Phage T7 Gp2 inhibition of *Escherichia coli* RNA polymerase involves misappropriation of σ^{70} domain 1.1

Brian Bae^a, Elizabeth Davis^a, Daniel Brown^b, Elizabeth A. Campbell^a, Sivaramesh Wigneshweraraj^b, and Seth A. Darst^{a,1}

^aLaboratory of Molecular Biophysics, The Rockefeller University, New York, NY 10065; and ^bMedical Research Council Centre for Molecular Bacteriology and Infection, Faculty of Medicine, Imperial College London, London SW7 2AZ, United Kingdom

Edited by E. Peter Geiduschek, University of California, San Diego, La Jolla, CA, and approved October 15, 2013 (received for review August 1, 2013)

Bacteriophage T7 encodes an essential inhibitor of the *Escherichia coli* host RNA polymerase (RNAP), the product of gene 2 (Gp2). We determined a series of X-ray crystal structures of *E. coli* RNAP holoenzyme with or without Gp2. The results define the structure and location of the RNAP σ^{70} subunit domain 1.1 ($\sigma_{1.1}^{70}$) inside the RNAP active site channel, where it must be displaced by the DNA upon formation of the open promoter complex. The structures and associated data, combined with previous results, allow for a complete delineation of the mechanism for Gp2 inhibition of *E. coli* RNAP. In the primary inhibition mechanism, Gp2 forms a protein–protein interaction with $\sigma_{1.1}^{70}$, preventing the normal egress of $\sigma_{1.1}^{70}$ from the RNAP active site channel. Gp2 thus misappropriates a domain of the RNAP holoenzyme, $\sigma_{1.1}^{70}$, to inhibit the function of the enzyme.

bacteriophage T7 Gp2 | X-ray crystallography | transcription

After infection of *Escherichia coli* by bacteriophage T7 (1), the first T7 genomic DNA to enter the bacterial cell contains four strong, early promoters transcribed by the host RNA polymerase (RNAP) σ^{70} -holoenzyme (E- σ^{70}). An early transcription product is T7 gene 1, coding for the single-subunit T7 RNAP that transcribes the middle and late phage genes. There appears to be no function for *E. coli* RNAP in middle and late infection, and indeed T7 encodes an essential inhibitor of *E. coli* RNAP, the 7.2-kDa product of gene 2 (Gp2). The essential role of Gp2 in T7 infection is to prevent unregulated transcription of the phage genome by the host RNAP in late stages of infection, which interferes with transcription of the phage genome by the speedier T7 RNAP (2). The requirement to shut off *E. coli* RNAP transcription is stringent. Gp2 is accordingly a potent inhibitor of E- σ^{70} transcription initiation (3, 4).

Gp2 binds to E- σ^{70} with sub-nanomolar K_d (5). The major binding determinant for Gp2 on the RNAP is the β' -jaw domain (Fig. 1) (4, 5). NMR structures of Gp2 and a Gp2- β' -jaw subcomplex, as well as molecular models of the Gp2-E- σ^{70} complex, indicate that Gp2 occupies a part of the RNAP downstream duplex DNA channel (5–7). Gp2 blocks the formation of the open promoter complex (RPo), but preformed RPo and RNAP elongation complexes (where the downstream duplex DNA already occupies the channel) are resistant to Gp2 (4, 6), suggesting a straightforward mechanism whereby Gp2 binding or downstream duplex DNA binding in the RNAP downstream channel are mutually exclusive. However, this mechanism does not explain observations that RNAP harboring an N-terminal truncation of domain 1.1 of σ^{70} ($\Delta 1.1\sigma^{70}$; Fig. 14) is very inefficiently inhibited by Gp2 (5, 7), indicating a more complex inhibition mechanism for Gp2.

Group I (primary) σ s, such as *E. coli* σ^{70} , are autoregulated by N-terminal $\sigma_{1.1}$ (Fig. 14), which serves to prevent σ^{70} interactions with promoter DNA in the absence of RNAP (8–10). $\sigma_{1.1}$ also plays a role in the formation of RPo (11, 12). Based on solution FRET studies, $\sigma_{1.1}$ is located within the RNAP active center channel in E- σ^{70} , but in RPo, $\sigma_{1.1}$ is displaced outside the

channel by the entering DNA (13). A 37-residue linker connects $\sigma_{1.1}^{70}$ to the rest of σ^{70} , facilitating the large movements of $\sigma_{1.1}^{70}$ during RPo formation. To better understand the role of $\sigma_{1.1}^{70}$ in Gp2 inhibition of RPo formation, we determined X-ray crystal structures of *E. coli* RNAP holoenzyme (14) harboring full-length or $\Delta 1.1\sigma^{70}$ and with or without bound Gp2. The results define the structure of *E. coli* $\sigma_{1.1}^{70}$, as well as a part of the linker segment, confirm the location of $\sigma_{1.1}^{70}$ in the RNAP downstream duplex DNA channel in the holoenzyme, and demonstrate that Gp2 uses stable, specific protein–protein interactions with both the β' -jaw as well as with $\sigma_{1.1}^{70}$ to hold $\sigma_{1.1}^{70}$ in the RNAP downstream duplex DNA channel, thereby blocking entry of the promoter DNA.

Results and Discussion

We determined a series of structures of *E. coli* RNAP holoenzyme (the target of Gp2) with and without bound Gp2. A total of four independent structures were refined [E- $\Delta 1.1\sigma^{70}$, 3.6-Å resolution; E-SelenoMethionine(SeMet) σ^{70} , 3.9-Å resolution; Gp2-E- $\Delta 1.1\sigma^{70}$, 3.9-Å resolution; and Gp2-E- σ^{70} , 3.8-Å resolution], and diffraction datasets from an additional crystal were also analyzed (Gp2-E-SeMet σ^{70} , 6.5-Å resolution). The highest resolution data were used for the refinement of E- $\Delta 1.1\sigma^{70}$ (R/R_{free} 0.246/0.288; rmsd bonds/angles 0.005 Å/1.187°; Ramachandran allowed/outliers 91.6%/0.7%). This model was then used as the starting point for refinement of the other models to equal or better parameters (Table S1 and Fig. S1).

Significance

After infection of *Escherichia coli* by bacteriophage T7, the host RNA polymerase (RNAP) produces early phage transcription products that encode the phages own RNAP (that transcribes subsequent phage genes) as well as Gp2, an essential inhibitor of the host RNAP. X-ray crystal structures of *E. coli* RNAP define the structure and location of the RNAP σ^{70} subunit domain 1.1 ($\sigma_{1.1}^{70}$) inside the RNAP active site channel, where it must be displaced by the DNA upon formation of the transcription complex. Gp2 binds inside the RNAP active site channel and also interacts with $\sigma_{1.1}^{70}$, preventing $\sigma_{1.1}^{70}$ from exiting the RNAP active site channel. Gp2 thus misappropriates a domain of the RNAP, $\sigma_{1.1}^{70}$, to inhibit the function of the enzyme.

Author contributions: B.B., S.W., and S.A.D. designed research; B.B., E.D., D.B., and E.A.C. performed research; B.B., D.B., S.W., and S.A.D. analyzed data; and B.B., E.A.C., S.W., and S.A.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes: 4LJZ (E- $\Delta 1.1\sigma^{70}$), 4LK1 (E-SeMet σ^{70}), 4LKO (Gp2-E- $\Delta 1.1\sigma^{70}$), and 4LLG (Gp2-E- σ^{70})].

See Commentary on page 19662.

¹To whom correspondence should be addressed. E-mail: darst@rockefeller.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1314576110/-DCSupplemental.

Fig. 1. Structure of *E. coli* E- σ^{70} . (A) Structural architecture of *E. coli* σ^{70} . The top bar represents the primary sequence of *E. coli* σ^{70} (residues 1–613). Every 100 residues are labeled on the bottom and marked with a thin white vertical line. The domain architecture is denoted by the thickness of the horizontal bar; thick regions represent structural domains ($\sigma_{1,1}^{70}$, σ_2^{70} , σ_3^{70} , and σ_4^{70}) and thin regions represent flexible loop regions (16). Hollow white regions denote disordered loops. The structural core of $\sigma_{1,1}^{70}$ and the linker connecting $\sigma_{1,1}^{70}$ with σ_2^{70} are colored dark orange. The rest of σ^{70} is colored light orange except for the nonconserved region σ^{70}_{NCR} (NCR inserted between conserved sequence regions 1.2 and 2.1) is colored light yellow (17, 29). The expanded view below denotes the secondary structure of $\sigma_{1,1}^{70}$ and the $\sigma_{1,1-12}^{70}$ linker (the thin line represents loop regions with dashed lines representing disordered segments not modeled; rectangles represent α -helices numbered α_1 – α_3 in the $\sigma_{1,1}^{70}$ structural core). The modeled portion of $\sigma_{1,1}^{70}$ is colored as a ramp from the N terminus (blue, residue 6) to the C terminus (red, residue 83). (B–D) Structural views of *E. coli* E- σ^{70} . The RNAP is shown as a molecular surface except $\sigma_{1,1}^{70}$ is shown as a backbone ribbon and color-ramped according to the lower part of A. Disordered connecting segments are denoted by orange spheres. The nomenclature for the β - and β' -lineage-specific inserts ($\beta i4$, $\beta i9$, and $\beta' i6$) is as in ref. 37. (B) View into the active site channel (channel view) showing $\sigma_{1,1}^{70}$ in the channel. (C) View down through the β -subunit (β -side view), with β removed to reveal the active site (Mg^{2+} ion, yellow sphere) and nucleic acid binding channels. Superimposed is the DNA from an initiation complex structure (downstream portion of the transcription bubble, and 10-bp of downstream duplex DNA; PDB ID code 4G7H) (38), shown as a gray cartoon (light gray, template strand; dark gray, nontemplate strand). $\sigma_{1,1}^{70}$ occupies the downstream duplex DNA binding channel with its center of gravity at approximately +8. (D) $\sigma_{1,1}^{70}$ is nestled in the RNAP channel, interacting with elements of the β_2 domain (β -residues 165–166, 197, and 202–203), the clamp (β' -residues 120, 132–133, and the rudder around residue 311), and other elements of the β' -pincer (β' -residues 1310–1311). The three methionine residues identified within $\sigma_{1,1}^{70}$ from selenomethionyl anomalous peaks (Fig. S2A) are shown (M47, M51, and M56).

Structure of E- σ^{70} . The structure of E- σ^{70} (containing full-length σ^{70}) reveals $\sigma_{1,1}$ in the context of the RNAP holoenzyme (Fig. 1 and Fig. S2). Whereas difference Fourier analyses revealed strong difference density corresponding to $\sigma_{1,1}^{70}$ (Fig. S2A and B), the electron density lacked features for most side chains, consistent with the dynamic nature of $\sigma_{1,1}^{70}$ during transcription initiation (13). $\sigma_{1,1}^{70}$ was modeled primarily as a poly-Ala backbone but with selected side chains as the density features allowed (i.e., side chains that, for the most part, interact with the RNAP). The sequence register was confirmed with the aid of the NMR structure of the structurally homologous *Thermatoga maritima* $\sigma_{1,1}^A$ [Protein Data Bank (PDB) ID code 2K6X] (10). In addition, anomalous Fourier difference maps revealed 44 of a possible 50 SeMet sites in the two copies of SeMet- σ^{70} in the crystallographic asymmetric unit, including six peaks attributed to methionines at positions 47, 51, and 56 in the two crystallographically independent copies of $\sigma_{1,1}^{70}$ (Fig. S2 and Table S2). Although the overall

fold of much of our $\sigma_{1,1}^{70}$ model is roughly consistent with a recently published structure of E- σ^{70} (PDB ID code 4IGC) (15), the folds deviate after the first methionine at position 47; the positions of methionines 51 and 56 in 4IGC are not consistent with anomalous Fourier peaks attributed to methionines 51 and 56 in our structure (Fig. S2A and B).

$\sigma_{1,1}^{70}$ comprises a core folded domain (residues 1–56) of three α -helices with similar topology to *Tma* $\sigma_{1,1}^A$ (Fig. 1 and Fig. S2) (10). $\sigma_{1,1}^{70}$ is linked to the N-terminal residue of the σ_2^{70} domain (residue 94 at the beginning of σ^{70} conserved region 1.2) (16, 17) by a 37-residue linker (residues 57–93) predicted to be mostly unstructured but with a weakly predicted α -helix (residues 74–80; Fig. S3) (18). The electron density maps contain difference density that is not accounted for by the $\sigma_{1,1}^{70}$ core (Fig. S2A), including an α -helical segment that we attribute to the linker helix (18). We could not confidently assign the sequence register of the linker segment, but based on correspondence between the observed and

predicted α -helix and the location of the α -helical density, we attribute the linker density to roughly residues 64–83. The linker segment wraps around the $\sigma_{1.1}^{70}$ core domain, with the linker helix interacting in the groove between α -helices 1 and 3 of the core domain to form a four-helix bundle (Fig. 1).

The resulting globular structural unit formed by $\sigma_{1.1}^{70}$ and the linker sits directly in the path of the downstream duplex DNA in the RNAP active site channel, with its center of gravity ~ 45 Å from the RNAP active site Mg^{2+} ion, corresponding to the location of base-pair +8, very close to the position (+9) proposed by FRET measurements (Fig. 1C) (13). The $\sigma_{1.1}^{70}$ /linker structural unit contacts elements of both the β -pincer (β_2 domain) and the β' -pincer from the inside of the RNAP active site channel (Fig. 1D).

Structure of Gp2–RNAP Complexes. In all four crystallographically independent complexes from the Gp2–E- $\Delta 1.1\sigma^{70}$ and Gp2–E- σ^{70} structures, the Gp2–RNAP contacts are essentially identical (Fig. 2 and Fig. S4). As expected, Gp2 interacts primarily with the β' -jaw as well as with some residues linking the β' -jaw to the rest of the RNAP, burying an accessible surface area of about 2,100 Å² with β' . The Gp2– β' interface includes five potential salt bridges (Gp2– β' : E28–R1149, D37–R1174, E44–K1170, R56–E1188, and R58–E1158; Fig. 2B). The charged residues participating in these salt bridges are all well conserved among Gp2 homologs and their

Fig. 2. Structure of Gp2–E- σ^{70} . (A) Overall structure, showing the channel view (same view as Fig. 1B). The RNAP is shown as a molecular surface except the β' -jaw (magenta) and $\sigma_{1.1}^{70}$ (orange) are shown as backbone ribbons with transparent molecular surfaces, as is Gp2 (green) sandwiched between them. Gp2 forms significant protein–protein interfaces with both β' -jaw and $\sigma_{1.1}^{70}$, serving as a bridge between them. (B) View of the $\sigma_{1.1}^{70}$ –Gp2– β' -jaw protein–protein–protein interaction, viewed from the RNAP active site outwards. The proteins are shown as backbone worms, with selected interacting side chains highlighted (also see Fig. S4). Note the preponderance of hydrophobic interactions in the $\sigma_{1.1}^{70}$ –Gp2 interface, and the five salt bridges characterizing the relatively polar Gp2– β' -jaw interface. Also shown is a segment of the β_2 domain that interacts with Gp2.

Fig. 3. Gp2 reorients $\sigma_{1.1}^{70}$ within the RNAP active site channel. (A) Cutaway views of the RNAP active site channel (similar to the view of Fig. 1C). (Left) E- σ^{70} , but with the position of Gp2 (from the Gp2–E- σ^{70} structure) outlined (green), illustrating the steric clash between $\sigma_{1.1}^{70}$ and Gp2. (Right) Gp2–E- σ^{70} , comparison with E- σ^{70} reveals the reorientation of the $\sigma_{1.1}^{70}$ structural core and the disordered $\sigma_{1.1-1.2}^{70}$ linker. (B) The $\sigma_{1.1}^{70}$ – $\sigma_{1.1-1.2}^{70}$ linker from E- σ^{70} and Gp2– σ^{70} were superimposed by the $\sigma_{1.1}^{70}$ structural core (orange), revealing that the $\sigma_{1.1-1.2}^{70}$ linker helix (magenta) and Gp2 (green) interact with the same hydrophobic surface of the $\sigma_{1.1}^{70}$ structural core between α -helices α_1 and α_3 . (C) Gp2[F27BpA] cross-links to $\sigma_{1.1}^{70}$. (Upper) Silver-stained SDS/PAGE gel showing the migration positions of β , β' , σ^{70} , $\Delta 1.1\sigma^{70}$, and Gp2[F27BpA] before (lanes 1, 3, 5, and 7) and after (lanes 2, 4, 6, and 8) UV-cross-linking. The reaction components in each lane are indicated at the top of C. (Lower) Western blot of the gel shown (Upper) with anti-Gp2 polyclonal antibodies. The Gp2[F27BpA]- σ^{70} complex is indicated. The expected position of the absent Gp2– $\Delta 1.1\sigma^{70}$ cross-linking product is also shown.

cognate RNAPs (Fig. S4) (19). The Gp2– β' R56–E1188 and R58–E1158 interactions are critical for Gp2 binding to, and inhibition of, RNAP (4, 6).

Gp2 also contacts residues from the β_2 domain of the β -pincer (Fig. 2 and Fig. S4). Thus, Gp2 bridges the β' - and β -pincers from inside the RNAP active site channel, explaining how the presence of Gp2 favors the closed clamp state of RNAP in solution (20).

Gp2 binds to the β' -jaw and thereby occupies the downstream duplex DNA-binding channel of the RNAP (Fig. S5), but Gp2 inhibits E- $\Delta 1.1\sigma^{70}$ only very inefficiently (5, 7). In the presence of Gp2, $\sigma_{1.1}^{70}$ in the RNAP active site channel must move to accommodate Gp2 due to a steric clash (Fig. 3A). The Gp2–E- σ^{70} structure reveals that $\sigma_{1.1}^{70}$ not only moves from its normal position in E- σ^{70} (Fig. 1), but reorients and forms a significant protein–protein interface (buried surface area of $\sim 1,400$ Å²) with Gp2 that is primarily hydrophobic in nature (Figs. 2 and 3A and Fig. S4). The reorientation of the $\sigma_{1.1}^{70}$ structural core, which involves a rotation of 115° and a small translation of about 4.5 Å, was confirmed by anomalous Fourier difference peaks marking the three selenomethionines in $\sigma_{1.1}^{70}$, calculated using data from Gp2–E-SeMet- $\sigma_{1.1}^{70}$ crystals (Fig. S6 and Table S1). The $\sigma_{1.1}^{70}$ structural core maintains the same structural fold in E- σ^{70} with or

without Gp2. In the presence of Gp2, the density maps for $\sigma_{1.1}^{70}$ were somewhat better, and more side chain information was included in the model, consistent with reduced mobility due to binding to Gp2.

Gp2 interacts with the same $\sigma_{1.1}^{70}$ surface as the $\sigma_{1.1-1.2}^{70}$ linker helix (Fig. 3B). The linker helix and, in fact, the whole linker segment is thus entirely displaced from around the $\sigma_{1.1}^{70}$ structural core, is not accounted for by any electron density, and is presumed disordered.

Gp2 Cross-Linking and Surface Conservation. To confirm the Gp2– $\sigma_{1.1}^{70}$ protein–protein interaction, we site-specifically incorporated the nonnatural, UV-cross-linkable amino acid p-benzoyl-L-phenylalanine (BpA) (21) at position F27 of Gp2, a residue that lies near the edge of, and participates in, the $\sigma_{1.1}^{70}$ protein–protein interaction (Fig. 2B and Fig. S4) but is not highly conserved so

is unlikely to be essential for the interaction (19). Moreover, phenylalanine replacement by BpA is a somewhat conservative substitution. The mutant Gp2 (Gp2[F27BpA]) was active in inhibiting E- σ^{70} in vitro (Fig. S7). Upon excitation with UV light, Gp2[F27BpA] in a Gp2[F27BpA]E- σ^{70} complex efficiently formed a specific covalent cross-link with σ^{70} , but no detectable cross-link was formed in a Gp2[F27BpA]–E- $\Delta 1.1\sigma^{70}$ with $\Delta 1.1\sigma^{70}$ (Fig. 3C), indicating that the cross-link from Gp2[F27BpA] to σ^{70} occurred within the $\sigma_{1.1}^{70}$ domain. The strict requirements for efficient BpA cross-linking—the C ζ atom of the substituted phenylalanine residue must be within ~ 2 – 4 Å of the carbon atom of a C–H bond (22)—make the observation of an adventitious cross-link due to close proximity of the two protein domains within the RNAP cleft unlikely. These results confirm a stable, specific protein–protein interaction between Gp2 and $\sigma_{1.1}^{70}$.

Fig. 4. Gp2 misappropriates $\sigma_{1.1}^{70}$ for E- σ^{70} inhibition. (A) The interaction surfaces on Gp2 for the β' -jaw and for $\sigma_{1.1}^{70}$ are conserved among Gp2 homologs. Shown are two views of the β' -jaw-Gp2- $\sigma_{1.1}^{70}$ interacting protein triplet. Gp2 is shown as a molecular surface color-coded according to conservation (39) within an alignment of 25 Gp2 homologs (19) as shown on the color scale below. The β' -jaw (magenta) and the $\sigma_{1.1}^{70}$ structural core (gray) are shown as backbone ribbons. The interaction surfaces of the Gp2 binding partners are outlined in black on the Gp2 surface. (B) Schematic illustrating the mechanism for Gp2 inhibition of E- σ^{70} . In E- σ^{70} (Upper Left), $\sigma_{1.1}^{70}$ occupies the RNAP downstream duplex DNA channel. Formation of RPo with promoter DNA causes the ejection of $\sigma_{1.1}^{70}$ from the channel (Right). In Gp2-E- σ^{70} (Lower Left), Gp2 (green) forms a protein bridge between β' and $\sigma_{1.1}^{70}$, preventing the egress of $\sigma_{1.1}^{70}$ from the channel and thereby blocking the entry of the promoter DNA.

The Gp2's reason for being thus appears to be to anchor $\sigma^{70}_{1.1}$ within the RNAP active site cleft by forming bridging protein–protein interactions between the β' -jaw and $\sigma^{70}_{1.1}$ (Figs. 24 and 34). Support for this supposition comes from mapping conserved residues on the surface of Gp2, which reveals two distinct patches of conserved surface. One patch matches the Gp2– β' -jaw interaction interface, whereas the other patch, on the opposite face of Gp2, matches the Gp2– $\sigma^{70}_{1.1}$ interaction interface (Fig. 44).

Conclusion

Our results, combined with available functional and genetic data (4–7), allow a complete delineation of the mechanism for Gp2 inhibition of *E. coli* E- σ^{70} (Fig. 4B). By binding to the β' -jaw inside the RNAP active site channel, Gp2 sterically (Fig. S5) and electrostatically (6, 7) interferes with the proper positioning of DNA in the RNAP downstream duplex DNA channel. However, in the absence of $\sigma^{70}_{1.1}$, promoter DNA can effectively compete with Gp2, so Gp2 inhibition is compromised (5–7). In the primary mechanism for inhibition, Gp2 forms a stable, specific protein–protein interaction with $\sigma^{70}_{1.1}$ (Figs. 2 and 3), preventing its normal egress from the RNAP active site channel (Fig. 4B). In this way, Gp2 misappropriates (23) a domain of the RNAP holoenzyme, $\sigma^{70}_{1.1}$, to inhibit the function of the holoenzyme.

Materials and Methods

Full details of the methods used are presented in the *SI Materials and Methods*.

Expression and Purification of *E. coli* RNAP, σ^{70} , $\Delta 1.1\sigma^{70}$, and Gp2. *E. coli* $\Delta\alpha$ -CTD-RNAP. For all of the structural experiments, we used *E. coli* core RNAP lacking the α CTD, prepared as described previously (24). The purified protein was buffer exchanged into storage buffer [10 mM Tris, pH 8.0, 0.3 M NaCl, 0.1 mM EDTA, 15% (vol/vol) glycerol, 5 mM DTT] and concentrated to 10 mg/mL by centrifugal filtration (VivaScience), and stored at -80°C until use.

***E. coli* σ^{70} and $\Delta 1.1\sigma^{70}$.** Full-length and $\Delta 1.1$ *E. coli* σ^{70} were both expressed from pET21a-based expression vectors encoding an N-terminal His₆-tag followed by a PreScission protease (GE Healthcare) cleavage site. The proteins were expressed using standard methods and purified by Ni²⁺-affinity chromatography, protease cleavage to remove the His₆-tag, anion exchange chromatography, and gel filtration chromatography. Purified σ^{70} or $\Delta 1.1\sigma^{70}$ was concentrated to 10 mg/mL by centrifugal filtration (VivaScience) and stored at -80°C . SeMet σ^{70} was produced by suppression of endogenous methionine biosynthesis (25) and purified as described for native σ^{70} . Purified SeMet σ^{70} was concentrated to 5 mg/mL and stored at -80°C .

Gp2. T7 Gp2 was expressed and purified as described (4). Purified T7 Gp2 (in TGED + 0.5 M NaCl) was concentrated to 10 mg/mL by centrifugal filtration (VivaScience) and stored at -80°C .

Crystallization of Complexes (E- $\Delta 1.1\sigma^{70}$, E-SeMet σ^{70} , Gp2-E- $\Delta 1.1\sigma^{70}$, Gp2-E- σ^{70} , and Gp2-E-SeMet σ^{70}). Before crystallization trials, aliquots of the purified components were thawed on ice and buffer-exchanged into crystallization buffer (20 mM Tris-HCl, pH 8.0, 0.2 M NaCl, 5 mM DTT). E- σ^{70} was formed by adding 1.2-fold molar excess of σ^{70} (or $\Delta 1.1\sigma^{70}$ for E- $\Delta 1.1\sigma^{70}$ or SeMet σ^{70} for E-SeMet σ^{70}) to the $\Delta\alpha$ CTD-core RNAP and incubated at room temperature for 15 min. Gp2–holoenzyme complexes were formed by adding fivefold molar excess of Gp2 to the holoenzyme and incubating at room temperature for 15 min. Before crystallization, the final concentration of complex was adjusted to 40 μM . The crystals of holoenzyme were grown via vapor

diffusion at 22°C by mixing 1 μL of sample with 1 μL of reservoir solution (0.1 M MES, pH 6, 0.1 M calcium acetate, 12–15% (vol/vol) PEG 400, 5 mM DTT) in a 48-well hanging drop tray (Hampton Research). Typically, the crystals took about 5 d to grow to a size of about 200 μm in the longest dimension. For cryoprotection, crystals were transferred into reservoir solution + 15% (vol/vol) ethylene glycol in two steps.

Data Collection, Refinement, and Model Building (E- $\Delta 1.1\sigma^{70}$, E-SeMet σ^{70} , Gp2-E- $\Delta 1.1\sigma^{70}$, Gp2-E- σ^{70} , and Gp2-E-SeMet σ^{70}). X-ray diffraction data were collected at Brookhaven National Laboratory National Synchrotron Light Source (NSLS) beamline X29 and at Argonne National Laboratory Advanced Photon Source (APS) beamlines 24ID-C and 24ID-E. Data were integrated and scaled using HKL2000 (26). The initial electron density maps were calculated by molecular replacement using Phaser (27) from a starting model of *E. coli* core RNAP (PDB ID code 3LU0) (24, 28) combined with available crystal structures of *E. coli* σ^{70} domains [σ^{70}_2 , PDB ID code 1S1G (25, 29); σ^{70}_4 , PDB ID code 2P7V (4, 30)] and a homology model for the rest of σ^{70} . Two molecules of holoenzyme were clearly identified in the asymmetric unit. The model was first refined with rigid body refinement of each RNAP molecule and subsequently of 20 individual mobile domains using Phenix (26, 31). The model was improved by iterative cycles of manual building with Coot (27, 32) and refinement with Phenix, combined with deformable elastic network refinement with noncrystallographic symmetry restraints using CNS 1.3 performed on the Structural Biology Grid cluster (33–35). The final E- $\Delta 1.1\sigma^{70}$ structure (3.6-Å resolution) was used as a starting model for refinement of E-SeMet σ^{70} , Gp2-E- $\Delta 1.1\sigma^{70}$, and Gp2-E- σ^{70} , which were refined using similar methods (Table S1 and Fig. S1).

Gp2[F27BpA] Cross-Linking. To introduce the photo-cross-linkable amino acid BpA into Gp2 at position F27, we introduced an amber stop codon TAG at position 27 using standard oligonucleotide-directed mutagenesis using plasmid pSW33:Gp2 (6), creating pSW33:Gp2(F27-TAG). The mutant Gp2 containing BpA at position F27 (Gp2[F27BpA]) was overexpressed and purified from J1 (DE3) *E. coli* strain (lacking the β' -jaw domain) as previously described (6, 36) by Ni²⁺-affinity chromatography. The ability of the Gp2 [F27BpA] to inhibit transcription by E- σ^{70} was confirmed using an in vitro transcription assay as previously described (6) (Fig. S7). Twenty-microliter cross-linking reactions were performed in cross-linking buffer (10 mM Tris-Cl, pH 8.0, 150 mM NaCl, 10% (vol/vol) glycerol, 0.1 mM EDTA, and 1 mM DTT) and comprised (as indicated in Fig. S7) 100 nM Gp2[F27BpA], 50 nM core RNAP, 200 nM σ^{70} , and 200 nM $\Delta 1.1\sigma^{70}$. The reaction was incubated at 20°C for 5 min to allow complex formation before cross-linking was carried out at a wavelength of 365 nm in a CL-1000 UV cross-linker (UVP). The reaction was stopped after 10 min by the addition of 2 $\times$ SDS/PAGE loading buffer (Sigma) and heated at 100°C for 5 min. Samples were separated on a 15%/10% (wt/vol) split-gradient SDS/PAGE gel and proteins were visualized by silver staining. Gp2[F27BpA]-containing complexes were visualized by Western blotting using an anti-Gp2 polyclonal antibody (Eurogentec).

ACKNOWLEDGMENTS. We thank K. R. Rajashankar and Igor Kourinov [APS Northeastern Collaborative Access Team (NE-CAT) beamlines] and W. Shi (NSLS beamline X29) for support with synchrotron data collection. This work was based, in part, on research conducted at the APS and NSLS and supported by the US Department of Energy, Office of Basic Energy Sciences. The NE-CAT beamlines at the APS are supported by Award RR-15301 from the National Center for Research Resources at the National Institutes of Health (NIH). B.B. was supported by a Merck Postdoctoral Fellowship (The Rockefeller University) and a National Research Service Award (NIH F32 GM103170). The work in the laboratory of S.W. was supported by Biotechnology and Biological Sciences Research Council Project Grant BB/K000233/1.

1. Molineux IJ (2005) *The Bacteriophages*, ed Calendar R (Oxford Univ Press, New York), pp 277–301.
2. Savalia D, Robins W, Nechaev S, Molineux I, Severinov K (2010) The role of the T7 Gp2 inhibitor of host RNA polymerase in phage development. *J Mol Biol* 402(1): 118–126.
3. Hesselbach BA, Nakada D (1975) Inactive complex formation between *E. coli* RNA polymerase and inhibitor protein purified from T7 phage infected cells. *Nature* 258(5533):354–357.
4. Nechaev S, Severinov K (1999) Inhibition of *Escherichia coli* RNA polymerase by bacteriophage T7 gene 2 protein. *J Mol Biol* 289(4):815–826.
5. James E, et al. (2012) Structural and mechanistic basis for the inhibition of *Escherichia coli* RNA polymerase by T7 Gp2. *Mol Cell* 47(5):755–766.

6. Cámara B, et al. (2010) T7 phage protein Gp2 inhibits the *Escherichia coli* RNA polymerase by antagonizing stable DNA strand separation near the transcription start site. *Proc Natl Acad Sci USA* 107(5):2247–2252.
7. Sheppard C, et al. (2011) Inhibition of *Escherichia coli* RNAP by T7 Gp2 protein: Role of negatively charged strip of amino acid residues in Gp2. *J Mol Biol* 407(5):623–632.
8. Camarero JA, et al. (2002) Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. *Proc Natl Acad Sci USA* 99(13):8536–8541.
9. Dombroski AJ, Walter WA, Record MT, Jr., Siegle DA, Gross CA (1992) Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell* 70(3):501–512.
10. Schwartz EC, et al. (2008) A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem Biol* 15(10):1091–1103.

11. Wilson C, Dombroski AJ (1997) Region 1 of sigma70 is required for efficient isomerization and initiation of transcription by Escherichia coli RNA polymerase. *J Mol Biol* 267(1):60–74.
12. Vuthoori S, Bowers CW, McCracken A, Dombroski AJ, Hinton DM (2001) Domain 1.1 of the σ (70) subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes. *J Mol Biol* 309(3):561–572.
13. Mekler V, et al. (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* 108(5):599–614.
14. Zuo Y, Wang Y, Steitz TA (2013) The mechanism of E. coli RNA polymerase regulation by ppGpp is suggested by the structure of their complex. *Mol Cell* 50(3):430–436.
15. Murakami KS (2013) The X-ray crystal structure of Escherichia coli RNA polymerase sigma70 holoenzyme. *J Biol Chem* 288(13):9126–9134.
16. Campbell EA, et al. (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell* 9(3):527–539.
17. Lonetto M, Gribskov M, Gross CA (1992) The sigma 70 family: Sequence conservation and evolutionary relationships. *J Bacteriol* 174(12):3843–3849.
18. Rost B, Yachdav G, Liu J (2004) The PredictProtein server. *Nucleic Acids Res* 32(Web Server Issue):W321–6.
19. Shadrin A, Sheppard C, Severinov K, Matthews S, Wigneshweraraj SR (2012) Substitutions in the Escherichia coli RNA polymerase inhibitor T7 Gp2 that allow inhibition of transcription when the primary interaction interface between Gp2 and RNA polymerase becomes compromised. *Microbiology* 158(Pt 11):2753–2764.
20. Chakraborty A, et al. (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science* 337(6094):591–595.
21. Chin JW, Martin AB, King DS, Wang L, Schultz PG (2002) Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc Natl Acad Sci USA* 99(17):11020–11024.
22. Sato S, et al. (2011) Crystallographic study of a site-specifically cross-linked protein complex with a genetically incorporated photoreactive amino acid. *Biochemistry* 50(2):250–257.
23. Hinton DM, et al. (2005) Transcriptional takeover by sigma appropriation: Remodeling of the sigma70 subunit of Escherichia coli RNA polymerase by the bacteriophage T4 activator MotA and co-activator AsiA. *Microbiology (Reading, Engl.)* 151(Pt 6):1729–1740.
24. Twist K-A, et al. (2011) A novel method for the production of in vivo-assembled, recombinant Escherichia coli RNA polymerase lacking the α C-terminal domain. *Protein Sci* 20(6):986–995.
25. Doublié S (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol* 276:523–530.
26. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data. *Methods Enzymol* 276:307–326.
27. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
28. Opalka N, et al. (2010) Complete structural model of Escherichia coli RNA polymerase from a hybrid approach. *PLoS Biol* 8(9):e1000483.
29. Malhotra A, Severinova E, Darst SA (1996) Crystal structure of a sigma 70 subunit fragment from E. coli RNA polymerase. *Cell* 87(1):127–136.
30. Patikoglou GA, et al. (2007) Crystal structure of the Escherichia coli regulator of σ 70, Rsd, in complex with σ 70 domain 4. *J Mol Biol* 372(3):649–659.
31. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
32. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
33. Schröder GF, Levitt M, Brunger AT (2010) Super-resolution biomolecular crystallography with low-resolution data. *Nature* 464(7292):1218–1222.
34. Brünger AT, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54(Pt 5):905–921.
35. O'Donovan DJ, et al. (2012) A grid-enabled web service for low-resolution crystal structure refinement. *Acta Crystallogr D Biol Crystallogr* 68(Pt 3):261–267.
36. Farrell IS, Toroney R, Hazen JL, Mehl RA, Chin JW (2005) Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nat Methods* 2(5):377–384.
37. Lane WJ, Darst SA (2010) Molecular evolution of multisubunit RNA polymerases: Sequence analysis. *J Mol Biol* 395(4):671–685.
38. Zhang Y, et al. (2012) Structural basis of transcription initiation. *Science* 338(6110):1076–1080.
39. Landau M, et al. (2005) ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* 33(Web Server Issue):W299–302.