7974–7988 *Nucleic Acids Research, 2016, Vol. 44, No. 16* *Published online 25 July 2016*
doi: 10.1093/nar/gkw656

Visualizing the phage T4 activated transcription complex of DNA and *E. coli* RNA polymerase

Tamara D. James[1,2], Timothy Cardozo[2,*], Lauren E. Abell[1], Meng-Lun Hsieh[1], Lisa M. Miller Jenkins[3], Saheli S. Jha[1] and Deborah M. Hinton[1,*]

[1]Gene Expression and Regulation Section, Laboratory of Cell and Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA, [2]Department of Biochemistry and Molecular Pharmacology, NYU Langone Medical Center, New York University School of Medicine, 180 Varick Street, Room 637, New York, NY 10014, USA and [3]Collaborative Protein Technology Resource, Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received February 13, 2015; Revised July 04, 2016; Accepted July 05, 2016

ABSTRACT

The ability of RNA polymerase (RNAP) to select the right promoter sequence at the right time is fundamental to the control of gene expression in all organisms. However, there is only one crystallized structure of a complete activator/RNAP/DNA complex. In a process called σ appropriation, bacteriophage T4 activates a class of phage promoters using an activator (MotA) and a co-activator (AsiA), which function through interactions with the σ^{70} subunit of RNAP. We have developed a holistic, structure-based model for σ appropriation using multiple experimentally determined 3D structures (*Escherichia coli* RNAP, the *Thermus aquaticus* RNAP/DNA complex, AsiA /σ^{70} Region 4, the N-terminal domain of MotA [$MotA^{NTD}$], and the C-terminal domain of MotA [$MotA^{CTD}$]), molecular modeling, and extensive biochemical observations indicating the position of the proteins relative to each other and to the DNA. Our results visualize how AsiA/MotA redirects σ, and therefore RNAP activity, to T4 promoter DNA, and demonstrate at a molecular level how the tactful interaction of transcriptional factors with even small segments of RNAP can alter promoter specificity. Furthermore, our model provides a rational basis for understanding how a mutation within the β subunit of RNAP (G1249D), which is far removed from AsiA or MotA, impairs σ appropriation.

INTRODUCTION

As viruses are intracellular parasites that cannot replicate by themselves, they rely on an immense diversity of clever hijacking schemes to direct host cell machinery towards their own development. Understanding the mechanism of these hijackings often sheds light on the function and susceptibility of the targeted host mechanism. Bacteriophage T4 does exactly this via appropriation of the σ subunit of *Escherichia coli* RNA polymerase (RNAP) [reviewed in (1,2)]. In this system, novel interactions among *E. coli* RNAP, the T4 MotA activator, and the T4 AsiA co-activator redirect RNAP from the expression of host to T4 promoter DNA. Thus, visualization of the molecular mechanism of this RNAP hijacking illuminates a process by which the specificity of an RNAP can be altered.

Bacterial RNAP contains both a core (subunits α_1, α_2, β, β' and ω), which has RNA synthesizing activity, and a specificity subunit, σ, which recognizes promoters at the transcription start site [reviewed in (2–4)] (Figure 1A). In *E. coli*, σ^{70} is the primary σ, needed for exponential growth and transcription of housekeeping genes. All primary σ's, including σ^{70}, share four conserved domains based on structure and sequence [reviewed in (5,6)]. Regions 2, 3 and 4 recognize and bind conserved promoter sequences, the −10 element, the $^{-15}TG^{-14}$ (extended −10), and the -35 element, respectively (reviewed in (2,3,7)). σ^{70} also has extensive and specific interactions with the core. Structures of bacterial RNAP's (8–10) indicate that σ Region 2 binds to the β' clamp helix, a coiled-coil domain, while Region 4 plus Helix 5 (H5; σ residues 600–613), the C-terminus of σ, interact with the flap domain of β (β-flap). These interactions position the DNA binding Regions 2 and 4 in 3D space to allow

*To whom correspondence should be addressed. Tel: +1 301 496 9885; Fax: +1 301 402 0053; Email: dhinton@helix.nih.gov
Correspondence may also be addressed to Timothy Cardozo. Tel: +1 212 263 6337; Fax: +1 646 754 9632; Email: timothy.cardozo@med.nyu.edu
Present addresses:
Tamara James, GeneCentrix, Inc. 175 Varick St., 4th Floor New York, NY 10014, USA.
Meng-Lun Hsieh, Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA.
Lauren Abell, Mitochondria and Metabolism Center, Pathology Department, University of Washington, Seattle, WA 98109, USA.
Saheli Jha, EpicentRx, Mountain View, CA 94040, USA.

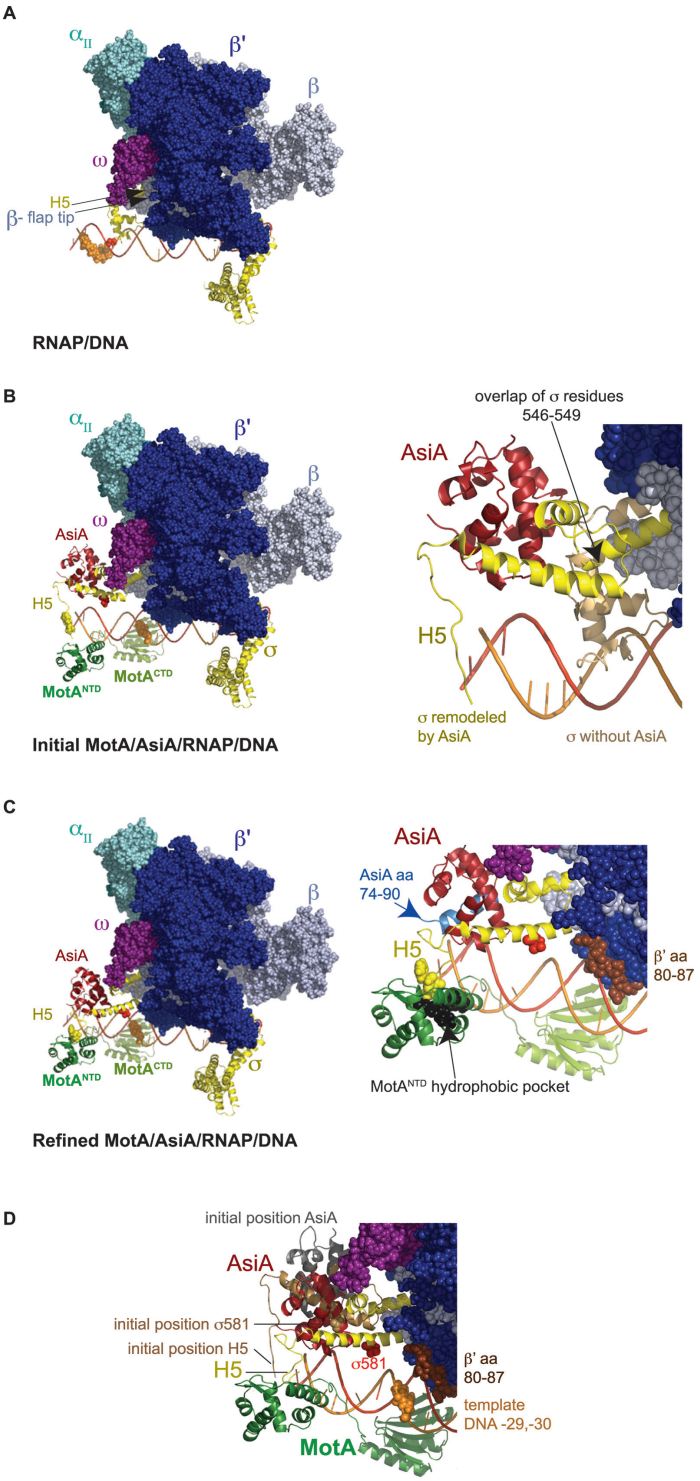

Figure 1. (A–C) Structural models of T4 middle promoter DNA with (**A**) *E. coli* RNAP alone, (**B**) *E. coli* RNAP/MotA/AsiA (initial) and (**C**) *E. coli* RNAP/MotA/AsiA (refined). Colors used are β', cobalt; β, ice blue; ω, magenta; α_{II}, teal; σ, yellow, $MotA^{NTD}$, forest green; $MotA^{CTD}$, light green; AsiA, firebrick red; DNA nontemplate strand, bright red; template stand, orange. (α_I cannot be seen in this orientation.) The position of H5 (the C-terminal residues of σ) is indicated, σD581 is shown as a red sphere, and the template cleavage sites obtained with σD581C-FeBABE are shown as orange spheres. The position of the β-flap tip is shown in (A) and F610 and L611 of H5 are shown as yellow spheres in (B) and (C). Right panel of (B) shows Region 4/H5 of both the σ present in *E. coli* RNAP (tan) and the AsiA-remodeled σ (yellow). σ residues 546–549 were overlapped to position AsiA-remodeled σ; the ω subunit has been removed in order to see the overlap. Right panel of (C) shows the C-terminal 17 residues of AsiA (74-90) in light blue, β' residues 80–87 in brown, and $MotA^{NTD}$ residues I7, I24, A83 and Y86, which are located in the hydrophobic pocket, as black spheres. (**D**) Alignment of the initial and refined models. The positions of AsiA (grey), σ Region 4 (tan), and σ position 581 (tan sphere) in the initial model and AsiA (red), σ Region 4 (yellow), and σ position 581 (red sphere) in the refined model are indicated. β' residues 80–87 (brown spheres) and template DNA positions −29, −30 (orange spheres) are also shown. See text for more details.

simultaneous recognition of the −10 and −35 promoter elements on the DNA (11,12).

Transcriptional activators can change the specificity of RNAP, and many activators specifically target σ^{70} (reviewed in (2,13)). For example, Class II activators contact the DNA just upstream or overlapping the -35 element. Acidic residues within Class II activators engage basic amino acids of σ^{70} Region 4, which lie on a surface-exposed helix. This interaction between a Class II activator and Region 4 then forces σ^{70} to interact with promoter elements whose sequences deviate from consensus and/or facilitates the isomerization of the initial RNAP/DNA complex to a transcriptionally active form. Although the binding of the activator may somewhat distort σ Region 4 (14), the overall conformation of Region 4 is still retained. The recent structure of the Class II complex containing TAP, an analog of *E. coli* CRP (c-AMP receptor protein), *Thermus thermophilus* RNA polymerase, the unwound DNA, and a small complementary RNA has provided a picture of the molecular contacts needed for this activation (15).

In contrast, in order to recognize T4 middle promoters, rearrangement of σ itself is exploited by the phage. In this fundamentally different process, called σ appropriation, MotA uses a basic/hydrophobic cleft in its N-terminal domain (NTD) to engage H5 [(16–18); reviewed in (1)], while its C-terminal domain (CTD) binds the -30 region (MotA box) of a T4 middle promoter (19,20). MotA activation also requires a small phage-encoded co-activator, AsiA, which binds tightly to σ^{70} Region 4, structurally remodeling this region (21,22). Although hundreds of activators of bacterial polymerase have been identified, MotA/AsiA is the only system known to function in this way. However, the unusual 'double-wing' motif seen in the structure of $MotA^{CTD}$ (23) has also been found in structures of some bacterial proteins of unknown function (24,25), suggesting that there could be unidentified systems that bear resemblance to the T4 system.

Despite understanding how fragmented pieces of MotA, AsiA and RNAP connect during σ appropriation, obtaining a complete picture of this process has been challenging as there is no structure for MotA with the DNA. Furthermore, some findings are puzzling. For example, work has suggested that contacts between the C-terminal region of AsiA and both the β-flap of core and MotA are needed for σ appropriation (26,27), even though substitutions and deletions of C-terminal AsiA residues do not significantly affect AsiA function *in vitro* (28,29). Furthermore, the G1249D substitution within the β subunit specifically impairs middle promoter activation *in vivo* (30) without affecting host transcription or transcription from early T4 promoters. However, the closest σ residues to G1249 lie within σ Region 3.2 rather than Region 4, which is targeted by AsiA. Thus, the ability to visualize a complete RNAP/MotA/AsiA/DNA complex could reveal the cooperative protein-protein and protein-DNA interactions needed for this process.

Despite extensive biochemical and genetic analyses and the reported structures of thermophilic RNAP (8,10), the structure and structure-model of *E. coli* RNAP (9,31), and structures of various activators (detailed in (32)), only one crystal structure of a complete activated transcription complex has been obtained (15). In the case of σ appropriation, available structures include AsiA in a complex with σ^{70} Region 4/H5 (21) and structures of $MotA^{CTD}$ (23) and $MotA^{NTD}$ (33). However, there are no structures of the MotA linker, full length MotA, or structures that include DNA. We hypothesized that careful integration of the structure of *E. coli* RNAP (9) with that of σ^{70} Region 4 in complex with AsiA (21) could result in a structural model of an AsiA-associated *E. coli* RNAP. Furthermore, in previous work we have biochemically mapped the orientation of MotA relative to the DNA within the σ-appropriated complex using MotA conjugated with the chemical cleaving reagent iron bromoacetamidobenzyl-EDTA (FeBABE) (34). This work then provides a framework within which to incorporate other analyses indicating how $MotA^{NTD}$ interacts with σ^{70} H5 (16,35). Rational combination of these heterogeneous data affords a complete snapshot of the protein-protein and protein-DNA interactions within the σ appropriation process. Our work depicts how AsiA/MotA redirects σ, and therefore RNAP activity, to a T4 middle promoter DNA and how the flexibility of σ Region 4/H5 is likely crucial for this process. Our model suggests that the βG1249D substitution impairs this process by hampering this needed flexibility.

MATERIALS AND METHODS

Molecular model of AsiA-appropriated RNAP/MotA/T4 middle promoter DNA

Two template structures were used to build a homology model of the modified AsiA-associated σ^{70} subunit: *E.coli* RNAP holoenzyme complex crystal (PDB ID: 4IGC (9)) and the NMR structure of *E. coli* σ^{70} Region 4/H5 bound to AsiA (PDB ID: 1TLH (21)). The structures were superimposed at residues 546–549 to allow different portions of the query σ^{70} sequence to align based on the following assignments: The σ^{70} query sequence to be built included the 4IGC chain y residues 95–609 plus σ^{70} residues 610–613. Residues 95–545 of the query sequence were tethered to 4IGC (chain y residues 95–545), and the σ^{70} query sequence residues 546–613 were tethered to 1TLH (chain b residues 546–613). An extended full-atom 3D model of a polypeptide chain with idealized covalent geometry was built for σ^{70} based on its sequence; torsion angles for the aligned portions of the backbone and identical aligned residues were assigned to those in the template structures; the most likely rotamer was assigned to non-identical aligned residue side chains; near-extended backbone torsion values were assigned to the loops; the sum of the physical energy and quadratic restraints between corresponding atoms in the model and template structures were iteratively minimized, reducing the strength of the restraint potential with each iteration; and finally non-identical side-chains and loops were subjected to Biased-Probability Monte Carlo (BPMC) sampling (36) to produce the lowest energy structure matching the template coordinates. The final root-mean-square deviation (RMSD) for each model was less than 0.5 Å to its template for matched segments. This resulted in a 3-D model of σ^{70} from threading a model of its sequence onto the indicated template fragments and through the junctions between them. The newly built σ^{70} was reincorporated into

the *E.coli* RNAP complex based on the position of the PDB ID: 4IGC, chain y.

The location of AsiA was simply inherited by its location in its NMR structure with *E. coli* σ Region 4/H5 (PDB ID: 1TLH (21)), since the σ portion of this structure was used as a template for the model.

DNA, extending from -7 to -41 on the nontemplate strand and -12 to -41 on the template strand, was incorporated into the model using two structures. The first structure was *T. aquaticus* holoenzyme containing structured and extended modeled DNA (chain t and chain u) (PDB ID: 1L9Z (37)). Neighboring residues surrounding 7Å of the template and nontemplate DNA positions −7 to −13 were limited to σ Region 2 (1L9Z PDB chain h residues 241–285). These residues were identified, and superimposed, based on sequence alignment, with our 3D homology model of σ^{70} Region 2 (residues 418–462). By superimposing the modeled duplex DNA containing the MotA box element (34) from positions −19 to −41 onto our promoter model, we extended the DNA to −41 using modeled B form DNA that included both $MotA^{CTD}$ and $MotA^{NTD}$ that had been placed on the DNA through FeBABE analyses (34).

A multi-template homology model of full length MotA was made using the PDB ID: 1I1S (33), chain a as a template for the $MotA^{NTD}$, and PDB ID: 1KAF, chain a as a template for the $MotA^{CTD}$ (23). The full length MotA sequence was aligned to the NTD and CTD templates and a 3D model of full length MotA was built by multi-template homology modeling as for the RNAP components. To accommodate a linker that connects the NTD and CTD, an unstructured loop was generated between residues 93 to 105 using a previously described loop modeling approach (38).

To refine the model based on the biochemical constraints imposed by analyses with σD581C conjugated with FeBABE or containing the photocrosslinkable amino acid analog benzoyl-phenylalanine (BpA), the following steps were taken: quadratic distance restraints as implemented in ICM-Pro imposing a penalty outside of an 8–11 Å range were set between position 581 of σ^{70} and the amino acids in β' that were shown to be near σ 581 by crosslink/mass spectrometry analyses and DNA positions −29/−30, and σ F610 and L611 were tethered to their locations in the $MotA^{NTD}$ hydrophobic cleft using the 'set tether' command of ICM-Pro (36). The φ and ϕ backbone dihedral angles of σ^{70} amino acids 550–551 and 595–604, which are hinge or random coil regions of Region 4, were unfixed and sampled by BPMC (36) of van der Waals, electrostatic, torsion, hydrogen bonding, solvation, entropy and distance restraint terms for 1 million calls until a negative energy value was reached indicating a low energy structure that satisfied the distance restraints. This approach left all of the crystallographic and NMR folded domain structures completely unaltered.

DNA

pDKT90 (19), which contains the T4 middle promoter P_{uvsX}, and $pP_{uvsX\text{-}sigma}$ (39), in which a consensus -35 element (TTGACA) replaces the MotA box present in P_{uvsX}, were digested with BsaAI to generate linear templates for transcription. 5′-^{32}P end-labeled P_{uvsX} DNA (a 200 bp fragment containing P_{uvsX} sequences from −94 to +83 relative to the start of transcription) that was used for the FeBABE cleavage reactions was generated by PCR and purified as described (19) except that Pfu polymerase (Agilent Technologies) was used. The 100 bp P_{uvsX} DNA, used in some of the crosslinking analyses and for transcription, was generated by annealing the following single-stranded oligonucleotides (Operon, purified by Reversed-Phase/Ion Exchange chromatography), containing positions from −66 to +34 relative to the start of transcription (phage sequences begin at position −34):

Non-template, 5′GCATGCCTGCAGGTCGACTCTA GAGGATCCTATTTGCTTAATAATCCATATGGTTA TAATAGAAATAAACCATCACATAAACGTGACC CAATAATGTGGG3′; Template,

5′CCCACATTATTGGGTCACGTTTATGTGATG GTTTATTTCTATTATAACCATATGGATTATTAAG CAAATAGGATCCTCTAGAGTCGACCTGCAGGC ATGC3′.

$pHis_6$AsiA, which contains the wt *asiA* gene with an N-terminal His_6-tag, and its derivative, $pHis_6\Delta$74-90, which lacks the C-terminal 17 codons of *asiA*, have been described (29).

pIA1000 containing genes encoding the subunits of *E. coli* RNAP (α_2, His_6−tagged β, β', ω; gift from Irina Artisimovich (40)) was transformed into BL21(DE3) cells (41). The single point mutation G1249D (G→A at position 3746 of the *rpoB* gene) was generated by Bioinnovatise within the His_6-tagged *rpoB* gene present in pIA1000. DNA sequence analysis (performed by Macrogen USA) confirmed the sequences of the wt and mutant *rpoB* gene.

$pHis_6\sigma$D581TAG is identical to $pET\sigma^{fl}CF$ (42) except that a TAG stop replaces the aspartic acid codon encoding σ residue 581. This plasmid was constructed by Genscript. pEVOL (43), which allows the incorporation of the amino acid analog BpA into *E. coli* proteins using a constructed *Methanocaldococcus jannaschii* aminoacyl-tRNA synthetase/suppressor tRNA pair, was the generous gift of Peter Schultz (Scripps Research Institute, La Jolla, CA).

Proteins

The purification of σ^{70} (44), MotA (16), N-terminal His_6-tagged σC132S/C291S/C295S/D581C (referred to as σD581C) (42), AsiA (45) and C-terminal His_6-tagged AsiA (29) have been described. *In vitro* transcription reactions comparing the activities of the conjugated proteins to that of wt were performed and analyzed as described (16).

Wild type and βG1249D mutant RNAP cores were purified as follows. Cells, started from single colonies of pIA1000/BL21(DE3) or pIA1000βG1249D/BL21(DE3), were grown in 1 L LB broth plus 50 μg/ml of kanamycin at 37°C with shaking to mid-log phase (OD_{600} of 0.5). IPTG was added to a final concentration of 1 mM, and the cell pellet was collected by centrifugation at 5000 × g for 10 min at 4°C. Lysis and purification were carried out following a modified protocol kindly provided by I. Artsimovitch (Ohio State University) and carried out on ice or at 4°C. The cell pellet was resuspended in 100 ml of lysis buffer, [50 mM Tris–Cl (pH 6.9), 500 mM NaCl, 5% glycerol] supplemented with 1 mg/ml lysozyme (Sigma L-7651) and 2 mM

benzamidine. Cells were homogenized in a Dounce homogenizer, left on ice for 20 min and supplemented with Tween 20 to a final concentration of 0.2%; further cell disruption was carried out by sonication until the OD_{600} was reduced to ∼30% of the starting value. After centrifuging two times at 27,500 x g for 30 min, a clear supernatant was obtained to which imidazole was added to a final concentration of 10 mM before application onto a HisTrap FF 5 ml column (GE Healthcare), previously equilibrated with binding buffer (lysis buffer plus 10 mM imidazole), on the ÄKTA (GE) system. Proteins were eluted with a gradient of 10–250 mM imidazole in binding buffer (20 ml) at a flow rate of 0.5 ml/min. Fractions containing RNAP (peak fractions at ∼ 200 mM imidazole) were pooled and dialyzed overnight into HepA buffer [50 mM Tris–HCl (pH 6.9), 5% glycerol, 0.5 mM EDTA, 1 mM DTT] containing 75 mM NaCl, and loaded onto a HiTrap Heparin HP 5 ml column (GE Healthcare) equilibrated in HepA buffer plus 75 mM NaCl. Protein was eluted with a gradient of 75 mM to 1.5 M NaCl in HepA buffer (100 ml) at a flow rate of 1 ml/min. Fractions enriched for RNAP (peak fractions at ∼1 M NaCl) were pooled, dialyzed into HepA buffer plus 75 mM NaCl, and loaded onto an equilibrated MonoQ 5/50 GL column (GE Healthcare). Proteins were eluted with a gradient of 75 mM to 1.5 M NaCl in HepA buffer (200 ml) at a flow rate of 1 ml/min. Highly purified RNAP (peak fractions at ∼0.4 M NaCl) was dialyzed against RNAP storage buffer [50 mM Tris–HCl (pH 7.5), 50% glycerol, 250 mM NaCl, 0.1 mM EDTA, 1 mM DTT]. The concentration of RNAP was determined after SDS-PAGE with Colloidal Coomassie Blue staining (Invitrogen) by comparing the intensity of the protein bands to known amounts of *E. coli* RNAP (Epicentre).

N-terminal His_6-tagged σD581BpA was purified from pHis$_6$σD581TAG/pEVOL/ BL21(DE3). Cells were grown with shaking at 37°C in 250 ml LB broth containing 25 μg/ml chloramphenicol and 50 μg/ml carbenicillin until mid-log phase. BpA (Bachem; 2.69 ml of 25 mg/ml in 1 M HCl, for 1 mM final) was added together with 2.69 ml of 1 M NaOH (to neutralize the HCl), 1.25 ml of 0.2 M IPTG, and 2.5 ml of 20% arabinose (wt/vol) to induce synthesis of the σ protein. Unless otherwise indicated the following steps were performed at 4°C. Cells were harvested after 90 min by centrifugation for 10 min at 13,000 × g. After resuspension in 10 ml Binding Buffer [20 mM Tris–HCl (pH 7.9), 500 mM NaCl, 5 mM imidazole, 1 mM phenylmethanesulfonyl fluoride (PMSF)], cells were lysed by sonication. The pellet obtained after centrifugation at 17,500 x g for 20 min was resuspended in 10 ml Binding Buffer and centrifuged again at 17,500 × g for 15 min. The pellet was resuspended in 4 ml of His-bind load buffer (Binding Buffer containing 6 M urea and lacking PMSF). The suspension was rotated slowly for 1 h and then centrifuged at 17,500 × g for 15 min. The supernatant was transferred to a clean tube, centrifuged again as above, filtered through a 0.8 μm syringe filter, and stored at −80°C. A 2 ml thawed aliquot of the supernatant was mixed with 2 ml of a 50% slurry of His-binding resin (Pierce HisPur Ni-NTA resin) previously equilibrated in His-bind load buffer with gentle rocking for 1 hr before transferring the resin to a 10 ml column, which was allowed to flow by gravity. The column was washed with 4 ml each of His-bind load buffer containing increasing concentrations of imidazole: 5 mM, 25 mM, 60 mM, 250 mM and 1 M. The purest protein fractions were found in the 60 and 250 mM eluents. These fractions were pooled and concentrated in an Ultra-15 centrifugal filter unit (Amicon, 30KDa MWCO) to a final volume of 3 ml, then transferred to a D-Tube Maxi Dialyzer (Novagen, 12–14 kDa MWCO) and dialyzed for at least 1 h against 1 liter of each of the following buffers to refold the protein: 1× Reconstitution Buffer [RB; 50 mM Tris–HCl (pH 8.0), 1 mM EDTA, 0.1% Triton X-100, 20% glycerol, 1 mM DTT] with 6 M urea (three times); $\frac{1}{2}$× RB with 6 M urea; $\frac{1}{2}$× RB with 3 M urea; 1× RB with 3 M urea; $\frac{1}{2}$× RB with 3 M urea; $\frac{1}{2}$× RB; 1× RB (3 times); TGED/NaCl/Triton [50 mM Tris–HCl (pH 8.0), 50 mM NaCl, 0.1 mM EDTA, 0.01% Triton X-100, 50% glycerol, 0.1 mM DTT] (four times) before storage at −20°C. *In vitro* transcription analyses using P_{uvsX} indicated that at a His_6σD581BpA:core ratio of 10:1, the unactivated level of P_{uvsX} RNA made with the mutant σ was ∼65% and the activated level (+MotA/AsiA) was 53% of that observed with wt.

FeBABE analyses

σD581C was conjugated with FeBABE ((iron (*S*)-1-(*p*-bromoacetamidobenzyl)ethylenediaminetetraacetate, Dojindo Laboratories) and FeBABE cleavage reactions were performed as described (42) except reactions contained 0.5 pmol P_{uvsX} DNA that had been ^{32}P-labeled on the template strand, 1.7 pmol of core, 1.7 pmol σD581C-FeBABE, and when indicated, 26 pmol of AsiA and/or 17 pmol of MotA in a reaction containing 11 mM Tris–Cl (pH 8), 37 mM NaCl, 20 mM potassium phosphate (pH 6.5), 0.4 mM EDTA, 0.1 mM EGTA, 0.2 mM DTT, 0.02 mM β-mercaptoethanol, 40 mM Tris-acetate (pH 7.9), 4 mM magnesium acetate, 100 μg/ml BSA, 0.0005% Triton X-100 and 15% glycerol. (AsiA and sigma were pre-incubated together first at 37°C for 10 min to ensure formation of the AsiA-sigma complex. The complex was then incubated with core at 37°C for 10 min to form the AsiA-associated complex before the addition of MotA and the DNA.)

Crosslinking with σD581BpA and native gels

For photo-crosslinking using σD581BpA and wt core in Figure 3, 38 pmol of σ (in 1 μl TGED/NaCl/Triton) were incubated with 1 μl AsiA buffer [10 mM Tris-Cl (pH 8), 0.1 mM EDTA, 50% glycerol, 500 mM NaCl, 0.1 mM DTT] with or without 65 pmol AsiA and 1 μl of 5× Kglu transcription buffer lacking BSA [40 mM Tris-acetate (pH 7.9), 150 mM potassium glutamate, 4 mM magnesium acetate, 0.1 mM EDTA, and 0.1 mM DTT] for 10 min at 37°C in 1.5 ml eppendorf tubes. RNAP core (4 pmol in 2 μl RNAP storage buffer) or buffer alone was then added and the solution incubated for 10 min at 37°C before collection on ice. Tubes were laid flat on the UV lamp for a total of 30 min at room temperature, turning the tubes to the opposite side after 15 min. Samples were electrophoresed on 10–20% Tris-tricine gels (Invitrogen) and stained with Colloidial Coomassie Blue (Invitrogen). Photocrosslinking reactions in Figure 6B were assembled similarly except re-

actions contained 120 pmol σD581BpA, 400 pmol AsiA, and 15 pmol of wt or mutant core in a total volume of 30 μl. A 20 μl aliquot was used for photocrosslinking, and a 5 μl aliquot was applied to a 4–12% Tris-glycine gel (Invitrogen/Thermo Fisher) run in 1× Native Tris-glycine buffer (Invitrogen/Thermo Fisher) and stained in Gel Code (Thermo Fisher) as described (46).

Identification of σD581BpA crosslink

Five photocrosslinked solutions containing a total of 20 pmol wt core, 190 pmol σD581BpA, and 325 pmol AsiA were combined, and the crosslinked species was separated from other species by SDS-PAGE and then in-gel trypsin-digested as described (47). A solution containing non-crosslinked control proteins was similarly reduced with DTT, alkylated with iodoacetamide, and digested in solution with trypsin at 37° C for 16 h. The resultant peptides were desalted using a C18 spin column (Thermo), diluted 1:1 in α-cyano-4-hydroxycinnamic acid, and then analyzed by MALDI-TOF mass spectrometry using positive reflectron mode on a MALDI microMX mass spectrometer (Waters).

RNA isolation/primer extension

BL21(DE3)/pLysE (41) cells containing either pet28(a+) (vector; Novagen, Inc.), pHis$_6$AsiA (29), or pHis$_6$Δ74–90 (29) were grown to mid-log phase at 37°C in LB broth containing 30 μg/ml kanamycin and 25 μg/ml chloramphenicol. After the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1.7 mM to induce the synthesis of the wild type or truncated AsiA protein, cells were grown for an additional 20 min. (Aliquots taken at this time and from uninfected cells grown for an additional 2 hr were run on protein gels to determine the levels of wt AsiA and AsiAΔ74–90.) T4 *asiA*$^-$ phage [*amS22* (48)] was added at a multiplicity of 10 and cell aliquots were taken at 2.5 and 7.5 min after infection. RNA was isolated essentially as described using Method I of Hinton (49). RNA isolated after a 4 min infection of wt T4D$^+$ of a NapIV suppressing strain has been described (50). The amount of RNA from the T4 middle promoters Pm$_{46}$ and Pm$_{55}$ was determined using primer extension analyses as described (30).

In vitro transcription assays

Transcription reactions (Figure 5B and Supplementary Figure S3) were assembled with 1 to 2 pmol of σ^{70}, 6.5×-fold AsiA (when indicated), 0.5×-fold core, 0.006–0.012 pmol DNA, and 1.83 pmol MotA (when indicated) in a solution (4 μl) containing 7 mM Tris–Cl (pH 7.9), 50 mM Tris-acetate (pH 7.9), 190 mM potassium glutamate, 5 mM magnesium acetate, 0.4 mM EDTA, 0.2 mM DTT, 125 μg/ml BSA, 50 mM potassium phosphate (pH 6.5), 0.05 mM EGTA, 22% glycerol, 0.1 mM β-mercaptoethanol, and 140 mM NaCl (or 90 mM NaCl for Supplementary Figure S3) (Unless otherwise indicated, AsiA was first incubated with σ^{70} for 10 min at 37°C and then the AsiA/σ complex was incubated at 37°C for 10 min with core to assure formation of the AsiA-associated RNAP). DNA and proteins were incubated for the indicated time at 37°C before the addition of a solution (1.0 μl) containing 0.5 μg heparin and NTPs (1 mM unlabeled ATP, GTP, CTP and 25 μM [α-^{32}P] UTP [∼1 × 10^5 dpm/pmol]) to initiate a single round of transcription. After 7.5 min at 37°C, 25 μl of gel load solution (1× TBE, 7 M urea, 1% bromophenol blue, 1% xylene cyanol FF) was added, and the reactions were placed on dry ice. Reactions with combined RNAPs (Figure 5B, lanes 9–12) contained an equal amount of wt and mutant core. RNA products were separated on 7 M urea, 4% polyacrylamide gels run in 0.5× TBE and visualized by autoradiography.

Promoter clearance assays (Supplementary Figure S2) were performed similarly except that the NTP/heparin solution contained 1 mM unlabeled GTP, CTP, UTP, 250 μM [α-^{32}P] ATP [∼1 × 10^4 dpm/pmol] and 0.5 μg heparin, and the reactions were collected on dry ice. RNA products were then treated with 10 units (1.0 μl) of Alkaline Phosphatase Calf Intestinal (CIP; New England Biolabs) for 15 min at 37°C before the addition of a solution (13 μl) containing 10 mM EDTA (pH 7), 1% bromophenol blue, and 1% xylene cyanol FF in deionized formamide. (The addition of phosphatase removes phosphate from any unincorporated [α-^{32}P] ATP, allowing one to observe the small abortive RNAs more clearly.) The RNA products were separated on 7 M urea, 23% polyacrylamide gels run in 0.5× TBE.

After autoradiography, levels of RNA products were quantified using a Powerlook 100XL densitometer and Quantity One software from Bio-Rad, Inc. The ratios of abortive products to full length RNA were determined as previously described (35).

RESULTS

Generation of a biochemically-constrained, biophysical model of the σ appropriated complex containing *E. coli* RNAP, AsiA, MotA, and T4 middle promoter DNA

The C-terminal portion of σ^{70} contains the highly conserved Region 4 (residues 540–599), consisting of helix 1, helix 2, helix 3, turn and helix 4, followed by the C-terminus that is the more variable helix 5 (H5; residues 600–613) (51). Within holoenzyme, Region 4 residues interact with the β-flap portion of core while H5 residues contact the β-flap tip (8–12,35) (Figure 1A; Supplemental Video 1). However, when bound to AsiA, residues within Region 4 are engaged by AsiA (18,21,52,53). We therefore used a multi-template homology modeling approach to produce a chimera of these two structures. Because the available structures indicated that σ^{70} residues 546 to 549 do not interact with either core or AsiA, we first superimposed these residues in the *E. coli* holoenzyme structure (9) with their equivalents in the AsiA/Region 4 structure (21), in order to reconcile the departure point of the two different structures (Figure 1B; Supplemental Video 1). Although this superposition is an assumption, one aim of this work was to test whether the hypothesis that the experimentally (X-ray, NMR) observed apo- and AsiA elements would be able to fit together well without alteration. To position DNA within this complex, we introduced the fork junction DNA present in the structure of the *Thermus aquaticus* RNAP/DNA complex (37). This was accomplished by retaining the highly conserved

contacts between residues in σ Regions 2.4/3 and the -10/extended -10 elements.

It was then necessary to add MotA^{NTD}, MotA^{CTD}, and to replace the upstream portion of the promoter DNA that contains the σ^{70}-dependent −35 element with sequence containing MotA box DNA. While it was known that MotA^{CTD} alone can weakly bind the MotA box (17), the MotA^{CTD} structure exhibits an unusual double wing motif (23) and thus, its position on the DNA could not be determined by analogy with other DNA-binding proteins. To determine how this domain interacts with the T4 middle promoter P$_{uvsX}$, we relied on fine resolution mapping obtained by conjugating specific residues within MotA with the cleaving reagent FeBABE (34). Importantly, this mapping was performed using the entire σ-appropriated complex rather than MotA and DNA alone. This work indicated that both wings of the MotA^{CTD} 'sit' within the major groove of the MotA box DNA. Cleavage sites generated by FeBABE conjugated at residues within MotA^{NTD} (D43, E93, K96) also allowed us to position the available structure of MotA^{NTD} relative to the DNA (34). Together, MotA^{NTD} and MotA^{CTD} are arranged on the DNA from positions -41 to -19. Since previous work has shown that within a T4 middle promoter, RNAP/DNA contacts downstream of ∼-20 are not affected by the presence of MotA/AsiA (54), we replaced the DNA introduced from the *T. aquaticus* structure with our mapped MotA^{CTD}/MotA^{NTD}/T4 middle promoter modeled B-form DNA starting at position -19. Finally, for visualization of full length MotA, we used loop modeling to generate the unsolved linker between the C-terminus of MotA^{NTD} and the N-terminus of MotA^{CTD}.

Previous work has defined the MotA and σ^{70} H5 interaction, indicating that F610 and L611 of σ^{70} H5 interact with a basic hydrophobic cleft within MotA^{NTD}, containing MotA residues I7, I24, A83 and Y86 (16,35). Inspection of our initial AsiA/RNAP/MotA/DNA model (Figure 1B; Supplemental Video 1) reveals that σ^{70} H5 now points toward this basic hydrophobic cleft. It should be noted that this position for H5 arose simply from the joining of the AsiA-remodeled distal σ conformation with the proximal holoenzyme conformation and the FeBABE derived location of MotA^{NTD}. It did not involve conformational manipulation of any part of the RNAP. Thus, a structure/model that significantly fulfills the protein-protein and protein-DNA requirements for σ appropriation can be assembled from crystallographic and NMR structures without molecular modeling intervention of the RNAP. In addition, since it is known that RNAP alone transcribes from the T4 middle promoter P$_{uvsX}$ (45), we also produced a structure/model of RNAP at this promoter by simply replacing the remodeled σ with the σ present in the *E. coli* structure and removing MotA and AsiA (Figure 1A and the first panel of Supplemental Video 1).

In basal transcription and in Class I and Class II activation, the C-terminal domains of the α subunits of RNAP (αCTDs) interact with DNA upstream of position -40 (55), but in σ appropriation, ADP-ribosylation of α residue Arg265, which occurs after T4 infection [reviewed in (1)], eliminates the ability of the αCTDs to interact with the DNA (54). There is no evidence that the αCTDs play a role in σ appropriation, and ADP-ribosylation is not needed

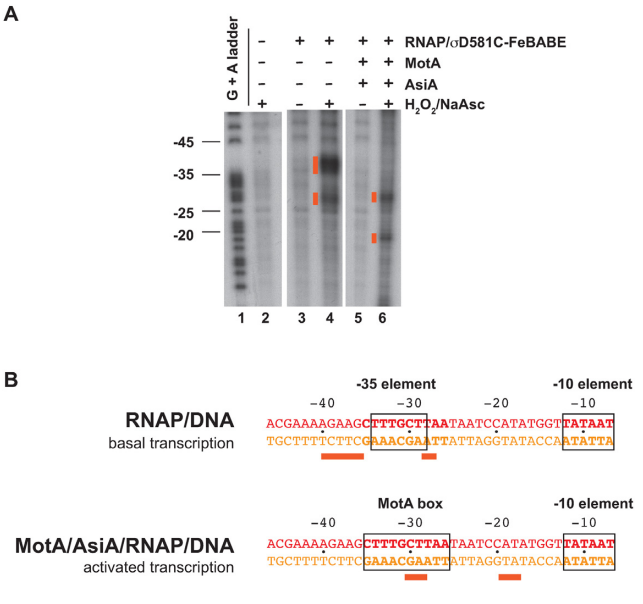

Figure 2. Binding of AsiA and MotA alters the position of σ residue 581 relative to the DNA. (**A**) Denaturing polyacrylamide gel showing DNA products generated by σD581C-FeBABE cleavage of ^{32}P 5′-end labeled template strand of P$_{uvsX}$ DNA. Lane 1 is the G + A marker ladder; lanes 2–6 are reactions containing the indicated reagents. Orange boxes indicate the cleavage products generated in the presence of H_2O_2 and sodium ascorbate (NaAsc). (**B**) Sequence of P$_{uvsX}$ DNA from -46 to -7, showing the positions of the σD581C-FeBABE generated cut sites on the template strand DNA for basal or activated transcription. The −10 element and the −35 element or the MotA box are boxed.

for MotA/AsiA activation (45), suggesting that the presence of either the unmodified or modified αCTDs do not significantly affect protein-protein or protein-DNA interactions. Consequently, because we do not have enough information at this time to position the αCTDs using either the TAP/RNAP/DNA structure (15) or the *E. coli* holoenzyme structure (9) as a guide, we have not included them in our structure/model.

Refinement of the σ appropriated model to position σ^{70} H5 within the MotA^{NTD} cleft

Because H5 residues F610 and L611 are the crucial residues for the interaction of σ with MotA^{NTD} and for MotA activation (35), it is reasonable to optimize the H5/MotA^{NTD} interaction by burying σ F610 and L611 within the MotA^{NTD} basic hydrophobic cleft. However, modeling would provide many solutions that would be consistent with such an interaction. Consequently, we determined additional biochemical constraints by obtaining the position of σ^{70} D581, which is located within the turn between helices H3 and H4 of Region 4 (56), relative to protein and to DNA.

RNAP that contains a mutant σ^{70} in which a cysteine at position 581 has been conjugated with FeBABE has been used several times previously to position σ Region 4 relative to promoter DNA (57–62). We took advantage of the fact that the T4 middle promoter P$_{uvsX}$ can be used *in vitro* as a −35/−10 promoter by RNAP alone (basal transcription) or as an activated middle promoter in the presence

of RNAP, MotA and AsiA (activated transcription) (54). This allowed us to compare the position of σ^{70} in the presence and absence of MotA/AsiA. Addition of H_2O_2 and sodium ascorbate to a stable transcription complex formed by RNAP alone with P_{uvsX} DNA resulted in major cleavage sites at −36 to −40 as well as a second set of cuts centered at ∼ -28 (Figure 2A, lane 4; Figure 2B). The major cut sites are predicted from the position of residue 581 relative to the DNA (Figure 1A; Supplemental Video 1; red sphere is D581, orange spheres show template DNA cut sites) and are similar to what has been seen previously with RNAP at this (42) and several other promoters (57–62). (The second set of cleavage sites at −27/−28, located a helical turn downstream, are typically seen in these reactions. They arise from the position of the conjugated residue relative to the major and minor grooves of the DNA (63).) However, in the presence of MotA and AsiA, the cut sites on the template strand shift ∼10 bp downstream to positions −30/−29 and −19/−18 (orange spheres in Figure 1B, left panel; Figure 2A, lane 6; Figure 2B; Supplemental Video 2).

To determine the position of σ residue 581 relative to protein, we first performed photocrosslinking experiments with σ^{70} containing the crosslinker maleimide benzophenone (MBP) conjugated through the thiol group at the cysteine at σ residue 581 (σD581C-MBP). Upon photoactivation, MBP crosslinks with moieties present within ∼8 to 11 Å by reacting with nucleophiles or by forming C-H insertion products (64–66). Using σD581C-MBP, we observed a very slowly migrating species, consistent with a crosslink between σD581C and either β or β' (Supplementary Figure S1A, lanes 6 and 8; Supplementary Figure S1B, lanes 6 and 7). Although the intensity of the band was low, which is a common problem with benzophenone crosslinking (67), the band was specific and reproducible. Western blot analyses were positive for both β' (Supplementary Figure S1A, lanes 14 and 16) and σ^{70} (Supplementary Figure S1B, lanes 10 and 11), but negative for β (not shown). This species was absolutely dependent on the presence of AsiA (Supplementary Figure S1A, lanes 4 versus 8), but was observed in the presence or absence of DNA (Supplementary Figure S1B, lanes 6 versus 7) or MotA (Supplementary Figure S1A, lanes 6 versus 8). Thus, the σD581C-MBP/β' crosslink is dependent on a conformational change imposed by AsiA binding to σ Region 4.

Our analysis with σD581C-MBP also detected a species of ∼80 kDa, consistent with a crosslink between σ and AsiA (smaller arrow, Supplementary Figure S1B, lanes 6 and 7). Although this species was dependent on the presence of AsiA and its size was consistent with a predicted σ^{70}-AsiA crosslink, we were unable to confirm that it contained AsiA by Western, presumably because its level was below our detection limit.

To increase the amount of crosslinking and to perform mass spectrometric analyses of the slowly migrating crosslinked species, we also performed photocrosslinking with σ^{70} protein containing the amino acid analog benzoylphenylalanine (BpA) at position 581. In this case, we obtained a higher yield of the large species, which migrated like the one seen with σD581C-MBP, and again it was dependent on the presence of AsiA (Figure 3A). A comparison of the in-gel tryptic peptides of this species with the tryptic peptides arising from digestion of a solution of the proteins that were not subjected to UV treatment revealed the unique product, β'HRGVICEK+σQF-BPA-VTR (Figure 3B). This indicates that the σ BpA at position 581 crosslinks to the region of β' containing residues 80–87.

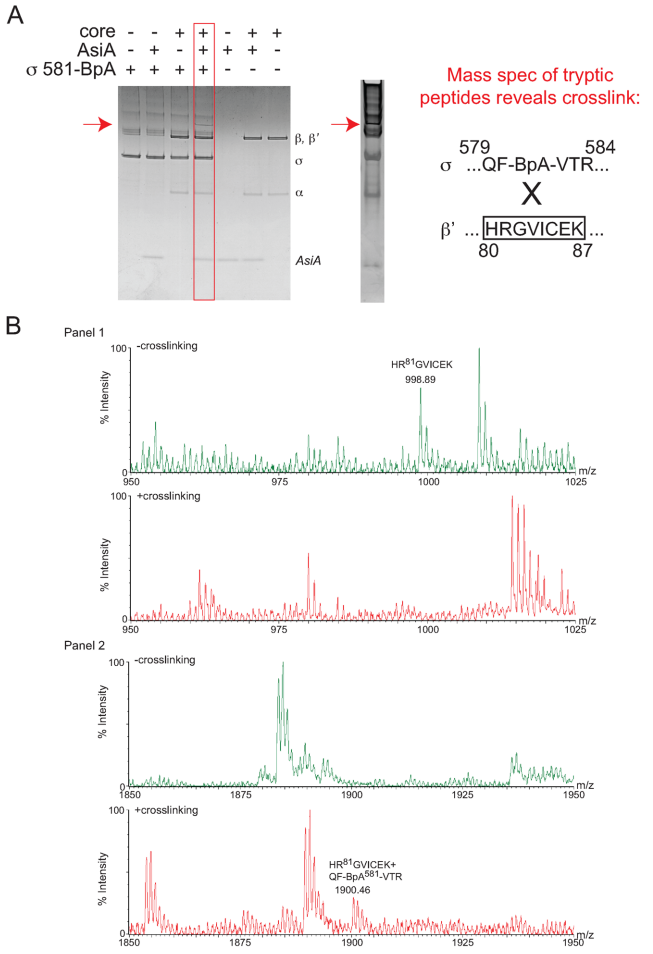

Figure 3. σD581BpA photocrosslinks to the region of β' containing residues 80–87. (**A**) Left, Coomassie-stained 10–20% Tris-tricine gel showing the species obtained after UV treatment of a solution containing the indicated proteins. The slowly migrating band observed in the presence of σD581BpA, core, and AsiA is indicated by the red arrow. Right, Coomassie-stained gel showing the species removed for digestion with typsin, followed by MALDI-TOF mass spectrometry. The sequence of the identified crosslinked peptide between σ and β' is shown. (**B**) Mass spectra showing crosslink formation between σD581BpA and β'. (Panel 1) Mass spectra of the tryptic digests of the non-crosslinked proteins (top) and crosslinked forms (bottom) from $m/z = 950–1025$. In the non-crosslinked spectrum, the mass of the β' peptide ^{80}HRGVICEK^{87} is labeled; this ion is not observed upon crosslinking. (Panel 2) Mass spectra of the tryptic digests of the non-crosslinked proteins (top) and crosslinked forms (bottom) from m/z = 1850-1950. In the crosslinked spectrum, a unique ion with $m/z = 1900.46$ is noted. This mass is consistent with that expected for β' ^{80}HRGVICEK^{87} crosslinked to sigma ^{579}QF-BpA-VTR^{584}.

We used a constrained energy minimization approach to determine if the initial 3D conformation of the σ-appropriated complex could accommodate these biochemical constraints. Conformations with good van der Waals, electrostatic and torsion energy that also satisfy the distance restraints derived from the biochemistry were sampled (see

Methods). A refined model, in which σ F610 and L611 are buried within the MotA^{NTD} hydrophobic cleft, was obtained with reasonable peptide geometry and minimal energy strain that also satisfied the distance restraints (Figure 1C and D; Supplemental Video 2). In this model, σD581 is 11Å from β' residue R81, DNA positions −29/−30 are the closest template bases to σD581, and σ H5 is buried within the hydrophobic cleft of MotA^{NTD}. It should be noted, though, that the constraints used here arise from the FeBABE and crosslinking analyses that would have captured the most stable complex. Thus, other positions may be possible during the various steps of transcription initiation.

As predicted by the model, interactions involving the C-terminal residues of AsiA are not essential for middle promoter activation

Although the N-terminal half of AsiA is required for its interaction with σ Region 4 (21,29,68), a function for the C-terminal region of the protein has not been clear. Our structure/model (Figure 1C, right) predicts that there are no close contacts between the C-terminal portion of AsiA (residues 74–90, shown in light blue) and either MotA or subunits of RNAP. This is consistent with NMR analyses, which detect the interaction of MotA with σ H5, but do not reveal additional chemical shifts in the AsiA protein when MotA is added to the AsiA/σ Region 4 complex (16). It is also consistent with *in vitro* transcription studies demonstrating that mutations within or a deletion of C-terminal residues 74–90 of AsiA do not significantly impair inhibition of transcription by AsiA in the absence of MotA or activation in the presence of MotA (28,29). However, we needed to consider conflicting evidence from other work concluding that a mutation at either R82 or M86 of AsiA disrupts an 'essential' interaction between AsiA and MotA (26) and that AsiA N74 lies at an interface between AsiA and the β-flap (27). We were concerned then that these interactions might be needed under the biological conditions of phage infection and so they were missed in the *in vitro* studies. If so, then our model would be insufficient. Therefore, we assayed the need for the C-terminal portion of AsiA *in vivo.*

We performed primer extension analyses to observe the effect of removing the C-terminal 17 residues (AsiAΔ74–90) on AsiA's ability to complement T4 *asiA am* phage under non-suppressing conditions in the activation of two T4 middle promoters, Pm_{46} and Pm_{55} (Figure 4). As controls, we monitored the level of transcripts from these promoters after infection of cells containing the vector or a plasmid with wt *asiA.* We also tested RNA isolated after a wt T4 infection that had been previously analyzed (50) to confirm the correct assignment of the primer extension products.

We found that AsiAΔ74-90 complements T4 *asiA am* for transcription from either promoter. For Pm_{55} RNA, the level obtained with either wt or AsiAΔ74-90 was similar at both 2.5 and 7.5 min post-infection (Figure 4, lanes 2, 3 versus 5, 6). Although 2.5 min after infection, the level of Pm_{46} RNA was lower with AsiAΔ74-90 (Figure 4, lanes 2 versus 3), at 7.5 min the level seen with either wt or AsiAΔ74-90 was again similar (Figure 4, lanes 5 versus 6). SDS-PAGE

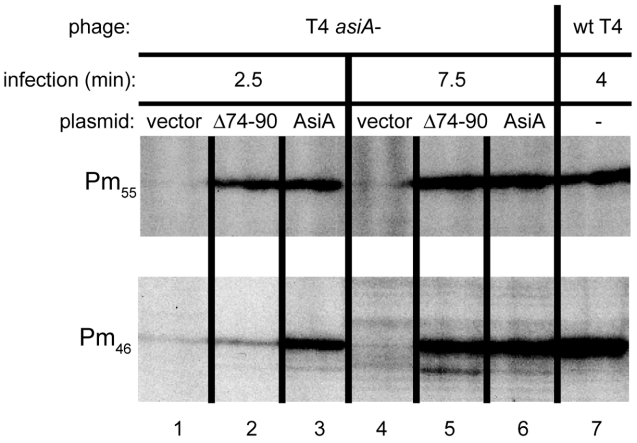

Figure 4. The C-terminal 17 residues of AsiA are not required for middle promoter activation *in vivo*. Shown are portions of gels displaying the primer extension products for the T4 middle promoters Pm_{55} and Pm_{46} obtained from RNA isolated after infection with either T4 *amS22 (asiA$^-$)* or wt T4. Cells contained the empty vector, a plasmid with wt *asiA*, or a plasmid with mutant *asiA* lacking the coding region for the 17 C-terminal residues (Δ74–90) and were infected for the indicated time.

(not shown) indicated that the level of AsiAΔ74-90 after induction is actually only about half that of wt. This indicated that the activity of the truncated protein is not simply because a high level of the mutant compensates for poor activity. Taken with our previous work, these results demonstrate that the C-terminal region of AsiA is not essential for σ appropriation *in vivo* or *in vitro.* The results support our modeled structure, which indicates that the C-terminal region of AsiA does not interact with MotA or the β-flap. However, we cannot rule out the possibility that non-essential contacts between AsiA and MotA or AsiA and the β-flap may form at some point during the various steps of transcription initiation.

Using the model to investigate how the βG1249D substitution impairs middle promoter activation

An amber mutation within either T4 *motA* or *asiA* generates very low levels of MotA or AsiA, respectively, resulting in poor, but detectable, phage growth after infection of wt *E. coli.* However, T4 *motA*am or *asiA*am phage yield only pinpoint plaques when plated on an *E. coli* mutant strain containing the G1249D substitution within β (30,48,69). Primer extension analyses have revealed that βG1249D specifically impairs transcription from phage middle promoters; after infection there is no decrease in transcription from early phage promoters (30). In addition, the βG1249D mutation does not exhibit a defect in host growth in the absence of infection (30).

The position of βG1249 within our model is far removed from AsiA, MotA, DNA, and σ H5 (Figure 5A, left panel), suggesting that the mutation should not directly interfere with interactions among these protein or DNA targets. However, we wanted to eliminate the formal possibilities that (i) there is an unidentified protein, involved in σ appropriation, which directly contacts βG1249 *in vivo* and/or

(ii) the modification of wt T4 DNA by the presence of glycosylated, hydroxymethylated cytosines is involved.

To investigate these possibilities, we asked whether the βG1249D impairment in middle promoter activation is recapitulated using purified mutant RNAP and unmodified DNA. As seen in Figure 5B, the effect of the mutant RNAP on middle promoter activation *in vitro* is dramatic. While the presence of MotA/AsiA significantly increased transcription when using wt β, there was practically no increase with βG1249D over the unactivated level (Figure 5B, lanes 4 versus lane 8). To make sure that our purified G1249D core did not contain a nonspecific transcription inhibitor that was absent in the purified wt core, we performed a reaction using an equal amount of wt and mutant polymerases. In this case, no impairment was observed (Figure 5B, lane 12). Interestingly, the presence of the mutation *in vitro* did result in a 2- to 4-fold decrease in basal P_{uvsX} transcription obtained in the absence of MotA/AsiA (Figure 5B, lane 1 versus lane 5; other experiments not shown), indicating that although we did not observe a growth defect with mutant cells *in vivo*, the mutation does somewhat impair RNAP function under our conditions *in vitro*.

Given that the impairment of σ appropriation by βG1249D does not require other factors besides RNAP, MotA, and AsiA and it does not require modification of T4 DNA, we considered other possibilities for how this substitution might be deleterious. Within the *E. coli* RNAP structure (9), βG1249 is ∼7 to 9 Å from three negatively charged σ residues located just before Region 4: D525, E529, and D533 (Figure 5A, right panel). Thus, it seemed possible that the substitution of the highly flexible glycine at 1249 with the negatively charged aspartic acid, which can adopt fewer unclashed φ and ϕ backbone dihedral angles, could electrostatically and/or sterically inhibit conformational changes needed for MotA/AsiA activated transcription. Thus, we asked whether the mutation affected (i) the ability of AsiA/σ to form a complex with core, (ii) the conformation of the AsiA-associated RNAP or (iii) promoter clearance.

The βG1249D does not impair promoter clearance

Within the structure of elongating RNAP, βG1249 is located at the upstream edge of the RNA/DNA hybrid, immediately adjacent to switch loop 3; this loop has been proposed to help guide the extruding RNA into the channel (70). Thus, it was possible that the presence of MotA and AsiA might limit the flexibility of sigma in such a way that the βG1249D substitution would inhibit promoter clearance. To measure the effect of βG1249D on promoter clearance, we performed single round *in vitro* transcription reactions using short templates that allowed us to measure both full length and abortive RNA products. In this assay, a lower ratio of abortives relative to full length RNA indicates more efficient promoter clearance (35,71,72). In the absence of MotA and AsiA (basal transcription), RNAP containing βG1249D was somewhat less efficient than wt in clearing the promoter, which could contribute to the decrease in basal transcription we observe *in vitro* (Supplementary Figure S2, basal transcription). However, in the presence of MotA/AsiA, the abortive/full length RNA ratio was similar when using either core (Supplementary Figure S2, activated transcription) even though the overall total amount of transcription was much lower when using βG1249D (Supplementary Figure S2A; note the differences in the intensity values for the wt vs. G1249D scans).

The βG1249D core generates a conformationally different complex with AsiA/σ compared to wt

The transcriptionally competent AsiA/RNAP/MotA/DNA complex is generated by a multi-step process. AsiA first binds to free σ, and then the AsiA/σ complex binds to core (39), generating AsiA-associated RNAP holoenzyme. AsiA-RNAP must then 'capture' MotA while it is at the promoter, since MotA rapidly associates/dissociates from the DNA (34).

To ask whether βG1249D interferes with the association of AsiA/σ with core, we employed a Tris-glycine native gel. In this system, holoenzyme migrates as a discrete band that is separated from core, AsiA, or σ^{70} [(46); Figure 6A, lanes 1, 2 versus 3). This analysis indicated that a similar amount of holoenzyme is generated with either wt or mutant core (Figure 6A, lanes 3, 5 versus 7). To investigate the generation of AsiA-associated RNAP in a totally different way, we used *in vitro* transcription reactions. We took advantage of a previous finding, demonstrating that AsiA-associated RNAP eventually forms an open complex with a strong −10/−35 promoter if given enough time (29,73). This property indirectly measures the presence of holoenzyme, since it is dependent on the generation of a holoenzyme containing core/AsiA/σ. We investigated the levels of transcription from the strong promoter $P_{uvsX\text{-}sigma}$, which has consensus −10 and −35 sequences. As seen in Supplementary Figure S3, the relative levels of transcription by AsiA-associated RNAP are similar with either the wt or mutant core. In addition, this behavior is the same whether AsiA/σ has been incubated with core for 10 or 60 min before transcription. Taken together, these analyses then indicate that both the wt and mutant core are able to generate an AsiA-associated complex.

Given that AsiA/σ can form a complex with either core, why then is the mutant defective in σ appropriation? To ask whether there is a difference in the conformation of the AsiA-associated RNAP when the substitution is present, we repeated the σ581BpA photocrosslinking experiments, comparing wt and βG1249D core (Figure 6B). Before the crosslinking, samples were split, and one aliquot was used for the native gel in Figure 6A. We found that even though the same amount of holoenzyme was formed with either core (Figure 6A), the β substitution resulted in ∼ 4-fold less of the crosslinked product. We conclude that even though the βG1249D core forms an AsiA-associated complex, the conformation of the complex differs from that made with wt core.

DISCUSSION

Modeling of σ appropriation allows visualization of a multifactor transcription complex

The elegant systems by which T4 hijacks the host RNAP offer insight into the basic mechanisms of transcription [re-

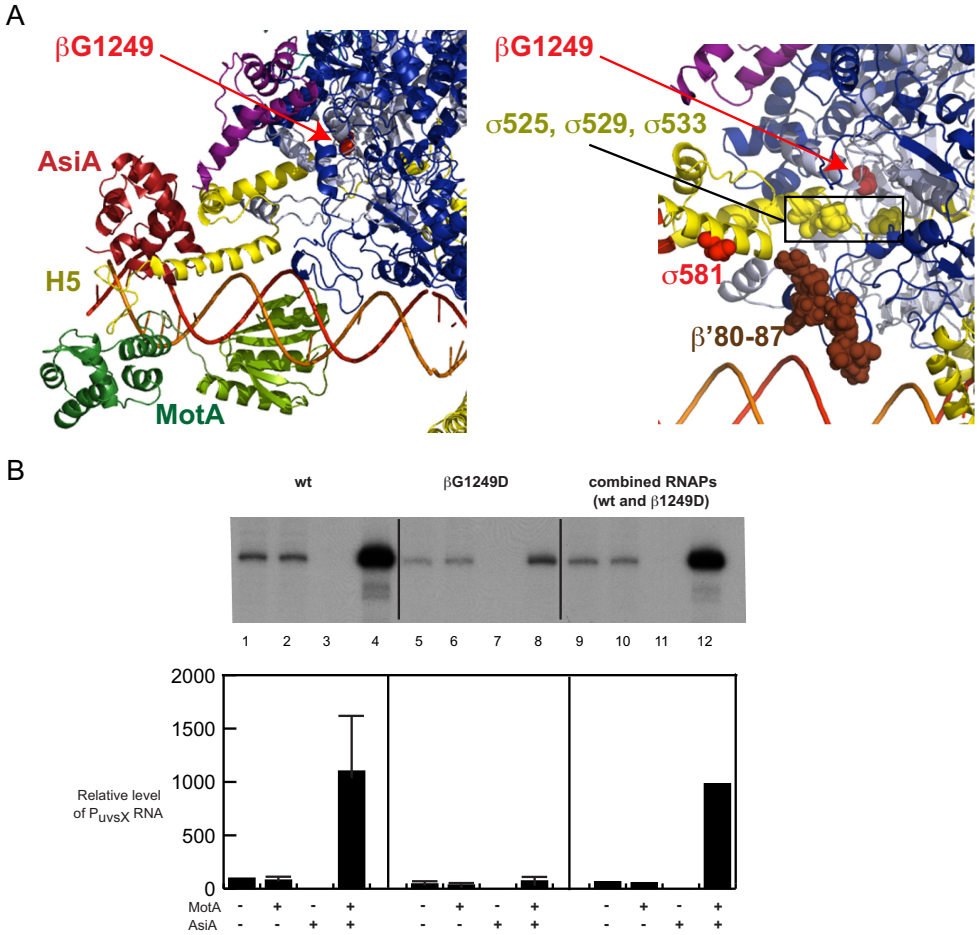

Figure 5. βG1249D impairs σ appropriation *in vitro*. (**A**) Left panel. Close-up of the refined structure/model of *E. coli* RNAP/AsiA/MotA/DNA (same as Figure 1C) showing the position of βG1249 as a red sphere. Right Panel. Close up showing the positions of σD525, σE529, and σD533 (boxed yellow spheres) relative to βG1249 (red sphere indicated with red arrow). The positions of β' residues 80–87 (brown spheres) and σD581 (red sphere) are also shown. (**B**) Comparison of *in vitro* transcription using core containing wt β or βG1249D. The top is a representative transcription gel, showing P_{uvsX} RNA obtained after single round transcription reactions in which the indicated proteins were incubated with the DNA for 1 min at 37°C before the addition of NTPs and heparin. Graph below indicates the relative level of P_{uvsX} RNA. Standard deviations shown in lanes 2–8 were obtained from 4 experiments; lanes 9–12 represent the average of 2 experiments.

viewed in (1,74)]. For example, the activation of T4 middle promoters through the process of σ appropriation highlights the importance of σ^{70} Region 4/H5 for contacting the DNA, RNAP core, and regulatory factors. Extensive biochemical, genetic, and structural analyses have identified crucial features as well as protein and DNA surfaces involved in σ appropriation. However, individual structures of $MotA^{NTD}$, of $MotA^{CTD}$, of AsiA bound to σ^{70}, and of RNAP holoenzyme by themselves cannot reveal how these various molecular surfaces work together to take over RNAP. Thus, the combination of extensive biochemical evidence together with structural pieces of the complex has provided a 'jigsaw puzzle' waiting to be pieced together. The structure and biochemistry-based modeling here provides a first visualization of this T4-appropriated transcriptional machine.

What are the new insights gained by having this visualization? First, the model together with biochemistry reveals that flexibility within the σ Region 4/core interaction is needed to facilitate the formation of the AsiA/RNAP/MotA/DNA complex. This conclusion arises from comparison of our initial model with the refined model. In the initial model, σ^{70} H5, which was previously buried under the β-flap helix tip within host RNAP (9), is now exposed in a position that is distant from the β subunit and pointed toward its target, the hydrophobic cleft within $MotA^{NTD}$ (Figure 1B). This model, which does not require distortion of the protein or DNA, places H5 in a reasonable position. However, it does not position σ^{70} residues F610 and L611, which are crucial for MotA activation and are needed for the interaction of H5 with MotA (16,35), directly within the $MotA^{NTD}$ pocket, and it does not correctly locate σ residue 581 relative to protein and DNA. To achieve these interactions, it is necessary to move the protein, DNA, or both. Previous work has indicated that MotA binding does not significantly distort the DNA (19,20,75). Thus, it appears that protein flexibility within σ Region 4/H5 is needed to bury σ^{70} F610/L611 within the $MotA^{NTD}$ cleft and to orient σ Region 4/H5 correctly. In this way, H5, once re-

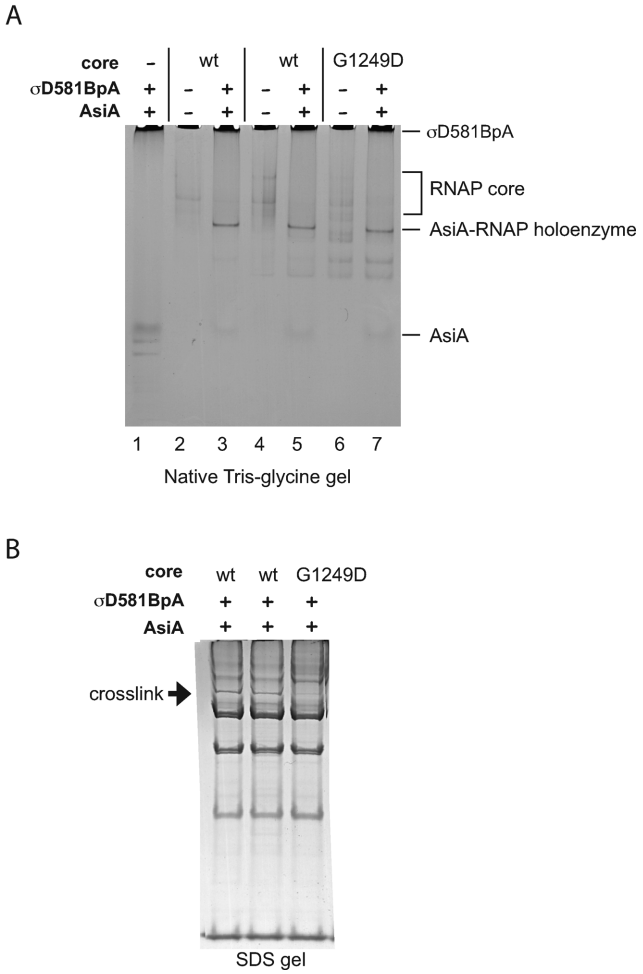

Figure 6. RNAP containing βG1249D generates holoenzyme with AsiA/σD581BpA, but is defective in generating the crosslink with β'. (**A**) Native Tris-glycine gel. Solutions were assembled with the indicated components. The positions of AsiA, RNAP core, AsiA-RNAP holoenzyme, and σD581BpA are marked. (The bands that migrate faster than AsiA seen in lane 1 are trace contaminants present in the σD581BpA preparation.) (**B**) SDS-PAGE gel showing the products obtained after photocrosslinking. Arrow points to the crosslink between σD581BpA and β' ^{80}HRGVICEK^{87} identified in Figure 3.

leased from the grip of the β-flap tip, can now become an adaptable target for regulation.

This conclusion is further supported by our analysis of how βG1249D, which is far removed from MotA and AsiA, significantly decreases MotA activation. Our work demonstrates the following: (i) The βG1249D substitution does not change the ratio of abortives to full length RNA synthesized by the MotA/AsiA/RNAP/DNA complex (Supplementary Figure S2). Thus, impairment by βG1249D must be at a step before the formation of the initiated complex. (ii) The βG1249D substitution does not impair the ability of AsiA/σ to bind to core (Figure 6A) or the ability of AsiA to inhibit transcription (Supplementary Figure S3). Thus, the impairment affects a step after formation of the AsiA-associated RNAP. From these experiments, we conclude that the βG1249D defect occurs at a step needed for AsiA-RNAP to associate with MotA/DNA. Finally, we observe that the AsiA-associated mutant RNAP generates much less of the crosslink between σD581BpA and β' than is seen with the AsiA-associated wt RNAP (Figure 6B). Taken all together, these results suggest that something different about the position of Region 4/H5 when G1249D is present leads to a defect in the association of AsiA-RNAP with MotA/DNA.

Because MotA associates and dissociates rapidly from the DNA (34), MotA cannot function by binding tightly to the DNA and then delivering the DNA to AsiA-RNAP. Instead, the AsiA-associated RNAP must 'capture' the MotA while MotA is on the DNA, through the interaction of σ H5 with MotA^{NTD}. Given the close proximity of βG1249 to the negatively charged σ residues D525, E529, and D533 (Figure 5A, right panel), it seems likely that the substitution of glycine with the negatively charged aspartic acid inhibits the flexibility needed to position Region4/H5 in the ideal location for the H5/MotA^{NTD} interaction. Thus, we speculate that the presence of the βG1249D substitution generates an AsiA-associated RNAP, whose H5 location is suboptimal for its connection to MotA^{NTD}.

Visualizing the T4 appropriated σ in context of the complete RNAP complex also explains why AsiA efficiently binds free σ, but not holoenzyme (39), since AsiA is not capable of simply replacing the β-flap *in situ*. Indeed, in our model, as in its isolated complex with σ^{70}, AsiA does not even contact the β subunit of the complex. Some work has suggested that the C-terminal portion of AsiA makes important contacts with the β-flap and with MotA (26,27). Although we cannot eliminate the possibility that these interactions can occur during the process, our model, other studies (16,28,29), and our primer extension results shown here argue that such interactions are not essential.

Our model reveals how the unusual DNA binding motif of MotA positions this activator relative to DNA, to its binding partner, σ H5, and to RNAP in general. The 'double wing' motif of MotA^{CTD} is novel among DNA binding proteins, and to date identification of σ appropriation has been limited to phage T4. However, besides the dozens of T4-like phage MotA homologs, the unusual MotA^{CTD} structural motif has been found in the structures of two proteins of unknown function, the conserved *E. coli* YjbR (25) and the *Pseudomonas syringae* protein Pspto_3016 (24). As the founding member of this family of protein motifs, MotA provides a basis for rational predictions of protein function for these and other as yet unidentified members of the family.

Finally, these results demonstrate the power of combining molecular modeling with extensive biochemical constraints to generate a meaningful and biologically relevant structure/model. Given the inability to obtain structures for many multi-subunit protein/DNA complexes, our work shows how one can develop a structure/model that satisfies detailed *in vitro* and *in vivo* findings.

Implications for σ^{70} function and RNAP

While our visualization is likely only one snapshot of a dynamic 'movie' of the σ^{70} appropriation process, our structure/model illustrates how σ^{70} Region 4/H5 can be displaced to a new location without significant alterations

of σ^{70} Region 2, which maintains its contacts with the downstream core promoter elements. Thus, the remodeling of a small portion of RNAP can completely alter its promoter specificity, switching polymerase from the recognition of σ^{70}-dependent host promoters to recognition of T4 middle promoters. In this regard, MotA and AsiA can be thought of as factors that together with σ^{70} create a new σ factor rather than simply as an individual activator and co-activator.

The ability to create a different σ depends on the flexibility of σ^{70} Region 4 and H5, but also on the continued maintenance of the typical contacts of Regions 2 and 3 with core and the DNA. Regions 2–4 are conserved among hundreds of identified σ factors (51). It seems likely then that other as yet unidentified factors may also take advantage of the fact that core can 'hold' Regions 2 and 3, while allowing a rearrangement of Region 4. Such a rearrangement may not maintain the same molecular details as those seen here. Nonetheless, the structured model for σ appropriation yields insights into the general features that such a remodeling could contain.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank K. Murakami for giving us access to unpublished data, Steve White for helpful discussions, Irina Artsimovitch for pIA1000 and the protocol for RNAP purification, Peter Schultz for pEVOL, and Leslie Knipling for technical assistance.

FUNDING

Intramural Research Program of the National Institutes of Health (to T.D.J., L.E.A., M.-L.H., L.M.M.J., S.S.J. and D.M.H.); National Institutes of Health [DP2 OD004631 to T.C.]. Funding for open access charge: The Intramural Research Program of the National Institutes of Health.
Conflict of interest statement. None declared.

REFERENCES

1. Hinton,D.M. (2010) Transcriptional control in the prereplicative phase of T4 development. *Virol J*, **7**, 289.
2. Decker,K.B. and Hinton,D.M. (2013) Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. *Annu. Rev. Microbiol.*, **67**, 113–139.
3. Saecker,R.M., Record,M.T. Jr and Dehaseth,P.L. (2011) Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J. Mol. Biol.*, **412**, 754–771.
4. Murakami,K.S. (2015) Structural biology of bacterial RNA polymerase. *Biomolecules*, **5**, 848–864.
5. Feklistov,A., Sharon,B.D., Darst,S.A. and Gross,C.A. (2014) Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.*, **68**, 357–376.
6. Paget,M.S. (2015) Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules*, **5**, 1245–1265.
7. Hook-Barnard,I.G. and Hinton,D.M. (2007) Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regul. Syst. Biol.*, **1**, 275–293.
8. Murakami,K.S., Masuda,S. and Darst,S.A. (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science*, **296**, 1280–1284.
9. Murakami,K.S. (2013) X-ray crystal structure of Escherichia coli RNA polymerase sigma70 holoenzyme. *J. Biol. Chem.*, **288**, 9126–9134.
10. Vassylyev,D.G., Sekine,S., Laptenko,O., Lee,J., Vassylyeva,M.N., Borukhov,S. and Yokoyama,S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature*, **417**, 712–719.
11. Kuznedelov,K., Minakhin,L., Niedziela-Majka,A., Dove,S.L., Rogulja,D., Nickels,B.E., Hochschild,A., Heyduk,T. and Severinov,K. (2002) A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition. *Science*, **295**, 855–857.
12. Geszvain,K., Gruber,T.M., Mooney,R.A., Gross,C.A. and Landick,R. (2004) A hydrophobic patch on the flap-tip helix of E.coli RNA polymerase mediates sigma(70) region 4 function. *J. Mol. Biol.*, **343**, 569–587.
13. Lee,D.J., Minchin,S.D. and Busby,S.J. (2012) Activating transcription in bacteria. *Annu. Rev. Microbiol.*, **66**, 125–152.
14. Blanco,A.G., Canals,A., Bernues,J., Sola,M. and Coll,M. (2011) The structure of a transcription activation subcomplex reveals how sigma(70) is recruited to PhoB promoters. *EMBO J.*, **30**, 3776–3785.
15. Feng,Y., Zhang,Y. and Ebright,R.H. (2016) Structural basis of transcription activation. *Science*, **352**, 1330–1333.
16. Bonocora,R.P., Caignan,G., Woodrell,C., Werner,M.H. and Hinton,D.M. (2008) A basic/hydrophobic cleft of the T4 activator MotA interacts with the C-terminus of E.coli sigma70 to activate middle gene transcription. *Mol. Microbiol.*, **69**, 331–343.
17. Pande,S., Makela,A., Dove,S.L., Nickels,B.E., Hochschild,A. and Hinton,D.M. (2002) The bacteriophage T4 transcription activator MotA interacts with the far-C-terminal region of the sigma70 subunit of Escherichia coli RNA polymerase. *J. Bacteriol.*, **184**, 3957–3964.
18. Baxter,K., Lee,J., Minakhin,L., Severinov,K. and Hinton,D.M. (2006) Mutational analysis of sigma(70) Region 4 needed for appropriation by the bacteriophage T4 transcription factors AsiA and MotA. *J. Mol. Biol.*, **363**, 931–944.
19. March-Amegadzie,R. and Hinton,D.M. (1995) The bacteriophage T4 middle promoter PuvsX: analysis of regions important for binding of the T4 transcriptional activator MotA and for activation of transcription. *Mol. Microbiol.*, **15**, 649–660.
20. Schmidt,R.P. and Kreuzer,K.N. (1992) Purified MotA protein binds the -30 region of a bacteriophage T4 middle-mode promoter and activates transcription in vitro. *J. Biol. Chem.*, **267**, 11399–11407.
21. Lambert,L.J., Wei,Y., Schirf,V., Demeler,B. and Werner,M.H. (2004) T4 AsiA blocks DNA recognition by remodeling sigma(70) region 4. *EMBO J.*, **23**, 2952–2962.
22. Stevens,A. (1976) A salt-promoted inhibitor of RNA polymerase isolated from T4 phage-infected E. coli. *RNA Polymerase*, 617–627.
23. Li,N., Sickmier,E.A., Zhang,R., Joachimiak,A. and White,S.W. (2002) The MotA transcription factor from bacteriophage T4 contains a novel DNA-binding domain: the 'double wing' motif. *Mol. Microbiol.*, **43**, 1079–1088.
24. Feldmann,E.A., Seetharaman,J., Ramelot,T.A., Lew,S., Zhao,L., Hamilton,K., Ciccosanti,C., Xiao,R., Acton,T.B., Everett,J.K. *et al.* (2012) Solution NMR and X-ray crystal structures of Pseudomonas syringae Pspto_3016 from protein domain family PF04237 (DUF419) adopt a "double wing" DNA binding motif. *J. Struct. Funct. Genomics*, **13**, 155–162.
25. Singarapu,K.K., Liu,G., Xiao,R., Bertonati,C., Honig,B., Montelione,G.T. and Szyperski,T. (2007) NMR structure of protein yjbR from Escherichia coli reveals 'double-wing' DNA binding motif. *Proteins*, **67**, 501–504.
26. Yuan,A.H. and Hochschild,A. (2009) Direct activator/co-activator interaction is essential for bacteriophage T4 middle gene expression. *Mol. Microbiol.*, **74**, 1018–1030.
27. Yuan,A.H., Nickels,B.E. and Hochschild,A. (2009) The bacteriophage T4 AsiA protein contacts the beta-flap domain of RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 6597–6602.
28. Pineda,M., Gregory,B.D., Szczypinski,B., Baxter,K.R., Hochschild,A., Miller,E.S. and Hinton,D.M. (2004) A family of anti-sigma70 proteins in T4-type phages and bacteria that are similar

to AsiA, a Transcription inhibitor and co-activator of bacteriophage T4. *J. Mol. Biol.*, **344**, 1183–1197.
29. Pal,D., Vuthoori,M., Pande,S., Wheeler,D. and Hinton,D.M. (2003) Analysis of regions within the bacteriophage T4 AsiA protein involved in its binding to the sigma70 subunit of E. coli RNA polymerase and its role as a transcriptional inhibitor and co-activator. *J. Mol. Biol.*, **325**, 827–841.
30. James,T.D., Cashel,M. and Hinton,D.M. (2010) A mutation within the beta subunit of Escherichia coli RNA polymerase impairs transcription from bacteriophage T4 middle promoters. *J. Bacteriol.*, **192**, 5580–5587.
31. Opalka,N., Brown,J., Lane,W.J., Twist,K.A., Landick,R., Asturias,F.J. and Darst,S.A. (2010) Complete structural model of Escherichia coli RNA polymerase from a hybrid approach. *PLoS Biol.*, **8**, e1000483.
32. Decker,K.B., James,T.D., Stibitz,S. and Hinton,D.M. (2012) The Bordetella pertussis model of exquisite gene control by the global transcription factor BvgA. *Microbiology*, **158**, 1665–1676.
33. Li,N., Zhang,W., White,S.W. and Kriwacki,R.W. (2001) Solution structure of the transcriptional activation domain of the bacteriophage T4 protein, MotA. *Biochemistry*, **40**, 4293–4302.
34. Hsieh,M.L., James,T.D., Knipling,L., Waddell,M.B., White,S. and Hinton,D.M. (2013) Architecture of the bacteriophage T4 activator MotA/promoter DNA interaction during sigma appropriation. *J. Biol. Chem.*, **288**, 27607–27618.
35. Bonocora,R.P., Decker,P.K., Glass,S., Knipling,L. and Hinton,D.M. (2011) The Bacteriophage T4 activator MotA and the {beta}-flap tip of RNA polymerase target the same Set of {sigma}70 C-terminal residues. *J. Biol. Chem.*, **286**, 39290–39296.
36. Abagyan,R. and Totrov,M. (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.*, **235**, 983–1002.
37. Murakami,K.S., Masuda,S., Campbell,E.A., Muzzin,O. and Darst,S.A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*, **296**, 1285–1290.
38. Cardozo,T., Totrov,M. and Abagyan,R. (1995) Homology modeling by the ICM method. *Proteins*, **23**, 403–414.
39. Hinton,D.M. and Vuthoori,S. (2000) Efficient inhibition of Escherichia coli RNA polymerase by the bacteriophage T4 AsiA protein requires that AsiA binds first to free sigma70. *J. Mol. Biol.*, **304**, 731–739.
40. Belogurov,G.A., Vassylyeva,M.N., Svetlov,V., Klyuyev,S., Grishin,N.V., Vassylyev,D.G. and Artsimovitch,I. (2007) Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol. Cell*, **26**, 117–129.
41. Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, **185**, 60–89.
42. Decker,K.B., Chen,Q., Hsieh,M.L., Boucher,P., Stibitz,S. and Hinton,D.M. (2011) Different requirements for sigma Region 4 in BvgA activation of the Bordetella pertussis promoters Pfim3 and PfhaB. *J. Mol. Biol.*, **409**, 692–709.
43. Young,T.S., Ahmad,I., Yin,J.A. and Schultz,P.G. (2010) An enhanced system for unnatural amino acid mutagenesis in E. coli. *J. Mol. Biol.*, **395**, 361–374.
44. Gerber,J.S. and Hinton,D.M. (1996) An N-terminal mutation in the bacteriophage T4 motA gene yields a protein that binds DNA but is defective for activation of transcription. *J. Bacteriol.*, **178**, 6133–6139.
45. Hinton,D.M., March-Amegadzie,R., Gerber,J.S. and Sharma,M. (1996) Bacteriophage T4 middle transcription system: T4-modified RNA polymerase; AsiA, a sigma 70 binding protein; and transcriptional activator MotA. *Methods Enzymol.*, **274**, 43–57.
46. Glaser,B.T., Bergendahl,V., Anthony,L.C., Olson,B. and Burgess,R.R. (2009) Studying the salt dependence of the binding of sigma70 and sigma32 to core RNA polymerase using luminescence resonance energy transfer. *PLoS One*, **4**, e6490.
47. Shevchenko,A., Tomas,H., Havlis,J., Olsen,J.V. and Mann,M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.*, **1**, 2856–2860.
48. Ouhammouch,M., Orsini,G. and Brody,E.N. (1994) The asiA gene product of bacteriophage T4 is required for middle mode RNA synthesis. *J. Bacteriol.*, **176**, 3956–3965.
49. Hinton,D.M. (1989) Transcript analyses of the uvsX-40-41 region of bacteriophage T4. Changes in the RNA as infection proceeds. *J. Biol. Chem.*, **264**, 14432–14439.
50. Marshall,P., Sharma,M. and Hinton,D.M. (1999) The bacteriophage T4 transcriptional activator MotA accepts various base-pair changes within its binding sequence. *J. Mol. Biol.*, **285**, 931–944.
51. Gruber,T.M. and Gross,C.A. (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.*, **57**, 441–466.
52. Minakhin,L., Camarero,J.A., Holford,M., Parker,C., Muir,T.W. and Severinov,K. (2001) Mapping the molecular interface between the sigma(70) subunit of E. coli RNA polymerase and T4 AsiA. *J. Mol. Biol.*, **306**, 631–642.
53. Gregory,B.D., Nickels,B.E., Garrity,S.J., Severinova,E., Minakhin,L., Urbauer,R.J., Urbauer,J.L., Heyduk,T., Severinov,K. and Hochschild,A. (2004) A regulator that inhibits transcription by targeting an intersubunit interaction of the RNA polymerase holoenzyme. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 4554–4559.
54. Hinton,D.M., March-Amegadzie,R., Gerber,J.S. and Sharma,M. (1996) Characterization of pre-transcription complexes made at a bacteriophage T4 middle promoter: involvement of the T4 MotA activator and the T4 AsiA protein, a sigma 70 binding protein, in the formation of the open complex. *J. Mol. Biol.*, **256**, 235–248.
55. Ross,W., Gosink,K.K., Salomon,J., Igarashi,K., Zou,C., Ishihama,A., Severinov,K. and Gourse,R.L. (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science*, **262**, 1407–1413.
56. Campbell,E.A., Muzzin,O., Chlenov,M., Sun,J.L., Olson,C.A., Weinman,O., Trester-Zedlitz,M.L. and Darst,S.A. (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell*, **9**, 527–539.
57. Colland,F., Fujita,N., Kotlarz,D., Bown,J.A., Meares,C.F., Ishihama,A. and Kolb,A. (1999) Positioning of sigma(S), the stationary phase sigma factor, in Escherichia coli RNA polymerase-promoter open complexes. *EMBO J.*, **18**, 4049–4059.
58. Marr,M.T., Roberts,J.W., Brown,S.E., Klee,M. and Gussin,G.N. (2004) Interactions among CII protein, RNA polymerase and the lambda PRE promoter: contacts between RNA polymerase and the -35 region of PRE are identical in the presence and absence of CII protein. *Nucleic Acids Res.*, **32**, 1083–1090.
59. Kumar,A. and Moran,C.P. Jr (2008) Promoter activation by repositioning of RNA polymerase. *J. Bacteriol.*, **190**, 3110–3117.
60. Bown,J.A., Owens,J.T., Meares,C.F., Fujita,N., Ishihama,A., Busby,S.J. and Minchin,S.D. (1999) Organization of open complexes at Escherichia coli promoters. Location of promoter DNA sites close to region 2.5 of the sigma70 subunit of RNA polymerase. *J. Biol. Chem.*, **274**, 2263–2270.
61. Owens,J.T., Chmura,A.J., Murakami,K., Fujita,N., Ishihama,A. and Meares,C.F. (1998) Mapping the promoter DNA sites proximal to conserved regions of sigma 70 in an Escherichia coli RNA polymerase-lacUV5 open promoter complex. *Biochemistry*, **37**, 7670–7675.
62. Marr,M.T., Datwyler,S.A., Meares,C.F. and Roberts,J.W. (2001) Restructuring of an RNA polymerase holoenzyme elongation complex by lambdoid phage Q proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 8972–8978.
63. Pai,K.S., Bussiere,D.E., Wang,F., White,S.W. and Bastia,D. (1996) Structure of the replication terminus-terminator protein complex as probed by affinity cleavage. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 10647–10652.
64. Tao,T., Lamkin,M. and Scheiner,C.J. (1985) The conformation of the C-terminal region of actin: a site-specific photocrosslinking study using benzophenone-4-maleimide. *Arch. Biochem. Biophys.*, **240**, 627–634.
65. Dorman,G. and Prestwich,G.D. (1994) Benzophenone photophores in biochemistry. *Biochemistry*, **33**, 5661–5673.
66. Giron-Monzon,L., Manelyte,L., Ahrends,R., Kirsch,D., Spengler,B. and Friedhoff,P. (2004) Mapping protein-protein interactions between MutL and MutH by cross-linking. *J. Biol. Chem.*, **279**, 49338–49345.
67. Kawamura,A., Hindi,S., Mihai,D.M., James,L. and Aminova,O. (2008) Binding is not enough: flexibility is needed for photocrosslinking of Lck kinase by benzophenone photoligands. *Bioorg. Med. Chem.*, **16**, 8824–8829.

68. Gilmore,J.M., Bieber Urbauer,R.J., Minakhin,L., Akoyev,V., Zolkiewski,M., Severinov,K. and Urbauer,J.L. (2010) Determinants of affinity and activity of the anti-sigma factor AsiA. *Biochemistry*, **49**, 6143–6154.
69. Pulitzer,J.F., Coppo,A. and Caruso,M. (1979) Host–virus interactions in the control of T4 prereplicative transcription. II. Interaction between tabC (rho) mutants and T4 mot mutants. *J. Mol. Biol.*, **135**, 979–997.
70. Vassylyev,D.G., Vassylyeva,M.N., Perederina,A., Tahirov,T.H. and Artsimovitch,I. (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*, **448**, 157–162.
71. Nickels,B.E., Garrity,S.J., Mekler,V., Minakhin,L., Severinov,K., Ebright,R.H. and Hochschild,A. (2005) The interaction between sigma70 and the beta-flap of Escherichia coli RNA polymerase inhibits extension of nascent RNA during early elongation. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 4488–4493.
72. Hsu,L.M. (2009) Monitoring abortive initiation. *Methods*, **47**, 25–36.
73. Orsini,G., Kolb,A. and Buc,H. (2001) The Escherichia coli RNA polymerase.anti-sigma 70 AsiA complex utilizes alpha-carboxyl-terminal domain upstream promoter contacts to transcribe from a -10/-35 promoter. *J. Biol. Chem.*, **276**, 19812–19819.
74. Geiduschek,E.P. (2010) An introduction to transcription and gene regulation. *J. Biol. Chem.*, **285**, 25885–25892.
75. Sharma,M., Marshall,P. and Hinton,D.M. (1999) Binding of the bacteriophage T4 transcriptional activator, MotA, to T4 middle promoter DNA: evidence for both major and minor groove contacts. *J. Mol. Biol.*, **290**, 905–915.