Structures of *E. coli* σ^S-transcription initiation complexes provide new insights into polymerase mechanism

Bin Liu[a,1], Yuhong Zuo[a,1,2], and Thomas A. Steitz[a,b,c,2]

[a]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520; [b]Howard Hughes Medical Institute, Yale University, New Haven, CT 06520; and [c]Department of Chemistry, Yale University, New Haven, CT 06520

Edited by Seth A. Darst, The Rockefeller University, New York, NY, and approved March 3, 2016 (received for review October 16, 2015)

In bacteria, multiple σ factors compete to associate with the RNA polymerase (RNAP) core enzyme to form a holoenzyme that is required for promoter recognition. During transcription initiation RNAP remains associated with the upstream promoter DNA via sequence-specific interactions between the σ factor and the promoter DNA while moving downstream for RNA synthesis. As RNA polymerase repetitively adds nucleotides to the 3′-end of the RNA, a pyrophosphate ion is generated after each nucleotide incorporation. It is currently unknown how the release of pyrophosphate affects transcription. Here we report the crystal structures of *E. coli* transcription initiation complexes (TICs) containing the stress-responsive σ^S factor, a de novo synthesized RNA oligonucleotide, and a complete transcription bubble (σ^S-TIC) at about 3.9-Å resolution. The structures show the 3D topology of the σ^S factor and how it recognizes the promoter DNA, including likely specific interactions with the template-strand residues of the −10 element. In addition, σ^S-TIC structures display a highly stressed pretranslocated initiation complex that traps a pyrophosphate at the active site that remains closed. The position of the pyrophosphate and the unusual phosphodiester linkage between the two terminal RNA residues suggest an unfinished nucleotide-addition reaction that is likely at equilibrium between nucleotide addition and pyrophosphorolysis. Although these σ^S-TIC crystals are enzymatically active, they are slow in nucleotide addition, as suggested by an NTP soaking experiment. Pyrophosphate release completes the nucleotide addition reaction and is associated with extensive conformational changes around the secondary channel but causes neither active site opening nor transcript translocation.

transcription initiation | RNA polymerase | σ^S factor | promoter recognition | pyrophosphate release

Cellular organisms transfer genetic information from DNA to RNA using multisubunit RNA polymerases (RNAPs) that are conserved from bacteria to humans (1, 2). In bacteria, a single five-subunit core enzyme of RNA polymerase ($\alpha_2\beta\beta'\omega$) is responsible for all RNA synthesis, whereas multiple σ factors compete to associate with the RNAP core enzyme to form a holoenzyme that is required for initiating the process at DNA promoter sites (3, 4). RNAP remains associated with the upstream promoter DNA during transcription initiation and moves downstream for RNA synthesis, causing DNA scrunching to form a stressed and unstable initiation complex (5–8). Processive RNA synthesis happens only after the initiation complex escapes the promoter as transcription progresses from initiation to elongation (9–11).

RNA synthesis in both transcription initiation and elongation involves repetitive cycles of nucleotide addition comprising translocation, NTP binding, catalysis, and pyrophosphate release steps. During this cycling process, the RNAP active site opens for NTP association and closes to align the incoming NTP with the RNA 3′ hydroxyl group for catalysis. Nucleotide addition extends the RNA by 1 nt and generates a pyrophosphate ion (PPi). Previous structural studies of transcription complexes frequently used DNA fragments that form a partial transcription bubble. Most of these complexes display an open active site in the posttranslocation state, and the active site-closed conformation has been observed only with NTP-bound precatalysis complexes in which the nucleotide addition reaction was prevented by using 3′-deoxy RNA or non-hydrolysable NTP analogs (Table S1). A PPi-associated complex has never been observed previously in cellular RNAP structures. How release of PPi affects the opening of the RNAP active site and transcription translocation remains to be established.

In this work we determined the crystal structures of the *E. coli* transcription initiation complexes (TICs) containing the σ^S factor, a de novo synthesized RNA oligonucleotide, and a complete transcription bubble (σ^S-transcription initiation complexes, σ^S-TICs). The σ^S factor controls the expression of many genes in response to general stresses, such as nutrient deprivation upon entry into stationary phase. The structures show the specific interactions of the σ^S factor with the promoter −10 element and provide insights into the mechanism of σ^S-dependent selective gene expression under stress conditions. In addition, the σ^S-TIC crystals display a pretranslocated initiation complex with a PPi associated at the active site that remains closed. The position of the PPi and the unusual phosphodiester linkage between the two terminal RNA residues suggest an unfinished nucleotide addition reaction that likely is at equilibrium between nucleotide addition and pyrophosphorolysis. The slow enzymatic activity of the σ^S-TIC crystals allowed us to observe PPi dissociation without nucleotide addition in an NTP-soaking experiment. PPi release appears to be associated with extensive conformational changes around the secondary channel but causes neither active site opening nor transcript translocation.

Significance

As RNA polymerase (RNAP) translocates along the DNA template for repetitive nucleotide additions, its active site opens and closes for NTP association and catalysis, and a pyrophosphate ion (PPi) is generated after each nucleotide incorporation. Understanding the role of PPi release is important for elucidating the polymerase mechanism. The structures of the σ^S-containing transcription initiation complexes (σ^S-TICs) provide insights into the mechanism of σ^S-dependent selective gene expression. In addition, the highly stressed σ^S-TICs trap a PPi at the RNAP active site, a previously unobserved but catalytically relevant functional state. Our study also demonstrates that PPi release is not directly related to either translocation or active site opening but causes extensive conformational changes on the periphery of the RNAP secondary channel.

Author contributions: B.L., Y.Z., and T.A.S. designed research; B.L. and Y.Z. performed research; B.L. and Y.Z. analyzed data; and B.L., Y.Z., and T.A.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The structure(s) reported in this paper have been deposited in the Protein Data Bank (PDB) database (PDB ID codes 5IPL, 5IPM, and 5IPN).

[1]B.L. and Y.Z. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: thomas.steitz@yale.edu or yuhong.zuo@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1520555113/-/DCSupplemental.

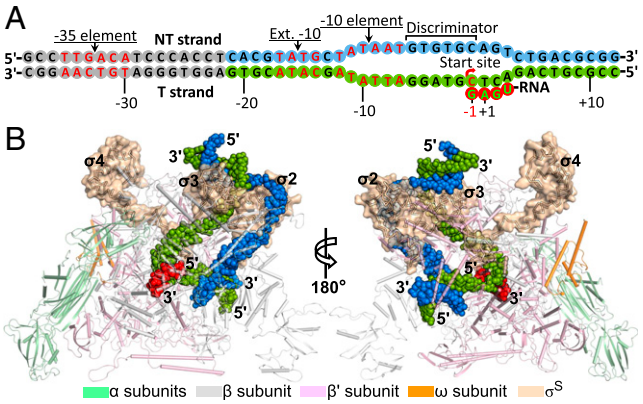

Fig. 1. Overall structure of the *E. coli* σ^S-TIC. (*A*) Schematic representation of the synthetic promoter DNA and a de novo synthesized RNA transcript in the σ^S-TIC crystals. RNA synthesis starts from the −1 position with a GTP residue as observed previously with related promoters in the σ^{70}-TICs (5). The disordered upstream DNA residues are shown as gray cycles. (*B*) Structure of the σ^S-TIC. The *E. coli* RNAP core enzyme is shown in a tube-and-arrow cartoon representation. The σ^S factor is shown as a Cα trace within a surface representation (wheat). The promoter DNA (NT strand, blue; T strand, green) and the nascent RNA (red) are shown as filled spheres.

Results and Discussion

Overall Structure of the *E. coli* σ^S-Based TIC. By using a synthetic DNA scaffold containing the promoter consensus −35 and −10 sequences, we assembled and crystallized the *E. coli* TIC that contains the general stress response σ^S factor. The σ^S-TIC crystals are enzymatically active for additional nucleotide incorporation. The structures were solved by molecular replacement, and the final models were refined to about 3.6-Å resolution (Table 1). The σ^S-TIC structures display a closed active site and a well-ordered nascent RNA–DNA hybrid in the pretranslocation register.

The σ^S-TIC that initially crystallized contains a 14-nt bubble and an RNA tetranucleotide synthesized de novo from three NTPs and traps a PPi at the active site (Fig. 1). Continuous electron densities are seen for both strands of the ssDNA in the bubble region (Fig. S1*A*). Much of the nontemplate (NT) strand of the bubble that is downstream of the −10 hexamer could not be resolved unambiguously in the density map, suggesting a slippery feature of this DNA segment that displays limited interactions with the protein. The ssDNA residues on the template (T) strand are much better resolved, although residues in the middle of this section appear disordered. The upstream promoter DNA runs into the downstream DNA duplex of a symmetry-related molecule, causing strand separation and disordering of a significant portion of the upstream DNA (Fig. S1*B*).

***E. coli* σ^S Factor and σ^S-RNAP Core Interactions.** Like the primary σ factors, the *E. coli* σ^S factor contains a highly negatively charged N-terminal domain ($\sigma_{1.1}$) that we have not been able to trace in the σ^S-TIC structures. Other than a nonconserved region (NCR) that is commonly present in primary σ factors, σ^S displays very high sequence identity with the primary σ factors (Fig. S2*A*) (10). Consistent with their sequence similarity, the σ^S factor in the σ^S-TIC forms essentially the same fold as the σ^{70} factor throughout the conserved regions, from region 1.2 to the very C terminus, including the long $\sigma_{3.2}$ loop that inserts deep into the RNAP active site chamber (Fig. 2*A*).

A major contribution to the σ-RNAP core association comes from interactions between the σ_2 domain and a helix-turn-helix formation, termed a "clamp helices," of the clamp domain of the β′ subunit. Compared with σ^{70}, which displays strong affinity for the RNAP core enzyme (K_d ~0.26 nM), σ^S binds relatively weakly to the RNAP core in the absence of nucleic acids (K_d ~4.26 nM) (12). This weaker affinity of σ^S might be attributed partly to its lack of an NCR domain that provides additional interactions with the clamp helices of the RNAP core enzyme (Fig. S2*B*).

Table 1. Data collection and refinement statistics

Parameters	σ^S-TIC1	σ^S-TIC2 1-h soaking with CTP only	σ^S-TIC3 1-h soaking with CTP, UTP, and GTP
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions, Å			
a	131.71	132.75	132.87
b	152.67	151.97	152.17
c	226.70	228.06	229.12
Resolution, Å	100–3.60 (3.66–3.60)	100–4.20 (4.27–4.20)	100–4.60 (4.68–4.60)
R_{sym}, %	12.5 (>100)	10.2 (>100)	8.5 (>100)
$I/\sigma I$*	9.87 (0.37)	8.44 (0.46)	10.06 (0.55)
Completeness, %	99.6 (99.4)	99.9 (100)	99.3 (99.6)
Redundancy	5.1 (4.2)	7.7 (7.9)	6.9 (6.8)
Refinement			
Resolution, Å	100–3.60	100–4.20	100–4.60
No. reflections	50,952	32,692	24,772
R_{work}/R_{free}, %	24.7/28.7	26.7/32.7	24.0/33.4
No. atoms	27,632	29,036	29,036
Protein	26,167	27,572	27,572
DNA/RNA/ions	1,465	1,464	1,464
B factors	129.5	177.4	183.1
Rmsd			
Bond length, Å	0.011	0.011	0.008
Bond angle, °	1.448	1.422	1.179

Values in parentheses are for highest-resolution shell. Each dataset was collected from a single crystal.
*$I/\sigma I$ = 2.0 at 4.2 Å (σ^S-TIC1), 5.6 Å (σ^S-TIC2), 6.2 Å (σ^S-TIC3); $I/\sigma I$ = 1.0 at 3.9 Å (σ^S-TIC1), 4.7 Å (σ^S-TIC2), 5.2 Å (σ^S-TIC3).

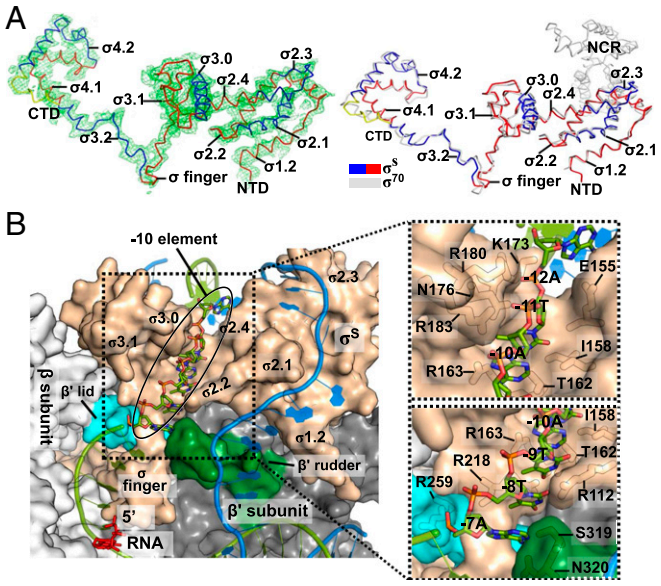

Fig. 2. *E. coli* σ^S factor. (*A*) Overall structure of the *E. coli* σ^S factor in the σ^S-TIC. The orientation of σ^S here is the same as in Fig. 1. (*Left*) The *Fo-Fc* electron density map (mesh, contoured at 2.0 σ) was calculated using the phases from the RNAP core-only model. (*Right*) The *E. coli* σ^S factor in the σ^S-TIC (colored) and the σ^{70} factor in the σ^{70}-holoenzyme (PDB ID code4JKR) (gray) are superimposed. (*B*) Recognition of the promoter −10 element in the σ^S-TIC. The N-terminal half of the β subunit (residues 2–667) is omitted for clarity. The nascent RNA (red) and the −10 hexamer residues of the T strand (green) are shown as sticks. The rest of the nucleic acid is shown in cartoon style (T strand, green; NT strand, marine). The RNA polymerase is shown in surface representation: σ^S, wheat; β subunit, light gray; β′ lid, cyan; β′ rudder, forest green; the rest of the β′ subunit is shown in dark gray. A few residues that interact directly with the −10 T-strand residues are labeled in the insets.

In *E. coli* and many γ-proteobacteria, a small protein called "Crl" was found to stimulate σ^S-dependent transcription by promoting the formation of the σ^S–RNAP holoenzyme (13, 14). It was shown previously that the σ^S–Crl interaction involves a general area on the surface of σ^S including residues of Asp87, Asp135, Pro136, and Glu137 (15). This general area corresponds to the attachment face of the NCR to the primary σ factors (Fig. S2*B*), suggesting that Crl might function like the NCR of a primary σ factor to help encompass the clamp helices and enhance the interactions of σ^S with the core enzyme.

Promoter Recognition by *E. coli* σ^S Factor. The *E. coli* σ^S factor was shown to recognize the same consensus −35 and −10 hexamers as the σ^{70} factor in an in vitro selection experiment (16). The σ^S-TIC structures show that σ^S recognizes the −10 hexamer of the NT strand with specific interactions with the bases of the −7T and −11A residues in the same manner as the primary σ factors (Fig. S2*C*) (5, 17–19). It was suggested recently that the −10 T strand passes through a cleft between the σ_2 and σ_3 domains in an initiation complex (5, 17), but it remained unclear how the −10 T strand interacts with the RNAP. The σ^S-TIC structures suggest that the single-stranded T-strand residues −7 to −11 stack on each other inside the narrow tunnel formed by the σ_2–σ_3 cleft and the β1 domain (Fig. 2*B*). The DNA backbone appears to make extensive interactions with positively charged side chains from the β1 domain and the σ_2 and σ_3 domains, and the bases may form several hydrogen-bonding interactions with the σ^S factor (Fig. S3), raising the possibility that the T-strand residues may contribute to the specific recognition of the −10 hexamer by the σ^{70} family factors. Additional evidence is required to determine whether the −10 template region contributes to the promoter recognition.

The few differences between σ^S and σ^{70} in their interactions with the extended −10 region include residues Ile169 and Lys173 in σ^S and the corresponding residues Val454 and Glu458 in σ^{70}. It was reported that Lys173 is responsible for a preference of cytidine by σ^S at the promoter −13 site (20). Lys173 potentially could interact with the −13 base of the T strand, but a specific recognition could not be established at the current resolution.

In the σ^S-TIC crystals shown here, the helix-turn-helix of the σ^S C-terminal domain (σ_4) that potentially interacts with the promoter −35 hexamer, including helix H42 and the N-terminal half of helix H43, is involved in crystal packing, and thus there is no space to accommodate the −35 region of the promoter DNA duplex. Although σ^S interactions with the promoter −35 element are not observed in the σ^S-TIC crystals, the sequence and structural conservation and the shared recognition of promoter sequences suggest that the σ^S and σ^{70} factors would interact with the DNA promoters in a very similar manner.

Selective Gene Expression Under Stress Conditions. An intriguing question is how σ^S achieves selective gene expression, given that it recognizes consensus −35 and −10 sequences essentially identical to those recognized by σ^{70}. The σ^S-dependent promoters display higher sequence deviations from the consensus −35 hexamer (21); these deviations, again, might be related to the absence of an NCR domain in the σ^S factor. Although the NCR domain in σ^{70} might make it interact better with the RNAP core enzyme, a stretch of acidic residues (18 of the *E. coli* σ^{70} residues 188–209) (Fig. S2*B*) is expected to inhibit promoter loading directly through interactions of the σ^{70}-holoenzyme with the −10 element (22), thus making interactions with the −35 element important for recruiting the σ^{70}-holoenzyme to the promoter. In contrast, the smaller σ_2 domain of the σ^S factor would allow it to interact directly with the promoter −10/extended −10 region, thus making the σ^S-holoenzyme less dependent on the −35 element for promoter loading.

The protein concentration of the σ^S factor in *E. coli* is tightly regulated at the levels of transcription, translation, and protein stability (23, 24). During exponential growth in rich medium, the σ^S protein level in *E. coli* is negligible, but when *E. coli* enters the stationary phase or under certain stress conditions the σ^S protein level might increase by a thousand-fold and reach a level comparable to that of σ^{70}. Easier access not only would provide the σ^S-holoenzyme an advantage in competing for promoters in the heavily packed DNA during the stationary phase but also would justify the requirement of tight regulation of the σ^S protein concentration under normal growth conditions.

Stressed TICs. The σ^{70}-TICs that we reported recently contain a complete transcription bubble and display a well-ordered nascent RNA–DNA hybrid lying at the pretranslocation position; we suggested that this pretranslocated hybrid may be a manifestation of the stressed feature of an initiation complex (5). Similar to the σ^{70}-TICs, the σ^S-TICs we report here also contain a complete transcription bubble and an RNA oligonucleotide synthesized de novo from NTPs, and the σ^S-TIC structures display a pretranslocated RNA–DNA hybrid as well (Fig. 3). Moreover, the active site of the σ^S-TICs is fully closed by the folding of the trigger loop (TL) into helices (TH). Although the helical conformation of the TL in a pretranslocated transcription complex also might exist in our low-resolution σ^{70}-TIC structures (5), it has never been observed previously in any other transcription complexes.

It has been shown that conformational changes of a flexible TL and the bridge helix (BH), which traverses across the active site cleft, remodel the active site of cellular RNA polymerases (1, 2, 25–27). An unfolded TL leaves the active site open to the secondary channel and thus allows NTP binding to the active site, whereas the folding of the TL into two α-helices (THs) closes the active site and helps align the incoming NTP with the RNA 3′-hydroxyl group and with a conserved histidine residue (β′His936

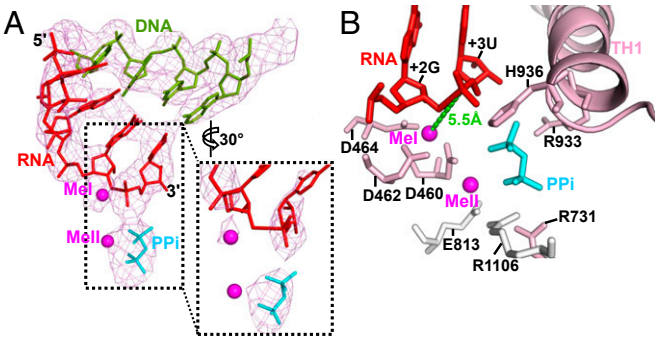

Fig. 3. σ^S-TIC active site. (*A*) A close-up view of the DNA–RNA hybrid and PPi at the active site in the structure before NTP soaking. The *Fo-Fc* electron density map (mesh, contoured at 3 σ) was calculated using the phases calculated from the protein-only model. The inset shows a zoomed-in view of the electron density (contoured at 5 σ) around the catalytic site and the RNA 3′-terminal residue. (*B*) A close-up view of the active site showing the major interactions involved in stabilizing the PPi. The metal ions (MeI and MeII) are shown as magenta spheres.

in *E. coli* RNAP) that serves as a general acid for catalysis (25, 27). The BH stacks against the RNA–DNA hybrid and could bend or bulge toward the template base, which forms the basis of the BH-controlled transcription translocation model (28, 29). The TL folding has a profound effect on the rate of transcription (30) and is affected by the BH conformation (25, 27). It was also suggested that transcription translocation might require an unfolded TL (31).

In the stressed σ^S-TIC, the BH displays a slight bending perpendicular to the translocation direction with a shift of about 3 Å around β′ residues 785–789 (Fig. S4*A*). This small shift in the BH appears to create enough space to allow the TL to fold into helices. TH formation significantly reduces the dimensions of the secondary channel entrance to the active site, from about 15 × 12 Å in the RNAP apo holoenzyme to about 8 × 6 Å in the σ^S-TIC (Fig. S4*B*), and thus even the dissociation of PPi might involve a concerted effort by the protein, involving more than amino acid side chain movements. Intriguingly, the phosphate of the RNA 3′-terminal residue appears not to interact directly with the metal ion (MeI) coordinated by the conserved carboxyl triad at the active center (Fig. 3*A*). The electron density contoured at higher levels displays some discontinuity between the two terminal residues, and the linkage of the terminal phosphate (+3U) to the penultimate RNA residue (+2G) also appears to be significantly distorted at the sugar ring of the penultimate residue. This finding suggests that the σ^S-TIC crystal might contain a mixture of initiation complexes, likely at equilibrium between the forward nucleotide addition reaction and presumably its reverse reaction, pyrophosphorolysis (32–34).

Unreleased Pyrophosphate in the σ^S-TIC Crystals. Because RNA polymerase adds nucleotides repetitively to the RNA 3′ end, one pyrophosphate ion is generated after each addition reaction. During this process, it is generally believed that the active site closes after NTP association and opens immediately after or concurrently with the dissociation of the PPi. For T7 RNA polymerase (T7RNAP), a PPi-associated complex has been obtained by supplementing the solution for crystallization with PPi; the PPi-bound and unbound states of T7RNAP were found to associate with the pretranslocation and posttranslocation states, which correspond to the closed and open conformations of the active site, respectively (35). However, a PPi-associated structure has never been observed for the cellular RNA polymerases (Table S1). It was shown previously that translocation occurs shortly after or concurrently with PPi release (36). However, how PPi release affects the opening of the RNAP active site and the translocation of the enzyme along the DNA template remains obscure.

The intrinsic abortive feature of transcription initiation would cause many rounds of oligonucleotide synthesis during the process of complex assembly and crystallization, and these multiple rounds of synthesis could lead to a significant accumulation of PPi in the crystallization drops that reaches a level comparable to or even exceeding the levels of NTPs. On the other hand, because of DNA scrunching, TICs are stressed and tend to rest at the pretranslocation state, a conformation clearly more favorable than others for PPi association. Not surprisingly, a well-ordered PPi remains associated with the σ^S-TIC active site.

The pyrophosphate in the σ^S-TIC interacts via a metal ion (MeII) with the carboxyl group of β′Asp460, one of the conserved carboxyl triad of the active center, with the side chains of the TL residues β′Arg933 and β′His936, and with the side chains of βArg1106 and β′Arg731 that line as the secondary channel (Fig. 3*B*). The observation that both the PPi and the phosphate of the RNA 3′-terminal residue interact with the side chain of β′His936 is consistent with the proposal that this conserved TL histidine residue is involved in both nucleotide addition and pyrophosphorolysis (25, 27). It is not clear whether the position and the network of PPi interactions we observed here also represent those of the β- and γ-phosphates of an incoming NTP before nucleotide incorporation, although it is likely that they do (Fig. S5).

Dissociation of PPi and RNAP Conformational Changes. To test the ability of the σ^S-TIC crystals to incorporate nucleotides, we soaked the crystals in solutions containing CTP that was omitted in the original complex assembly. Unlike the σ^{70}-TIC crystals that readily extend the RNA by 1 nt after soaking (5), the σ^S-TIC crystals

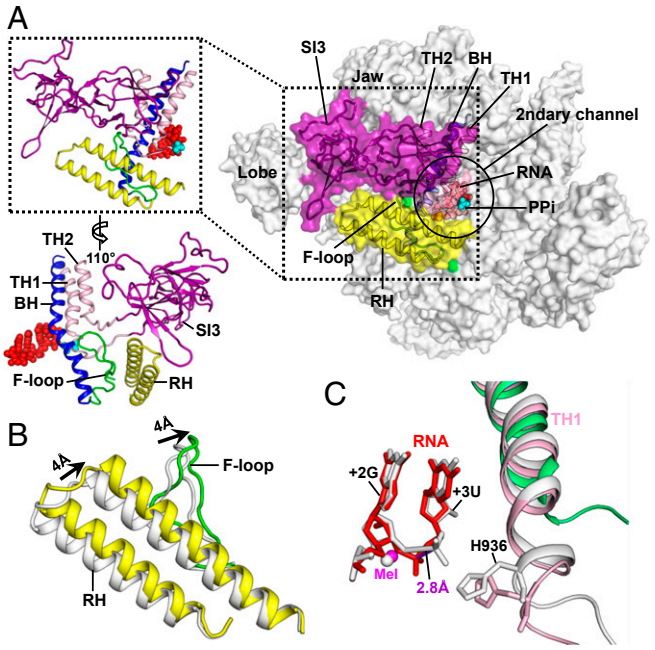

Fig. 4. Conformational changes after PPi release. (*A*) The secondary channel of the σ^S-TIC after the crystals were soaked in NTP-containing solutions. RNAP is shown in surface representation. The TL insertion domain (SI3) (purple), F-loop (green), RHs (yellow), BH (blue), and TH1 and TH2 (pink) are also presented in cartoon representations. The insets on the left are close-up views of the flexible structural elements. The nascent RNA (red) and Ppi (cyan) are shown as filled spheres. The PPi was modeled in as a reference. (*B*) A close-up view of the F-loop and RHs in the σ^S-TICs before (gray) and after (yellow) NTP soaking and PPi release. (*C*) A close-up view of the active site of the σ^S-TICs before (gray) and after (pink and red) PPi release. The corresponding region of the *E. coli* RNAP apo-holoenzyme (PDB ID code 4LJZ) (green) is shown for reference. The structures are superimposed on the β subunits.

appear to be much less efficient in RNA synthesis. Soaking σ^S-TIC crystals with CTP only (for a 1-nt extension) or with three NTPs (CTP, UTP, and GTP, for RNA extension of up to 3 nt) for 1 h did not result in noticeable nucleotide addition. RNA extension was observed only after NTP soaking for 2 h or longer. It remains unclear why the σ^S-TIC crystals display such a low reactivity, but it is worth mentioning that the concentration of free Mg^{2+} might be low during the soaking of NTP into the crystals.

Although soaking of the σ^S-TIC crystals in solutions containing NTPs for 1 h did not lead to nucleotide incorporation, the two NTP soaking experiments led to similar, significant conformational changes in both the protein and the nucleic acid in the σ^S-TIC crystals (Fig. 4). Because the soaking solution contains no PPi, soaking leads to the release of the associated PPi as expected, and coordinately, the phosphate group of the RNA 3′-terminal residue shifts toward MeI and makes a more normal phosphodiester linkage with the penultimate RNA residue, suggesting that completion of the nucleotide addition reaction might require PPi release. Interestingly, it appears that there is no noticeable NTP binding to the NTP "entry site" that would overlap with the PPi-binding site (27, 37).

E. coli RNA polymerase carries a large lineage-specific insertion, termed "sequence insertion 3" (SI3), in the middle of the TL (38). This TL insertion of 188 amino acids could make a large shift around the secondary channel (39) and has been shown to affect *E. coli* transcription because it disfavors TH formation and stimulates pausing (40). In all of the three σ^S-TIC crystals described here, SI3 is expected to make similar contacts with symmetry-related molecules and contributes negligibly to crystal packing. However, a long hairpin loop in the C-terminal domain of SI3 projects into an intermolecular space and likely prevents SI3 from making large movements. The single most prominent conformational change on the σ^S-TIC after NTP soaking appears to be that the normally disordered SI3 becomes ordered and visible in the electron density map. In this ordered conformation SI3 fills the gap between the lobe and jaw domains of *E. coli* RNAP and forms part of the wall that separates the primary downstream DNA channel and the secondary channel (Fig. 4*A*).

In addition to SI3 becoming localized and ordered, the rim helices (β′ residues 650–703) on the edge of the secondary channel are rotated by about 7° toward the SI3–TH linkers (a shift of about 4 Å at the far end). At the same time, the F-loop (β′ residues 742–762) that interacts with both BH and TH makes a similar rotation to press on the SI3–TH linkers (Fig. 4*B*). The conformational changes of the rim helices and the F-loop, which are likely induced by the release of PPi, appear to function together to tighten the SI3–TH linkers and to lock SI3 against the jaw domain with its cleft between two β-folds (Fig. S6).

Because NTP soaking caused very limited changes to the crystal packing, it is likely that the significant conformational changes of SI3, the rim helices, and the F-loop are directly related to the release of PPi. After PPi dissociation, there appears to be a very minor shift (~1–2 Å) at the C terminus of the first trigger helix (TH1) (Fig. 4*C*), and the σ^S-TIC remains at the pretranslocated position with a closed active site.

Concluding Remarks. The thermodynamic (Brownian ratchet) model of transcription translocation suggests that RNA polymerase oscillates between pre- and posttranslocation states with the forward movement biased by NTP binding. Interestingly, structures of cellular RNA polymerase complexes are overwhelmingly at the posttranslocation state with an open active site (Table S1). Although structural studies could not be used as an evidence for thermodynamic analysis, the structural preference suggests that the posttranslocation state with an open active site might be the thermodynamically more favored state. However, in both the σ^{70}-TIC and σ^S-TIC crystals that contain a complete bubble, all the hybrids that we observed lie at the pretranslocation position, and the active site remains closed. It is likely that the structural preference seen in previous structural studies is a biased representation directly related to the substrate designs, which frequently involve DNA fragments lacking the upstream portion of the transcription bubble. Apparently, transcription complexes containing a complete transcription bubble to showcase the thermodynamic barriers for translocation would be of more biological significance, as we show here with the σ^S-TICs.

Intriguing questions are how fast PPi release happens and how PPi release is related to active site opening and transcription translocation. For T7 RNA polymerase, it was suggested that PPi release is directly coupled with active site opening and translocation (35). For cellular RNA polymerases, PPi release was suggested to be a thermodynamically controlled quick process coupled to the RNAP conformational change that is associated with the binding of the next cognate nucleotide (32). The NTP soaking experiment presented here displays release of PPi by the σ^S-TIC crystals and shows that the active site of the σ^S-TIC remains closed after PPi release. Although PPi release appears to be associated with a small shift at the C terminus of TH1 (Fig. 4*C*), a movement that likely occurs during the unfolding of TH to TL, it causes neither active site opening nor transcript translocation. However, it is possible that a reversal of the observed conformational changes on the periphery of the secondary channel could be coupled to TL unfolding and active site opening.

Materials and Methods

Preparation and Crystallization of *E. coli* σ^S-TIC. To form the σ^S-TIC, we used a synthetic DNA scaffold corresponding to the promoter region between positions −38 and +12 relative to the expected transcription start site (Fig. 1*A*). The synthetic promoter, which contains the consensus −35 and −10 hexamers and the extended −10 motif, was prepared by annealing the NT strand to an equal molar amount of the T-strand DNA that is complementary to the NT strand except for a 6-nt discriminator region (Fig. 1*A*). The σ^S-TIC was assembled by directly incubating the σ^S-RNAP holoenzyme with a twofold molar excess of the preformed DNA promoter in buffer A [20 mM Tris (pH 7.5), 50 mM NaCl, 0.1 mM EDTA, 5 mM $MgCl_2$] at 37 °C for 20 min in the presence of ATP, GTP, and UTP (2 mM each). This mixture then was used for crystallization at room temperature by vapor diffusion with a reservoir containing 18% (wt/vol) PEG 3350, 0.1 M NaCl, and 0.1 M Hepes (pH 7.8). After the crystals grew to full size for (for about 1 wk), they were cryo-protected in the mother liquor containing 15% (wt/vol) ethylene glycol before flash-freezing in liquid nitrogen. The σ^S-TIC crystallizes in the orthorhombic $P2_12_12_1$ space group with one copy of the complex per asymmetric unit different from that of the σ^{70}-TICs (5).

NTP Soaking for RNA Synthesis in the Crystal. The initially obtained σ^S-TIC crystals containing an RNA tetranucleotide were soaked in the reservoir solution supplemented with either CTP only (for a 1-nt extension) or three NTPs (CTP, UTP, and GTP, for an RNA extension of up to 3 nt) (2 mM each) at room temperature for various time periods. The crystals then were cryo-protected and flash-frozen in the same manner as aforementioned.

Data Collection, Processing, and Structure Determination. X-ray diffraction data were collected at 100 K at the beamlines 24-ID-C and 24-ID-E at Argonne National Laboratory, Chicago, IL. All data were integrated and scaled with HKL2000 (41). The structures were solved by molecular replacement with PHASER (42) using a structure of the *E. coli* σ^{70}-TIC (5) as the starting model. The molecular replacement solution was subjected to rigid body refinement with Refmac5 (43) using multiple rigid groups, and the phases were improved by density modification. The maps were improved further by temperature factor sharpening that allowed building the σ^S factor and the nucleic acid models into the density using COOT (44). After model building in Coot, 10 cycles of TLS (translation libration screw-motion) and restrained refinement were performed using Refmac5 (43) in the CCP4 suite (45). Data collection and structural refinement statistics are summarized in Table 1. All figures were created using PyMOL (46).

ACKNOWLEDGMENTS. We thank the staff of Argonne National Laboratory beamlines 24-ID-C and 24-ID-E for help during data collection, the Center for Structural Biology Facility at Yale University for computational support, and Dr. Jimin Wang for contributions to our data processing and structural refinement. This work was supported by NIH Grant GM22778 (to T.A.S.). T.A.S. is an investigator of the Howard Hughes Medical Institute.

1. Murakami KS, Darst SA (2003) Bacterial RNA polymerases: The wholo story. *Curr Opin Struct Biol* 13(1):31–39.
2. Cramer P (2002) Multisubunit RNA polymerases. *Curr Opin Struct Biol* 12(1):89–97.
3. Saecker RM, Record MT, Jr, Dehaseth PL (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(5):754–771.
4. Burgess RR, Travers AA, Dunn JJ, Bautz EK (1969) Factor stimulating transcription by RNA polymerase. *Nature* 221(5175):43–46.
5. Zuo Y, Steitz TA (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol Cell* 58(3):534–540.
6. Revyakin A, Liu C, Ebright RH, Strick TR (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* 314(5802):1139–1143.
7. Kapanidis AN, et al. (2006) Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* 314(5802):1144–1147.
8. Straney DC, Crothers DM (1987) A stressed intermediate in the formation of stably initiated RNA chains at the *Escherichia coli* lac UV5 promoter. *J Mol Biol* 193(2): 267–278.
9. Goldman SR, Ebright RH, Nickels BE (2009) Direct detection of abortive RNA transcripts in vivo. *Science* 324(5929):927–928.
10. Gruber TM, Gross CA (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* 57:441–466.
11. Hansen UM, McClure WR (1979) A noncycling activity assay for the omega subunit of *Escherichia coli* RNA polymerase. *J Biol Chem* 254(13):5713–5717.
12. Maeda H, Fujita N, Ishihama A (2000) Competition among seven *Escherichia coli* sigma subunits: Relative binding affinities to the core RNA polymerase. *Nucleic Acids Res* 28(18):3497–3503.
13. Gaal T, Mandel MJ, Silhavy TJ, Gourse RL (2006) Crl facilitates RNA polymerase holoenzyme formation. *J Bacteriol* 188(22):7966–7970.
14. Pratt LA, Silhavy TJ (1998) Crl stimulates RpoS activity during stationary phase. *Mol Microbiol* 29(5):1225–1236.
15. Banta AB, et al. (2013) Key features of σS required for specific recognition by Crl, a transcription factor promoting assembly of RNA polymerase holoenzyme. *Proc Natl Acad Sci USA* 110(40):15955–15960.
16. Gaal T, et al. (2001) Promoter recognition and discrimination by EsigmaS RNA polymerase. *Mol Microbiol* 42(4):939–954.
17. Bae B, Feklistov A, Lass-Napiorkowska A, Landick R, Darst SA (2015) Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *eLife* 4:e08504.
18. Zhang Y, et al. (2012) Structural basis of transcription initiation. *Science* 338(6110): 1076–1080.
19. Feklistov A, Darst SA (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* 147(6):1257–1269.
20. Becker G, Hengge-Aronis R (2001) What makes an *Escherichia coli* promoter sigma(S) dependent? Role of the -13/-14 nucleotide promoter positions and region 2.5 of sigma(S). *Mol Microbiol* 39(5):1153–1165.
21. Wise A, Brems R, Ramakrishnan V, Villarejo M (1996) Sequences in the -35 region of *Escherichia coli* rpoS-dependent genes promote transcription by E sigma S. *J Bacteriol* 178(10):2785–2793.
22. 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.
23. Battesti A, Majdalani N, Gottesman S (2011) The RpoS-mediated general stress response in *Escherichia coli*. *Annu Rev Microbiol* 65:189–213.
24. Hengge R (2009) Proteolysis of sigmaS (RpoS) and the general stress response in *Escherichia coli*. *Res Microbiol* 160(9):667–676.
25. Vassylyev DG, et al. (2007) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448(7150):163–168.
26. Toulokhonov I, Zhang J, Palangat M, Landick R (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol Cell* 27(3):406–419.
27. Wang D, Bushnell DA, Westover KD, Kaplan CD, Kornberg RD (2006) Structural basis of transcription: Role of the trigger loop in substrate specificity and catalysis. *Cell* 127(5):941–954.
28. Vassylyev DG, et al. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* 417(6890):712–719.
29. Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 A resolution. *Science* 292(5523):1876–1882.
30. Mejia YX, Nudler E, Bustamante C (2015) Trigger loop folding determines transcription rate of *Escherichia coli*'s RNA polymerase. *Proc Natl Acad Sci USA* 112(3):743–748.
31. Brueckner F, Cramer P (2008) Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nat Struct Mol Biol* 15(8):811–818.
32. Johnson RS, Strausbauch M, Carraway JK (2011) Rapid pyrophosphate release from transcriptional elongation complexes appears to be coupled to a nucleotide-induced conformational change in *E. coli* core polymerase. *J Mol Biol* 412(5):849–861.
33. Rozovskaya TA, Chenchik AA, Beabealashvilli RSh (1982) Processive pyrophosphorolysis of RNA by *Escherichia coli* RNA polymerase. *FEBS Lett* 137(1):100–104.
34. Maitra U, Hurwitz J (1967) The role of deoxyribonucleic acid in ribonucleic acid synthesis. 13. Modified purification procedure and additional properties of ribonucleic acid polymerase from *Escherichia coli* W. *J Biol Chem* 242(21):4897–4907.
35. Yin YW, Steitz TA (2004) The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* 116(3):393–404.
36. Malinen AM, et al. (2012) Active site opening and closure control translocation of multisubunit RNA polymerase. *Nucleic Acids Res* 40(15):7442–7451.
37. Westover KD, Bushnell DA, Kornberg RD (2004) Structural basis of transcription: Nucleotide selection by rotation in the RNA polymerase II active center. *Cell* 119(4): 481–489.
38. Chlenov M, et al. (2005) Structure and function of lineage-specific sequence insertions in the bacterial RNA polymerase beta' subunit. *J Mol Biol* 353(1):138–154.
39. Yang Y, et al. (2015) TRANSCRIPTION. Structures of the RNA polymerase-σ54 reveal new and conserved regulatory strategies. *Science* 349(6250):882–885.
40. Windgassen TA, et al. (2014) Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by *Escherichia coli* RNA polymerase. *Nucleic Acids Res* 42(20):12707–12721.
41. Otwinowski Z, Minor W (1997) *Methods Enzymol*, eds Carter CW, Sweet RM (Academic, New York), pp 307–326.
42. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
43. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355–367.
44. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
45. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235–242.
46. DeLano WL (2002) *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA).
47. Liu B, Zuo Y, Steitz TA (2015) Structural basis for transcription reactivation by RapA. *Proc Natl Acad Sci USA* 112(7):2006–2010.
48. Opalka N, et al. (2010) Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS Biol* 8(9):e1000483.
49. Wang D, Zhu G, Huang X, Lippard SJ (2010) X-ray structure and mechanism of RNA polymerase II stalled at an antineoplastic monofunctional platinum-DNA adduct. *Proc Natl Acad Sci USA* 107(21):9584–9589.
50. Basu RS, et al. (2014) Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme. *J Biol Chem* 289(35):24549–24559.
51. Liu X, Bushnell DA, Silva DA, Huang X, Kornberg RD (2011) Initiation complex structure and promoter proofreading. *Science* 333(6042):633–637.
52. Cheung AC, Sainsbury S, Cramer P (2011) Structural basis of initial RNA polymerase II transcription. *EMBO J* 30(23):4755–4763.
53. Kinkelin K, et al. (2013) Structures of RNA polymerase II complexes with Bye1, a chromatin-binding PHF3/DIDO homologue. *Proc Natl Acad Sci USA* 110(38): 15277–15282.
54. Westover KD, Bushnell DA, Kornberg RD (2004) Structural basis of transcription: Separation of RNA from DNA by RNA polymerase II. *Science* 303(5660):1014–1016.
55. Kettenberger H, Armache KJ, Cramer P (2004) Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol Cell* 16(6):955–965.
56. Brueckner F, Hennecke U, Carell T, Cramer P (2007) CPD damage recognition by transcribing RNA polymerase II. *Science* 315(5813):859–862.
57. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, Artsimovitch I (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448(7150):157–162.
58. Damsma GE, Alt A, Brueckner F, Carell T, Cramer P (2007) Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat Struct Mol Biol* 14(12):1127–1133.
59. Lehmann E, Brueckner F, Cramer P (2007) Molecular basis of RNA-dependent RNA polymerase II activity. *Nature* 450(7168):445–449.
60. Dengl S, Cramer P (2009) Torpedo nuclease Rat1 is insufficient to terminate RNA polymerase II in vitro. *J Biol Chem* 284(32):21270–21279.
61. Sydow JF, et al. (2009) Structural basis of transcription: Mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol Cell* 34(6):710–721.
62. Damsma GE, Cramer P (2009) Molecular basis of transcriptional mutagenesis at 8-oxoguanine. *J Biol Chem* 284(46):31658–31663.
63. Walmacq C, et al. (2012) Mechanism of translesion transcription by RNA polymerase II and its role in cellular resistance to DNA damage. *Mol Cell* 46(1):18–29.
64. Sainsbury S, Niesser J, Cramer P (2013) Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. *Nature* 493(7432):437–440.
65. Bae B, et al. (2015) CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. *eLife* 4:e08505.
66. Barnes CO, et al. (2015) Crystal Structure of a Transcribing RNA Polymerase II Complex Reveals a Complete Transcription Bubble. *Mol Cell* 59(2):258–269.
67. Zhang Y, et al. (2014) GE23077 binds to the RNA polymerase 'i' and 'i+1' sites and prevents the binding of initiating nucleotides. *eLife* 3:e02450.
68. Tagami S, et al. (2010) Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. *Nature* 468(7326):978–982.
69. Weixlbaumer A, Leon K, Landick R, Darst SA (2013) Structural basis of transcriptional pausing in bacteria. *Cell* 152(3):431–441.
70. Wang D, et al. (2009) Structural basis of transcription: Backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* 324(5931):1203–1206.
71. Cheung AC, Cramer P (2011) Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* 471(7337):249–253.