked peptides from RNAP.

Analysis of foot-printing. Raw files were searched using MaxQuant against a protein database containing the sequences for purified RpoB and RpoC and a contaminant database (246 sequences). Trypsin was specified as the digestion enzyme with up to two missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification and protein N-terminal acetylation and methionine oxidation were variable modifications. We also included a variable modification of lysine to search for biotinylated residues. For quantification, the raw files and the list of identified peptides were imported into Skyline version 3.1, and subset for high quality peak matches among all runs (isotopic dot product score > 0.75) [32]. For each replicate, peptide intensity ratios were calculated for each peptide ion between the SutA and control samples. To account for variations in LC-MS/MS loading, all peptide intensity ratios for each experiment were normalized so that the median ratio was 1. Peptides whose intensities changed by 1.5-fold between the SutA and control samples in both replicates were classified as "obscured" (SutA < control) or "revealed (SutA > control).

Analytical size exclusion chromatography. RNAP (0.5 µM) and SutA (5.0 µM) were incubated together or separately for 2 h at 4 °C in 40 mM Tris pH 8.0, 30 mM NaCl, 5% glycerol, and 3 mM β-mercaptoethanol, then separated on a Superdex Increase 3.2/300 column (GE Healthcare).

Software analysis and data presentation. This section describes software packages that were not mentioned above. Data processing and statistical analysis were performed with Python version 2.7.9 with NumPy version 1.9.2, SciPy version 0.15.1, and Pandas version 0.16.1. Data were plotted with Matplotlib version 1.4.3 [33] and Seaborn version 0.5.1. Gel images were analyzed with ImageJ 64-bit version 1.45 [34]. RNAP structural analysis was performed with Biopython version

1.65 [35] and structural visualization was performed with Open-Source PyMOL version 1.3. Figures were assembled in Adobe Illustrator CS5.

3.8 Supplementary Figures

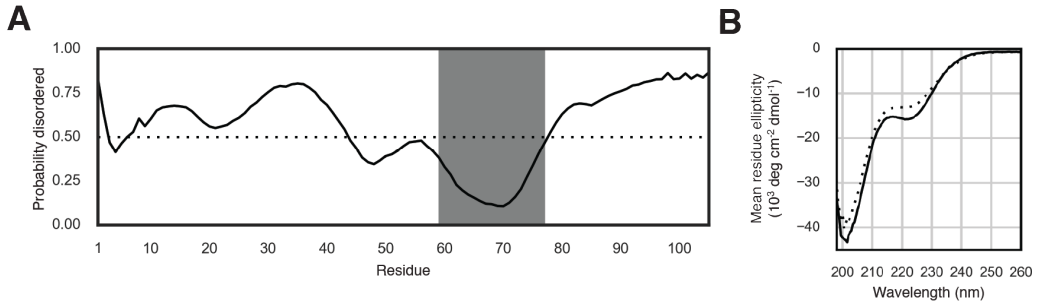

Figure 3.S1: **SutA structural predictions.** (A) Disordered regions were predicted using DisEMBL [11]. The region containing the predicted alpha helix is indicated with a gray box. (B) CD measurements of purified SutA at 4 °C (solid) and 22 °C (dotted).

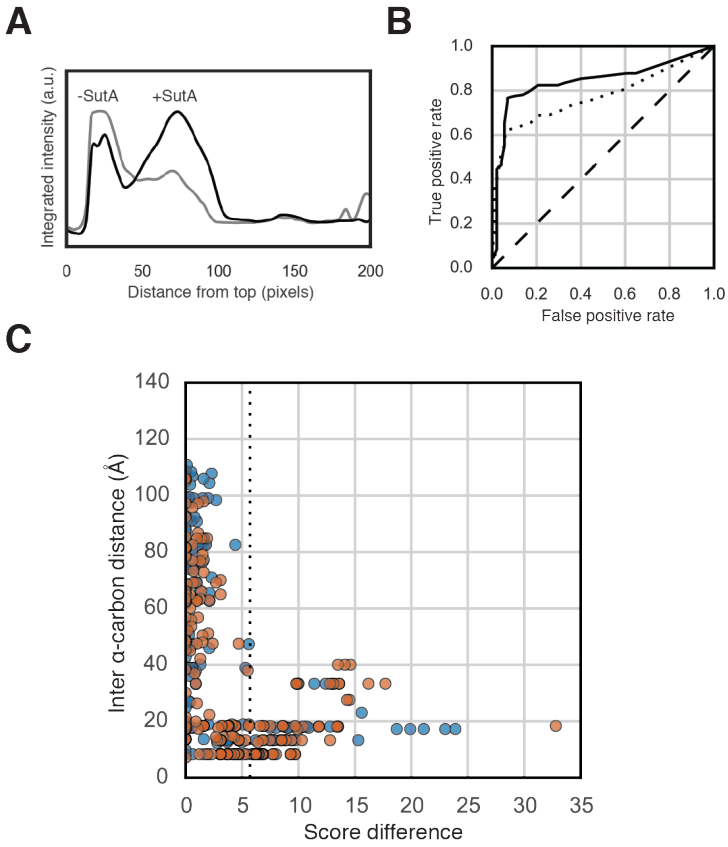

Figure 3.S2: **BS^3 cross-linking.** (A) Analysis of high molecular weight RNAP cross-linking products without (gray) and with SutA (black). Integrated density is plotted for the upper region of lanes from Figure 3.1B. (B) ROC curve for the "quality" cross-link model before (dotted) or after (solid) subsetting for paired precursor ions (i.e., found with both BS^3 isotopologs). (C) Inter α-carbon distance vs. Score Difference for cross-links from replicate 1 (blue) and replicate 2 (orange). The dotted line indicates the Score Difference cutoff used.

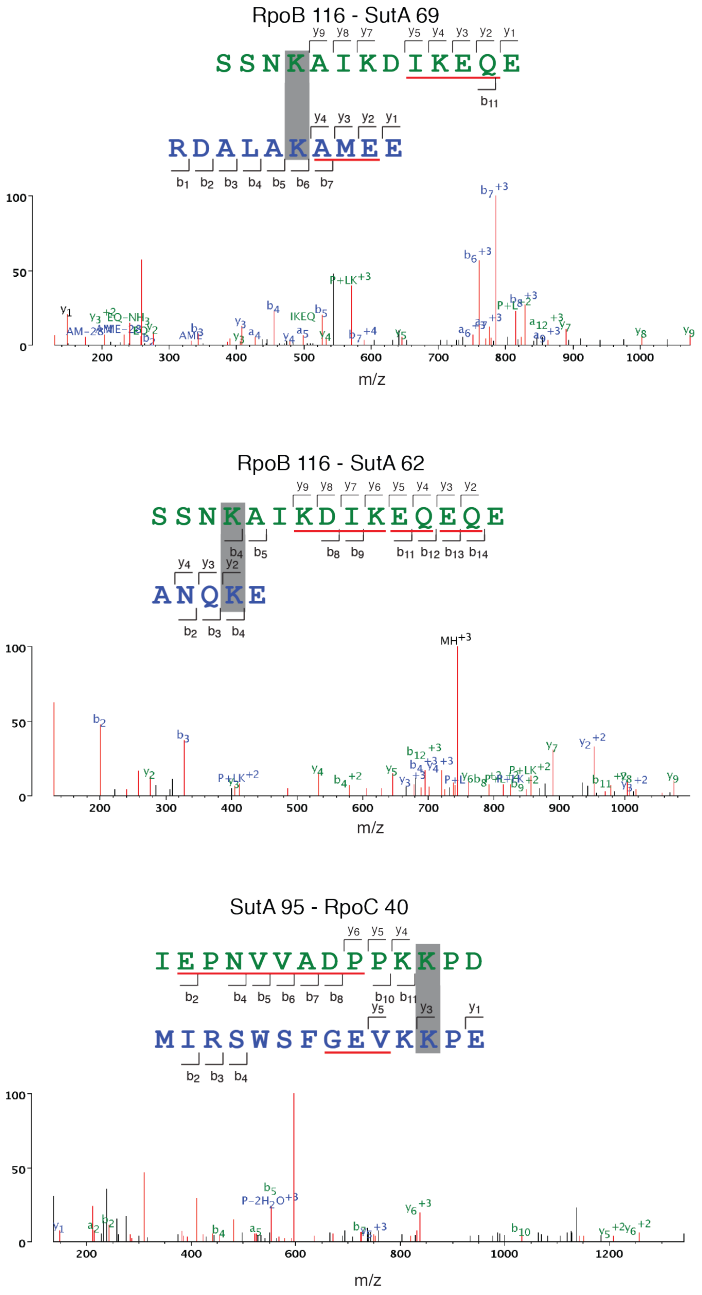

Figure 3.S3: **MS2 spectra for SutA-RNAP BS^3 cross-links.** HCD spectra for the highest scoring cross-links between SutA and RNAP. Matched fragment ions (b and y) are indicated above and below the peptide sequences and inter-peptide fragment ions by red underlines. The location of each cross-link is indicated by a gray box.

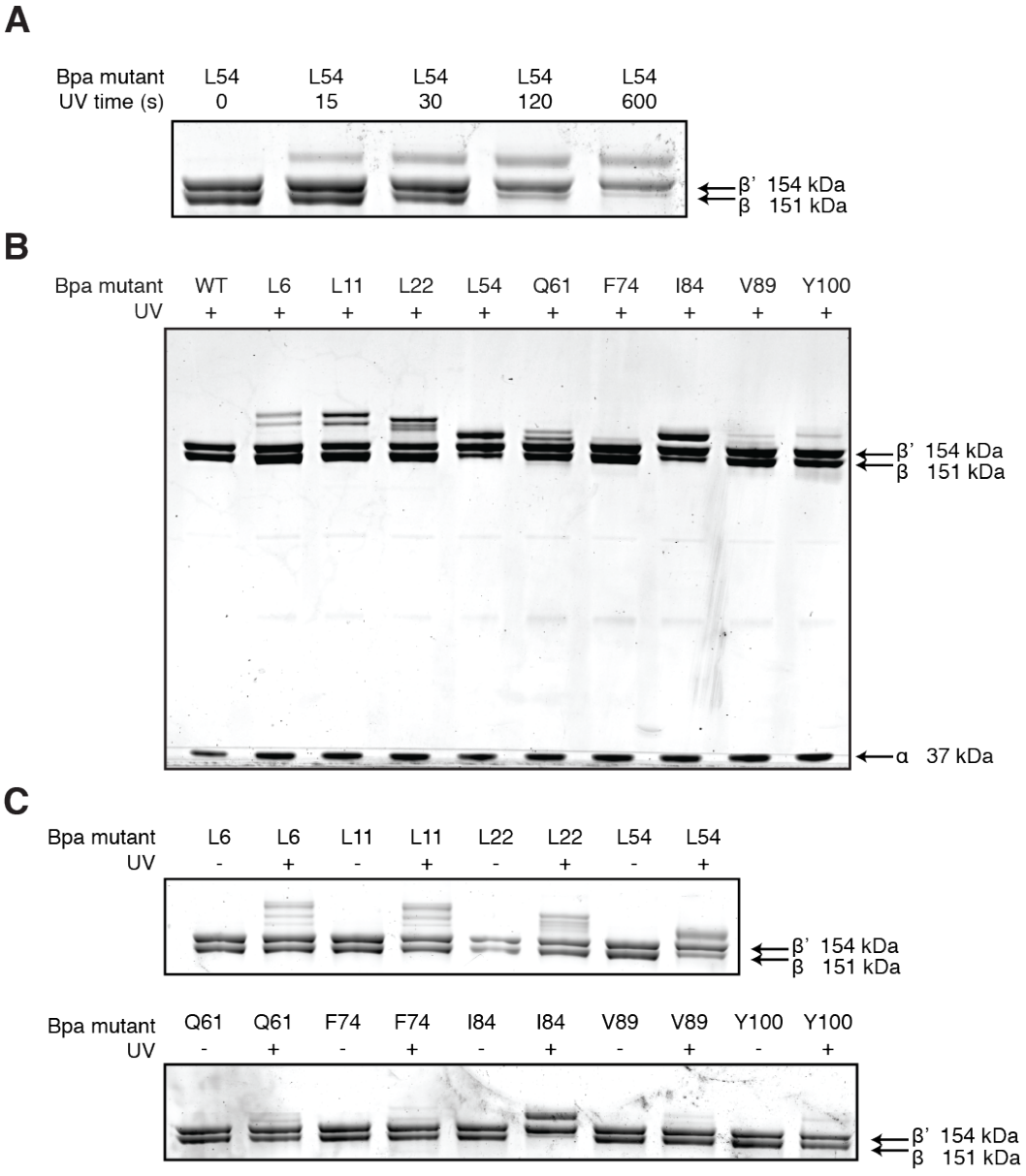

Figure 3.S4: **Bpa cross-linking.** SDS-PAGE analysis of Bpa cross-linking. (A) A time course of UV irradiation of RNAP mixed with SutA L54Bpa mutant. (B) Full gel from Figure 3.1C. (C) RNAP incubated with SutA mutants and irradiated in the absence of nucleic acids.

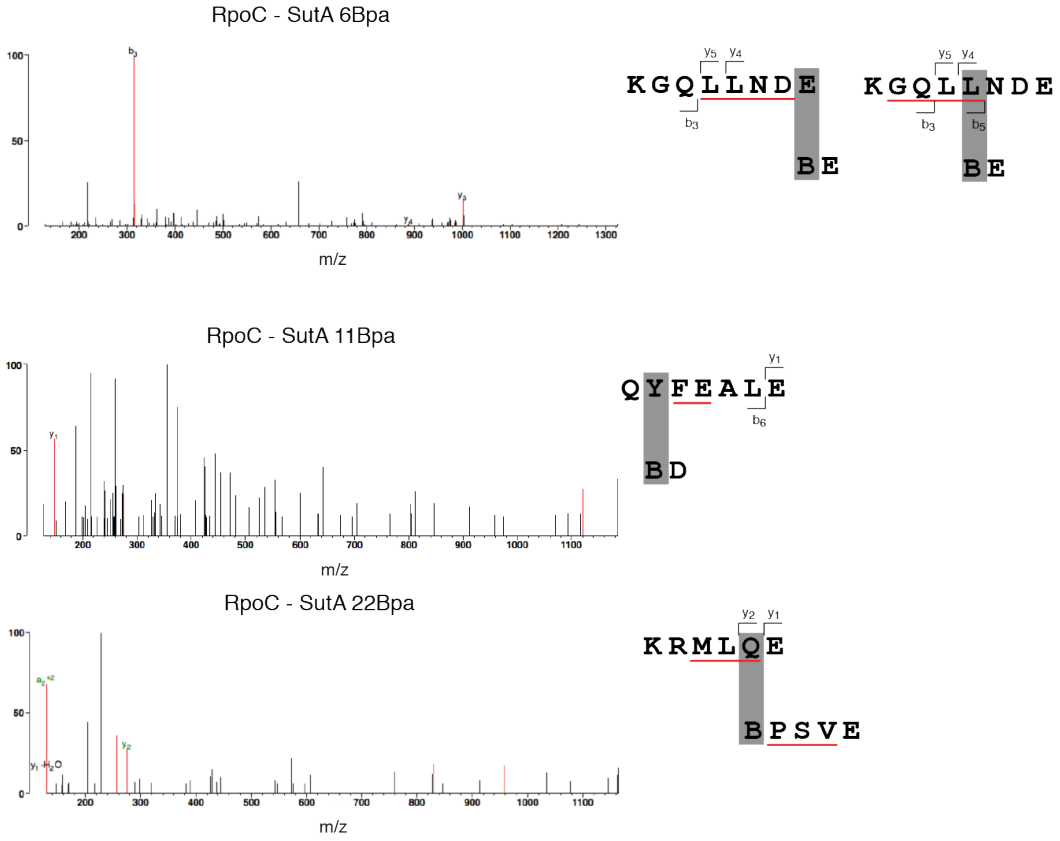

Figure 3.S5: **Bpa cross-link spectra.** HCD spectra for the best cross-links between SutA Bpa mutants and RNAP. Matched fragment ions (b and y) are indicated above and below the peptide sequences and inter-peptide fragment ions by red underlines. The location of each cross-link is indicated by a gray box. When evidence for multiple cross-links between the same peptides were found, all cross-link locations are shown.

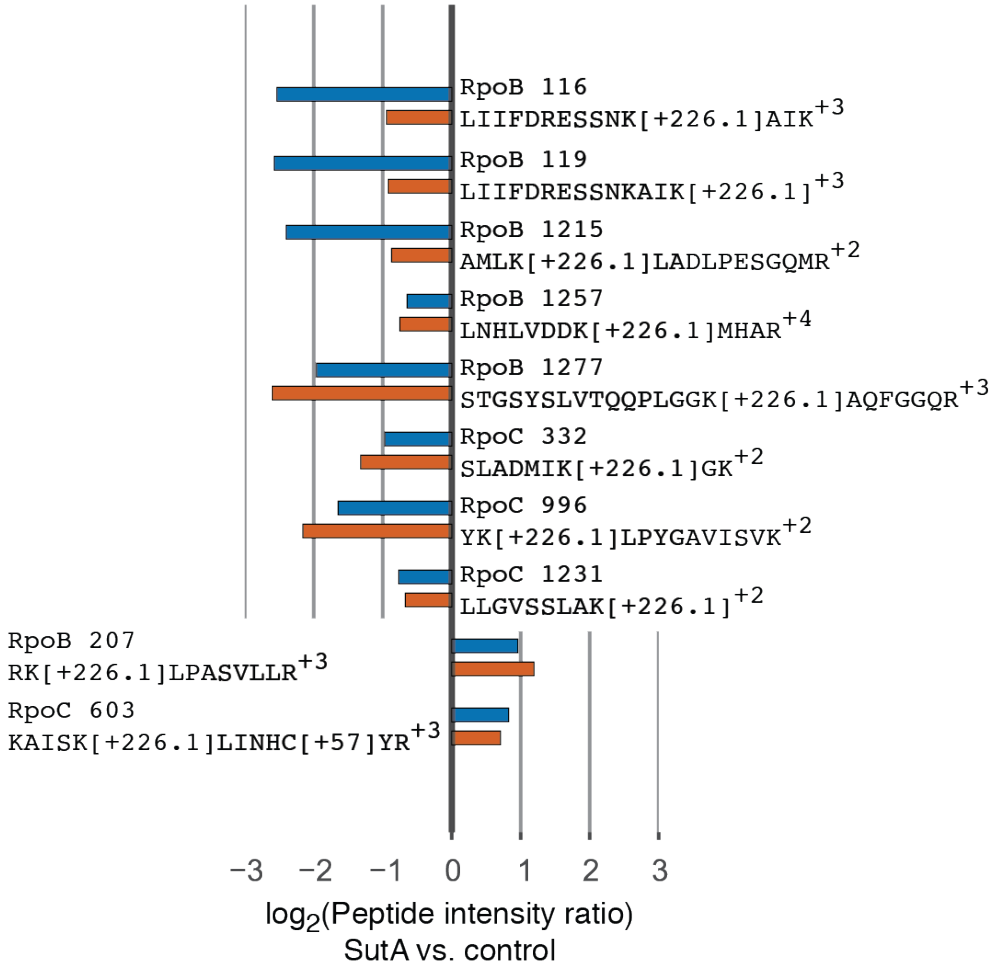

Figure 3.S6: **Protein foot-printing.** Lysine residues determined to be obscured (top) or revealed (bottom) in the presence of SutA. Peptide intensity ratios between RNAP with SutA vs. RNAP alone are shown for each replicate (blue or orange). The modified residue, modified peptide sequence, and charge state are listed next to the bars for each intensity ratio.

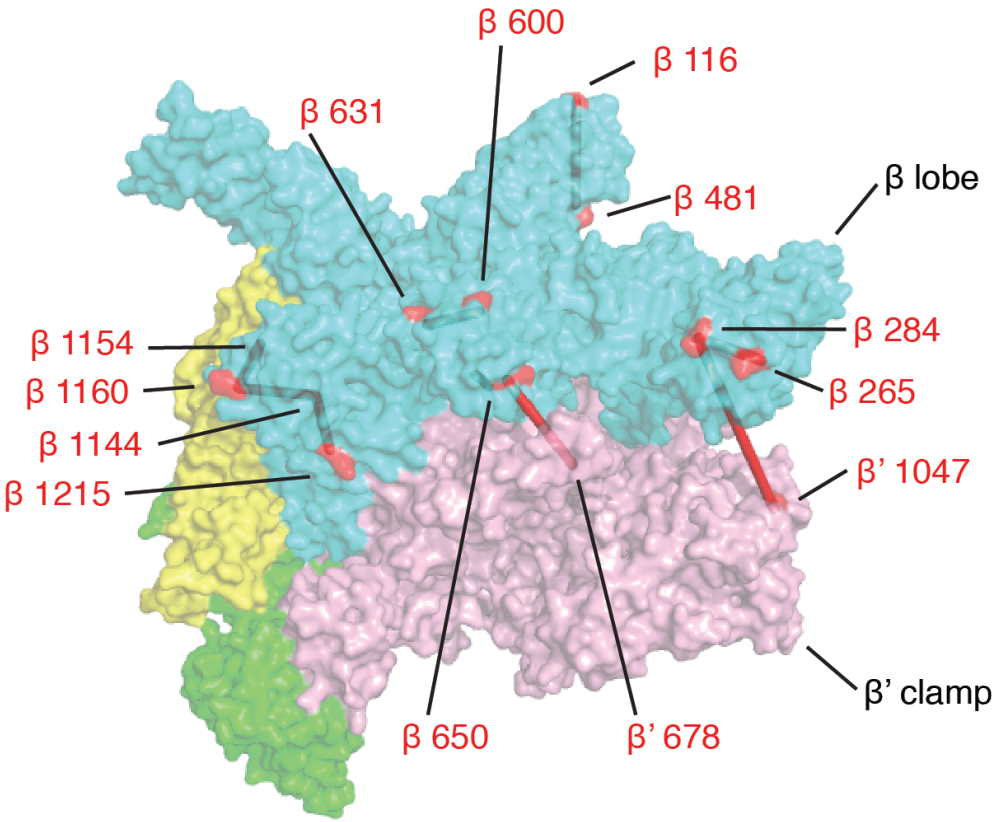

Figure 3.S7: **Intra-RNAP BS^3 cross-links.** Cross-links were mapped onto the *E. coli* structure. Cross-linked lysines are colored red, and inter α-carbon distances are displayed as red bars. Eight detected intra-RNAP cross-links are shown; the ninth is located on the opposite face of the structure. Subunits are colored as follows: α_I (yellow), α_{II} (green), β (cyan), and β' (pink).

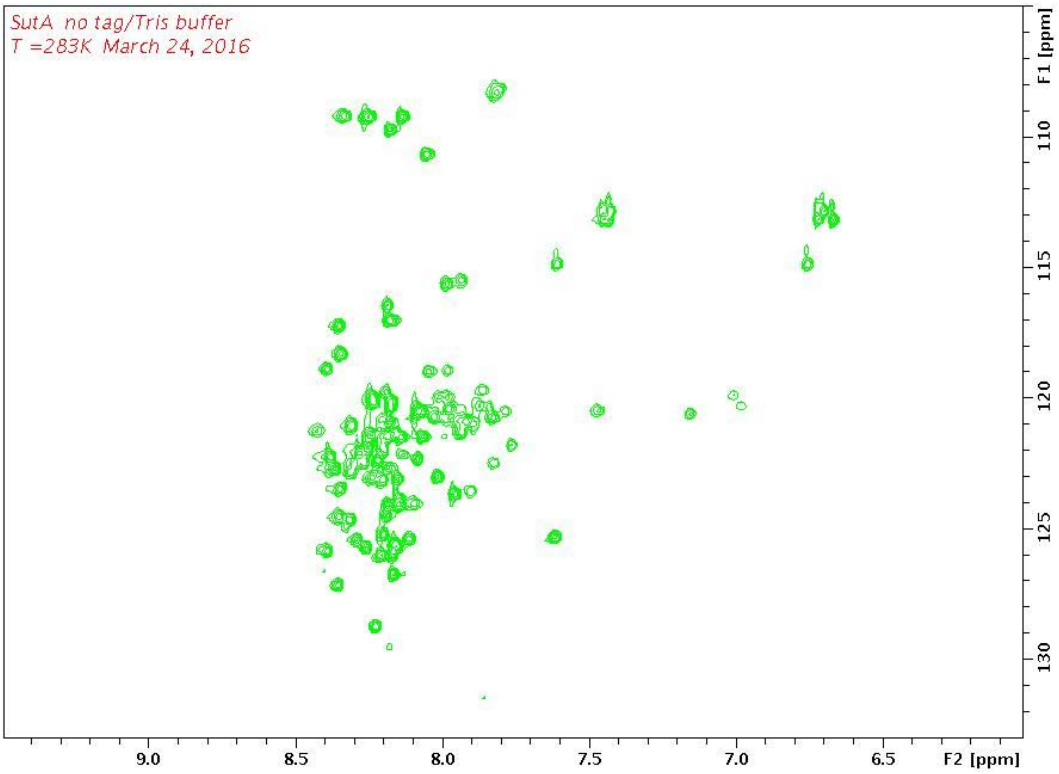

Figure 3.S8: **2D NMR of SutA.** HSQC spectrum of ^{15}N-labeled SutA. The horizontal and vertical axes represent ^{1}H and ^{15}N chemical shifts, respectively. Provided by Dr. Ben Ramirez.

3.9 Supplementary Tables.

Table 3.3: **Chapter 3: Strains and plasmids.** Strains and plasmids used in this study. Plasmids are stored as *E. coli* strains carrying the plasmid, and requests should be for the *E. coli* strain.

***Pseudomonas aeruginosa* Strains**		
Name	Genotype	Source
DKN263	*P. aeruginosa* UCBPP-PA14	
DKN1625	UCBPP-PA14 Δ*sutA*	[1]
***Escherichia coli* Strains**		
Name	Genotype	Source
DKN1640	Mach1 pMQ72_HA-sutA	[1]
DKN1687	Mach1 pMQ72_HA-sutA_Nterm	This Study
DKN1688	Mach1 pMQ72_HA-sutA_Cterm	This Study
DKN548	DH5α pMQ72	George O'Toole
DKN1697	BL21 DE3, pQE80L-6xHis-TEV-SutA	This Study
BMB14X	BL21 DE3, pQE80L-6xHis-TEV-SutA with XBpa	This Study
pEVOL-pBpF	BL21 DE3, p15A-pBpa synthetase and tRNA	[15]

Table 3.4: **Bpa mass modifications** Mass modifications that were used to search for Bpa cross-links. "B" represents the location of Bpa in the peptide sequence.

SutA Mutant	Peptide	Mass
L6	BE	398.1478
L11	BD	384.1321
L22	BAAAD	597.2434
L54	BPSVE	681.3010
I84	BE	398.1478

References

(1) Babin, B. M.; Bergkessel, M.; Sweredoski, M. J.; Moradian, A.; Hess, S.; Newman, D. K.; Tirrell, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, E597–E605.

(2) Dennis, P. P.; Bremer, H. *EcoSal Plus* **2013**, *1*.

(3) Paul, B. J.; Ross, W.; Gaal, T.; Gourse, R. L. *Ann. Rev. Genet.* **2004**, *38*, 749–770.

(4) Condon, C.; Squires, C.; Squires, C. L. *Microbiol. Rev.* **1995**, *59*, 623–645.

(5) Lemke, J. J.; Sanchez-Vazquez, P.; Burgos, H. L.; Hedberg, G.; Ross, W.; Gourse, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 5712–5717.

(6) Belogurov, G. A.; Artsimovitch, I. *Ann. Rev. Microbiol.* **2015**, *69*, 49–69.

(7) Zenkin, N.; Yuzenkova, Y. *Biomolecules* **2015**, *5*, 1195–1209.

(8) López de Saro, F. J.; Woody, A. Y.; Helmann, J. D. *J. Mol. Biol.* **1995**, *252*, 189–202.

(9) Withers, R.; Doherty, G. P.; Jordan, M.; Yang, X.; Dixon, N. E.; Lewis, P. J. *Mol. Microbiol.* **2014**, *93*, 1130–1143.

(10) Srivastava, D. B.; Leon, K.; Osmundson, J.; Garner, A. L.; Weiss, L. A.; Westblade, L. F.; Glickman, M. S.; Landick, R.; Darst, S. A.; Stallings, C. L.; Campbell, E. A. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 12619–12624.

(11) Linding, R.; Jensen, L. J.; Diella, F.; Bork, P.; Gibson, T. J.; Russell, R. B. *Structure* **2003**, *11*, 1453–1459.

(12) Kelley, L. A.; Sternberg, M. J. E. *Nat. Protoc.* **2009**, *4*, 363–371.

(13) Chen, Z. A.; Jawhari, A.; Fischer, L.; Buchen, C.; Tahir, S.; Kamenski, T.; Rasmussen, M.; Lariviere, L.; Bukowski-Wills, J.-C.; Nilges, M.; Cramer, P.; Rappsilber, J. *EMBO J.* **2010**, *29*, 717–726.

(14) Lane, W. J.; Darst, S. A. *J. Mol. Biol.* **2010**, *395*, 671–685.

(15) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020–11024.

(16) Liu, B.; Zuo, Y.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 2006–2010.

(17) Shell, S. M.; Hess, S.; Kvaratskhelia, M.; Zou, Y. *Biochemistry* **2005**, *44*, 971–978.

(18) Opalka, N.; Brown, J.; Lane, W. J.; Twist, K.-A. F.; Landick, R.; Asturias, F. J.; Darst, S. A. *PLoS Biol.* **2010**, *8*, e1000483.

(19) Lennon, C. W.; Ross, W.; Martin-Tumasz, S.; Toulokhonov, I.; Vrentas, C. E.; Rutherford, S. T.; Lee, J.-H.; Butcher, S. E.; Gourse, R. L. *Genes Dev.* **2012**, *26*, 2634–2646.

(20) Murakami, K. S.; Masuda, S.; Campbell, E. A.; Muzzin, O.; Darst, S. A. *Science* **2002**, *296*, 1285–1290.

(21) Vassylyev, D. G.; Vassylyeva, M. N.; Perederina, A.; Tahirov, T. H.; Artsimovitch, I. *Nature* **2007**, *448*, 157–162.

(22) Chakraborty, A.; Wang, D.; Ebright, Y. W.; Korlann, Y.; Kortkhonjia, E.; Kim, T.; Chowdhury, S.; Wigneshweraraj, S.; Irschik, H.; Jansen, R.; Nixon, B. T.; Knight, J.; Weiss, S.; Ebright, R. H. *Science* **2012**, *337*, 591–595.

(23) Furman, R.; Tsodikov, O. V.; Wolf, Y. I.; Artsimovitch, I. *J. Mol. Biol.* **2013**, *425*, 82–93.

(24) Lennon, C. W.; Gaal, T.; Ross, W.; Gourse, R. L. *J. Bacteriol.* **2009**, *191*, 5854–5858.

(25) Sevostyanova, A.; Belogurov, G. A.; Mooney, R. A.; Landick, R.; Artsimovitch, I. *Mol. Cell* **2011**, *43*, 253–262.

(26) Burmann, B. M.; Schweimer, K.; Luo, X.; Wahl, M. C.; Stitt, B. L.; Gottesman, M. E.; Rösch, P. *Science* **2010**, *328*, 501–504.

(27) Chambers, M. C. et al. *Nat. Biotechnol.* **2012**, *30*, 918–920.

(28) Kuznedelov, K.; Semenova, E.; Knappe, T. A.; Mukhamedyarov, D.; Srivastava, A.; Chatterjee, S.; Ebright, R. H.; Marahiel, M. A.; Severinov, K. *J. Mol. Biol.* **2011**, *412*, 842–848.

(29) Kalli, A.; Hess, S. *Proteomics* **2012**, *12*, 21–31.

(30) Trnka, M. J.; Baker, P. R.; Robinson, P. J. J.; Burlingame, A. L.; Chalkley, R. J. *Mol. Cell Proteomics* **2014**, *13*, 420–434.

(31) Kim, S.; Pevzner, P. A. *Nat. Commun.* **2014**, *5*, 5277.

(32) Schilling, B.; Rardin, M. J.; MacLean, B. X.; Zawadzka, A. M.; Frewen, B. E.; Cusack, M. P.; Sorensen, D. J.; Bereman, M. S.; Jing, E.; Wu, C. C.; Verdin, E.; Kahn, C. R.; Maccoss, M. J.; Gibson, B. W. *Mol. Cell Proteomics* **2012**, *11*, 202–214.

(33) Hunter, J. D. *Comput. Sci. Eng.* **2007**, *9*, 90–95.

(34) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. *Nat. Methods* **2012**, *9*, 671–675.

(35) Hamelryck, T.; Manderick, B. *Bioinformatics* **2003**, *19*, 2308–2310.