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 10 factor, a de novo synthesized RNA oligonucleotide, and a complete transcription bubble (σ 1-TIC) at about 3.9-Å resolution. The structures show the 3D topology of the σ 1 factor and how it recognizes the promoter DNA, including likely specific interactions with the template-strand residues of the −10 element. In addition, σ 1-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 1-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.

**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 1-containing transcription initiation complexes (σ 1-TICs) provide insights into the mechanism of σ 1-dependent selective gene expression under stress conditions. In addition, the σ 1-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 1-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.

**Author contributions:** B.L., Y.Z., and T.A.S. designed research; B.L. and Y.Z. performed transcription initiation | RNA polymerase | σ 1 factor | promoter recognition | pyrophosphate release

**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).

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

**T.A.S.** was supported by the National Science Foundation (grant MCB-1021042) and the National Institutes of Health (grant GM084266), and T.A.S. is a Pew Scholar in the Biological Sciences and an Investigator of the Howard Hughes Medical Institute. B.L. was supported by graduate fellowships from the National Science Foundation and the University of California, Santa Cruz.

**Data availability:** Freely available online through the PNAS open access option.

**Conflict of interest statement:** No competing interests declared.

**As the 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 1-containing transcription initiation complexes (σ 1-TICs) provide insights into the mechanism of σ 1-dependent selective gene expression under stress conditions. In addition, the highly stressed σ 1-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 transcription or active site opening but causes extensive conformational changes on the periphery of the RNAP secondary channel.**

**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 1-containing transcription initiation complexes (σ 1-TICs) provide insights into the mechanism of σ 1-dependent selective gene expression under stress conditions. In addition, the highly stressed σ 1-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 transcription 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.**

**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).

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

**T.A.S.** was supported by the National Science Foundation (grant MCB-1021042) and the National Institutes of Health (grant GM084266), and T.A.S. is a Pew Scholar in the Biological Sciences and an Investigator of the Howard Hughes Medical Institute. B.L. was supported by graduate fellowships from the National Science Foundation and the University of California, Santa Cruz.

**Data availability:** Freely available online through the PNAS open access option.

**Conflict of interest statement:** No competing interests declared.

**As the 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 1-containing transcription initiation complexes (σ 1-TICs) provide insights into the mechanism of σ 1-dependent selective gene expression under stress conditions. In addition, the highly stressed σ 1-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 transcription or active site opening but causes extensive conformational changes on the periphery of the RNAP secondary channel.**

**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 1-containing transcription initiation complexes (σ 1-TICs) provide insights into the mechanism of σ 1-dependent selective gene expression under stress conditions. In addition, the highly stressed σ 1-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 transcription 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.**

**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).

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

**T.A.S.** was supported by the National Science Foundation (grant MCB-1021042) and the National Institutes of Health (grant GM084266), and T.A.S. is a Pew Scholar in the Biological Sciences and an Investigator of the Howard Hughes Medical Institute. B.L. was supported by graduate fellowships from the National Science Foundation and the University of California, Santa Cruz.

**Data availability:** Freely available online through the PNAS open access option.

**Conflict of interest statement:** No competing interests declared.

**As the 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 1-containing transcription initiation complexes (σ 1-TICs) provide insights into the mechanism of σ 1-dependent selective gene expression under stress conditions. In addition, the highly stressed σ 1-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 transcription or active site opening but causes extensive conformational changes on the periphery of the RNAP secondary channel.**

**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 1-containing transcription initiation complexes (σ 1-TICs) provide insights into the mechanism of σ 1-dependent selective gene expression under stress conditions. In addition, the highly stressed σ 1-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 transcription 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.**

**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).

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

**T.A.S.** was supported by the National Science Foundation (grant MCB-1021042) and the National Institutes of Health (grant GM084266), and T.A.S. is a Pew Scholar in the Biological Sciences and an Investigator of the Howard Hughes Medical Institute. B.L. was supported by graduate fellowships from the National Science Foundation and the University of California, Santa Cruz.

**Data availability:** Freely available online through the PNAS open access option.

**Conflict of interest statement:** No competing interests declared.
Results and Discussion

Overall Structure of the E. coli σ^5-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^5 factor. The σ^5-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 structures were solved unambiguously in the density map, suggesting a slippery DNA hybrid in the pretranslocation register.

Table 1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Parameters</th>
<th>σ^5-TIC1</th>
<th>σ^5-TIC2 1-h soaking with CTP only</th>
<th>σ^5-TIC3 1-h soaking with CTP, UTP, and GTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2_12_1</td>
<td>P2_12_1</td>
<td>P2_12_1</td>
</tr>
<tr>
<td>Cell dimensions, Å</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>131.71</td>
<td>132.75</td>
<td>132.87</td>
</tr>
<tr>
<td>b</td>
<td>152.67</td>
<td>151.97</td>
<td>152.17</td>
</tr>
<tr>
<td>c</td>
<td>226.70</td>
<td>228.06</td>
<td>229.12</td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>100-3.60 (3.66-3.60)</td>
<td>100-4.20 (4.27-4.20)</td>
<td>100-4.60 (4.68-4.60)</td>
</tr>
<tr>
<td>R_{sym}, %</td>
<td>12.5 (&gt;100)</td>
<td>10.2 (&gt;100)</td>
<td>8.5 (&gt;100)</td>
</tr>
<tr>
<td>I/σ</td>
<td>9.87 (0.37)</td>
<td>8.44 (0.46)</td>
<td>10.06 (0.55)</td>
</tr>
<tr>
<td>Completeness, %</td>
<td>99.6 (99.4)</td>
<td>99.9 (100)</td>
<td>99.3 (99.6)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>5.1 (4.2)</td>
<td>7.7 (7.9)</td>
<td>6.9 (6.8)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>100-3.60</td>
<td>100-4.20</td>
<td>100-4.60</td>
</tr>
<tr>
<td>No. reflections</td>
<td>50,952</td>
<td>32,692</td>
<td>24,772</td>
</tr>
<tr>
<td>R_{work}/R_{ref}</td>
<td>24.7/28.7</td>
<td>26.7/32.7</td>
<td>24.0/33.4</td>
</tr>
<tr>
<td>No. atoms</td>
<td>27,632</td>
<td>29,036</td>
<td>29,036</td>
</tr>
<tr>
<td>Protein</td>
<td>26,167</td>
<td>27,572</td>
<td>27,572</td>
</tr>
<tr>
<td>DNA/RNA/ions</td>
<td>1,465</td>
<td>1,464</td>
<td>1,464</td>
</tr>
<tr>
<td>B factors</td>
<td>129.5</td>
<td>177.4</td>
<td>183.1</td>
</tr>
<tr>
<td>Rmsd</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond length, Å</td>
<td>0.011</td>
<td>0.011</td>
<td>0.008</td>
</tr>
<tr>
<td>Bond angle, °</td>
<td>1.448</td>
<td>1.422</td>
<td>1.179</td>
</tr>
</tbody>
</table>

Values in parentheses are for highest-resolution shell. Each dataset was collected from a single crystal.

σ=2 domain and a helix-turn-helix formation, termed a “clamp helices,” of the clamp domain of the β subunit. Compared with σ^30, which displays strong affinity for the RNAP core enzyme (Kd ~0.26 nM), σ^5 binds relatively weakly to the RNAP core in the absence of nucleic acids (Kd ~4.26 nM) (12). This weaker affinity of σ^5 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. S2B).

E. coli σ^5 Factor and σ^5-RNAP Core Interactions. Like the primary σ factors, the E. coli σ^5 factor contains a highly negatively charged N-terminal domain (σ_1,1) that we have not been able to trace in the σ^5-TIC structures. Other than a nonconserved region (NCR) that is commonly present in primary σ factors, σ^5 displays very high sequence identity with the primary σ factors (Fig. S2A) (10). Consistent with their sequence similarity, the σ^5 factor in the σ^5-TIC forms essentially the same fold as the σ^1 factor throughout the conserved regions, from region 1.2 to the very C terminus, including the long σ_1,2 loop that inserts deep into the RNAP active site chamber (Fig. 2A).

A major contribution to the σ-RNAP core association comes from interactions between the σ_2 domain and a helix-turn-helix motif of the clamp domain of the clamp protein. The σ^5 binds relatively weakly to the RNA polymerase core in the absence of nucleic acids (Kd ~4.26 nM) (12). This weaker affinity of σ^5 might be attributed partly to its lack of an NCR domain that provides additional interactions with the clamp helices of the RNA polymerase core (Fig. S2B).
It was shown previously that the domain and the extensive interactions with positively charged side chains from the general area on the surface of several hydrogen-bonding interactions with the attachment face of the NCR to the primary Asp135, Pro136, and Glu137 (15). This general area corresponds to the specific recognition of the −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 σ70 interactions with the promoter −35 element are not observed in the σ70-TIC crystals, the sequence and structural conservation and the shared recognition of promoter sequences suggest that the σ70 and σ30 factors would interact with the DNA promoters in a very similar manner.

Selective Gene Expression Under Stress Conditions. An intriguing question is how σ70 achieves selective gene expression, given that it recognizes consensus −35 and −10 sequences essentially identical to those recognized by σ30. The σ70-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 σ70 factor. Although the NCR domain in σ70 might make it interact better with the RNAP core enzyme, a stretch of acidic residues (18 of the E. coli σ70 residues 188–209) (Fig. S2B) is expected to inhibit promoter loading directly through interactions of the σ70-holoenzyme with the −10 element (22), thus making interactions with the −35 element important for recruiting the σ70-holoenzyme to the promoter. In contrast, the smaller σ30 domain of the σ70 factor would allow it to interact directly with the promoter −10/extended −10 region, thus making the σ70-holoenzyme less dependent on the −35 element for promoter loading.

The protein concentration of the σ70 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 σ30 protein level in E. coli is negligible, but when E. coli enters the stationary phase or under certain stress conditions the σ30 protein level might increase by a thousand-fold and reach a level comparable to that of σ70. Easier access not only would provide the σ70-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 σ70 protein concentration under normal growth conditions.

Stressed TICs. The σ70-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 σ70-TICs, the σ70-TICs we report here also contain a complete transcription bubble and an RNA oligonucleotide synthesized de novo from NTPs, and the σ70-TIC structures display a pretranslocated RNA-DNA hybrid as well (Fig. 3). Moreover, the active site of the σ70-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 σ70-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 (βH1His936...
sugar ring of the penultimate residue. This finding suggests that the linkage of the terminal phosphate (PPi) displays some discontinuity between the two terminal residues, and the RNA is extended perpendicular to the translocation direction with a shift of about 3 Å (Fig. 3A). The BH stacks against the RNA hybrid and PPi 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 σS-TIC (Fig. S4B), 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. 3A). 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 σS-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 σS-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 be associated 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 σS-TIC active site.

The pyrophosphate in the σS-TIC interacts via a metal ion (MeII) with the carboxyl group of β′Arg462, one of the conserved carboxyl triad of the active center, with the side chains of the TL residues β′Arg333 and β′His936, and with the side chains of β′Arg1106 and β′Arg731 that line as the secondary channel (Fig. 3B). 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).
appear to be much less efficient in RNA synthesis. Soaking σ5-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 σ5-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 σ5-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 σ5-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 McE 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 the σ5-TIC insertion of 188 amino acids could make a large shift around the secondary channel (39) and has been shown to affect the coordination of the phosphate group of the RNA 3′-terminal residue (39). The conformational changes in both the protein and the nucleic acid in the σ5-TIC crystals and active site opening. 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 the coordination of the phosphate group of the RNA 3′-terminal residue (39). The conformational changes in both the protein and the nucleic acid in the σ5-TIC crystals and active site opening. 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).

Materials and Methods
Preparation and Crystallization of E. coli σ5-TIC. To form the σ5-TIC, we used a synthetic DNA scaffold corresponding to the promoter region between positions 38–50 and +12 relative to the expected transcription start site (Fig. 1A). 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. 1A). The σ5-TIC was assembled by directly incubating the σ5-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 MgCl₂ 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% (w/v) 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% (w/v) ethylene glycol before flash-freezing in liquid nitrogen. The σ5-TIC crystallizes in the orthorhombic P2₁2₁2₃ space group with one copy of the complex per asymmetric unit different from that of the σ5-TIC (5).

NTP Soaking for RNA Synthesis in the Crystal. The initially obtained σ5-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 σ5-TIC (5) as the starting model. The molecular replacement solution was subjected to several rounds of 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 σ5 factor and the nucleic acid models into the density using COOT (44). After model building in Coot, 10 cycles of TLS (translation liberation 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.