Structural and mechanistic basis of reiterative transcription initiation

Yu Liu[a,b] , Libing Yu[a,b], Chirangini Pukhrambam[b,c] , Jared T. Winkelman[a,b,c], Emre Firlar[d] , Jason T. Kaelber[d] , Yu Zhang[e] , Bryce E. Nickels[b,c,1] , and Richard H. Ebright[a,b,1]

[a]Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854; [b]Waksman Institute of Microbiology, Rutgers University, Piscataway, NJ 08854; [c]Department of Genetics, Rutgers University, Piscataway, NJ 08854; [d]Rutgers CryoEM and Nanoimaging Facility, Institute for Quantitative Biomedicine, Rutgers University, Piscataway, NJ 08854; and [e]Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China

Reiterative transcription initiation, observed at promoters that contain homopolymeric sequences at the transcription start site, generates RNA products having 5′ sequences noncomplementary to the DNA template. Here, using crystallography and cryoelectron microscopy to define structures, protein–DNA photocrosslinking to map positions of RNAP leading and trailing edges relative to DNA, and single-molecule DNA nanomanipulation to assess RNA polymerase (RNAP)–dependent DNA unwinding, we show that RNA extension in reiterative transcription initiation 1) occurs without DNA scrunching; 2) involves a short, 2- to 3-bp, RNA–DNA hybrid; and 3) generates RNA that exits RNAP through the portal by which scrunched nontemplate-strand DNA exits RNAP in standard transcription initiation. The results establish that, whereas RNA extension in standard transcription initiation proceeds through a scrunching mechanism, RNA extension in reiterative transcription initiation proceeds through a slippage mechanism, with slipping of RNA relative to DNA within a short RNA–DNA hybrid, and with extrusion of RNA from RNAP through an alternative RNA exit.

RNA polymerase | transcription | reiterative transcription initiation | transcriptional slippage | DNA scrunching

In standard transcription initiation, RNA polymerase (RNAP) holoenzyme binds to promoter DNA, unwinds ∼13 base pairs (bp) of promoter DNA to form an RNAP–promoter open complex (RPo) containing a single-stranded "transcription bubble," selects a transcription start site, and synthesizes the first 10 nucleotides (nt) of the RNA product as an RNAP–promoter initial transcribing complex (RPitc) (1–4). Standard transcription initiation proceeds through a "DNA scrunching" mechanism, in which RNAP unwinds additional DNA, pulls the additional unwound DNA past its active center, and accommodates the additional unwound DNA as single-stranded bulges within the transcription bubble (5–8). During standard transcription initiation, each step of RNA extension involves 1) unwinding of 1 bp of DNA downstream of the RNAP active center, expanding the transcription bubble by 1 bp; 2) translocation of DNA and RNA together by 1 bp relative to the RNAP active center; 3) binding, through base pairing, of a complementary nucleoside triphosphate (NTP) to the DNA template strand in the RNAP active center; and 4) phosphodiester bond formation, resulting in addition of a nucleotide to the RNA 3′ end (2). Standard transcription initiation yields an RNA product having a sequence fully complementary to the DNA template strand. Furthermore, during standard transcription initiation, the RNA product remains fully base paired to the DNA template strand as an RNA–DNA "hybrid."

In an alternative pathway of transcription initiation, termed "reiterative transcription initiation," "transcriptional stuttering," or "pseudo-templated transcription," an RNAP–promoter reiteratively transcribing complex (RPrtc) synthesizes an RNA product having a 5′-end sequence that contains a variable number, up to tens to hundreds, of nucleotides not complementary to the DNA

template (9–11). Reiterative transcription initiation, which was first observed six decades ago (12, 13), competes with standard transcription initiation, both in vitro and in vivo (14–24). Reiterative transcription initiation occurs at promoters that contain homopolymeric sequences at, or immediately downstream of, the transcription start site, resulting in low yields of standard, full-length RNA products at such promoters (9–11). The extent of reiterative transcription initiation relative to standard initiation can change in response to changes in NTP concentrations, enabling transcription-factor-independent regulation of gene expression (9, 11, 25). Classic examples of genes regulated through changes in the extent of reiterative transcription initiation relative to standard transcription initiation in response to changes in NTP concentrations are the *Escherichia coli* pyrimidine biosynthetic gene *pyrBI* and the *Bacillus subtilis* pyrimidine biosynthetic gene *pyrG* (14, 18).

The mechanism of reiterative transcription initiation has not been firmly established. It has been hypothesized that reiterative transcription initiation involves an "RNA slipping" mechanism, in which RNA extension does not involve translocation of DNA relative to the RNAP active center, but, instead, involves

---

Significance

In standard transcription initiation, RNA polymerase (RNAP) synthesizes an initial RNA product complementary to the DNA template. In an alternative pathway of transcription initiation, termed "reiterative transcription initiation," observed at promoters containing homopolymeric sequences at the transcription start site, RNAP synthesizes an initial RNA product having a 5′ sequence containing a variable number, up to tens to hundreds, of nucleotides noncomplementary to the DNA template. Here, using crystallography, cryoelectron microscopy, photocrosslinking, and single-molecule DNA nanomanipulation, we show that RNA extension in reiterative transcription initiation involves slipping of RNA relative to DNA within a short RNA–DNA hybrid and generates RNA that exits RNAP through a previously unobserved RNA path ("alternative RNA path") and previously unobserved RNA exit ("alternative RNA exit").

translocation of RNA—"slippage"—relative to both DNA and the RNAP active center (9–11, 17, 19, 21, 24, 26–28). However, direct evidence for RNA slipping has not been presented, and the mechanism by which long reiteratively transcribed RNA products—RNA products tens to hundreds of nucleotides in length—leave the RNAP active center has not been defined.

Crystal structures have been reported of RPrtc (27, 28). However, in those structures, 1) transcription-bubble non–template-strand DNA was extensively disordered (3 to 8 nt disordered), complicating assessment of non–template-strand DNA scrunching; 2) transcription-bubble template-strand DNA was partly missing and partly disordered (5 nt missing; 3 nt disordered), precluding assessment of template-strand DNA scrunching; 3) the position of the RNA 3′-end relative to the DNA template strand did not permit further reiterative transcription (template-strand homopolymeric sequence not aligned with RNAP active-center addition site), precluding assessment of the mechanism of RNA extension; and 4) only complexes with short RNA products (6 to 8 nt) were analyzed, precluding assessment of how long reiterative-transcription–generated RNA products exit the RNAP active-center cleft (27, 28). As a result of these limitations, the previous crystal structures did not enable determination of the roles of DNA scrunching and RNA slipping, the length of the RNA–DNA hybrid, and the RNA-exit path in reiterative transcription initiation.

Here, we report crystal structures of RNAP engaged in standard transcription initiation of short RNA products on a template containing a nonhomopolymeric sequence and of RNAP engaged in reiterative transcription initiation of short RNA products on templates containing template-strand GGG and CCC homopolymeric sequences. In addition, we report a cryoelectron microscopy (cryo-EM) structure of RNAP engaged in reiterative transcription initiation of long—up to at least 50 nt—RNA products on a template containing a template-strand CCC homopolymeric sequence. The structures reveal that, whereas RNA extension in standard transcription initiation involves DNA scrunching, RNA extension in reiterative transcription initiation does not. The structures further reveal that only two template-strand nucleotides (in a post-translocated state) or three template-strand nucleotides (in a pretranslocated state) are positioned to be able to base pair to the RNA product, resulting in a short RNA–DNA hybrid. The cryo-EM structure of RNAP engaged in reiterative transcription initiation of long RNA products further reveals that reiterative-transcription–generated RNA exits RNAP using the path by which scrunched non–template-strand DNA exits RNAP in standard transcription initiation, instead of the path by which RNA exits RNAP in standard transcription initiation. Results of two independent orthogonal approaches, site-specific protein–DNA photocrosslinking and single-molecule DNA nanomanipulation, confirm the observed scrunching patterns. Taken together, our results establish that, whereas RNA extension in standard initiation involves DNA scrunching, RNA extension in reiterative initiation involves RNA slipping, with sliding of the RNA product relative to the DNA template strand within a short RNA–DNA hybrid, and with extrusion of RNA from the RNAP active-center cleft through an alternative RNA exit.

Results

Crystal Structures of RPrtc,4 and RPrtc,5: RNA Extension through RNA Slipping without DNA Scrunching.

We determined crystal structures of *Thermus thermophilus* RPrtc in which non–template-strand DNA is fully ordered, enabling assessment of DNA scrunching, and in which the template-strand homopolymeric sequence is aligned with the RNAP active-center

addition site, enabling assessment of RNA slipping (Figs. 1 and 2). For both the crystal structures of this section and the cryo-EM structure below, we analyzed *T. thermophilus* RNAP because atoms of this hyperthermophilic bacterial RNAP show lower thermal motions at structure-determination temperatures than atoms of mesophilic bacterial RNAP, enabling determination of structures having high order, higher resolution, and superior map quality (27–35). We obtained a crystal structure of RPrtc containing a 5-nt RNA product by incubation of a nucleic-acid scaffold having a $G_{+1}G_{+2}G_{+3}$ template-strand homopolymeric sequence, where +1 is the transcription start site, with RNAP and CTP [RPrtc,5 $(G_{+1}G_{+2}G_{+3})$; Fig. 1*C* and *SI Appendix*, Fig. S1*A* and Table S1], and we obtained a crystal structure of RPrtc containing a 4-nt RNA product by incubation of a nucleic-acid scaffold having a $C_{+1}C_{+2}C_{+3}$ template-strand homopolymeric sequence with RNAP and GTP [RPrtc,4 $(C_{+1}C_{+2}C_{+3})$; Fig. 1*D* and *SI Appendix*, Fig. S1*A* and Table S1]. For reference, we compared these structures to a previously reported crystal structure of *T. thermophilus* RPo (RPo; Fig. 1*A* and ref. 34) and a crystal structure of *T. thermophilus* RPitc containing a 5-nt RNA product obtained by incubation of a nucleic-acid scaffold lacking a template-strand homopolymeric sequence with RNAP, ATP, UTP, and CTP (RPitc,5; Fig. 1*B* and *SI Appendix*, Fig. S1*A* and Table S1).

The crystal structure of RPo shows ordered density for all nucleotides of the transcription-bubble nontemplate strand: 5 nt in the −10 element (a promoter element recognized by the conserved region 2 of transcription initiation factor σ; ref. 36), 4 nt in the discriminator element (another promoter element recognized by conserved region 2 of transcription initiation factor σ; ref. 36), and 4 nt between the discriminator element and downstream double-stranded DNA (dsDNA) (Fig. 1*A* and *SI Appendix*, Fig. S2*A*, blue, light blue, and pink; ref. 34).

The crystal structure of RPitc,5 shows an initially transcribing complex with a 5-nt RNA product and an RNAP active-center post-translocated state (Fig. 1*B*; *SI Appendix*, Fig. S2*B*). The RNA product is fully base paired to the DNA template strand as an RNA–DNA hybrid, with the RNA 3′ nucleotide and the corresponding DNA template-strand nucleotide located in the RNAP active-center product site ("P site"), and the next DNA template-strand nucleotide in the RNAP active-center addition site ("A site") available for base pairing with an incoming NTP. The positions of the RNA and DNA relative to the RNAP active center indicate that, as compared to those in RPo, 4 bp of downstream dsDNA have been unwound, 4 nt of each strand has been translocated relative to the RNAP active center, and the RNA product has been translocated 4 nt in lock-step register with template-strand DNA. The crystal structure of RPitc,5 indicates that the 4 nt of non–template-strand DNA translocated relative to the RNAP active center are accommodated through DNA scrunching, with bulging of the non–template-strand DNA segment between the discriminator element and downstream dsDNA. Thus, the crystal structure of RPitc,5 shows ordered density for all transcription-bubble non–template-strand nucleotides of the −10 element and the upstream half of the discriminator element, with the same positions and the same σ-DNA interactions as in RPo, and shows disorder for 8 nt of non–template-strand DNA, corresponding to the downstream half of the discriminator element and DNA immediately downstream of the discriminator element (Fig. 1*B* and *SI Appendix*, Fig. S2 *B*, cyan boxes). The 8-nt segment of disordered non–template-strand DNA in RPitc,5 has exactly the same endpoints, and spans exactly the same distance, as a 4-nt segment of ordered non–template-strand DNA in RPo (Fig. 1 *A* and *B*; *SI Appendix*, Fig. S2 *A* and *B*), indicating that ~4 nt of the 8-nt segment of disordered non–template-strand DNA are flipped out and/or bulged out relative to the path of the nontemplate strand in RPo (Fig. 1*B* and *SI Appendix*,
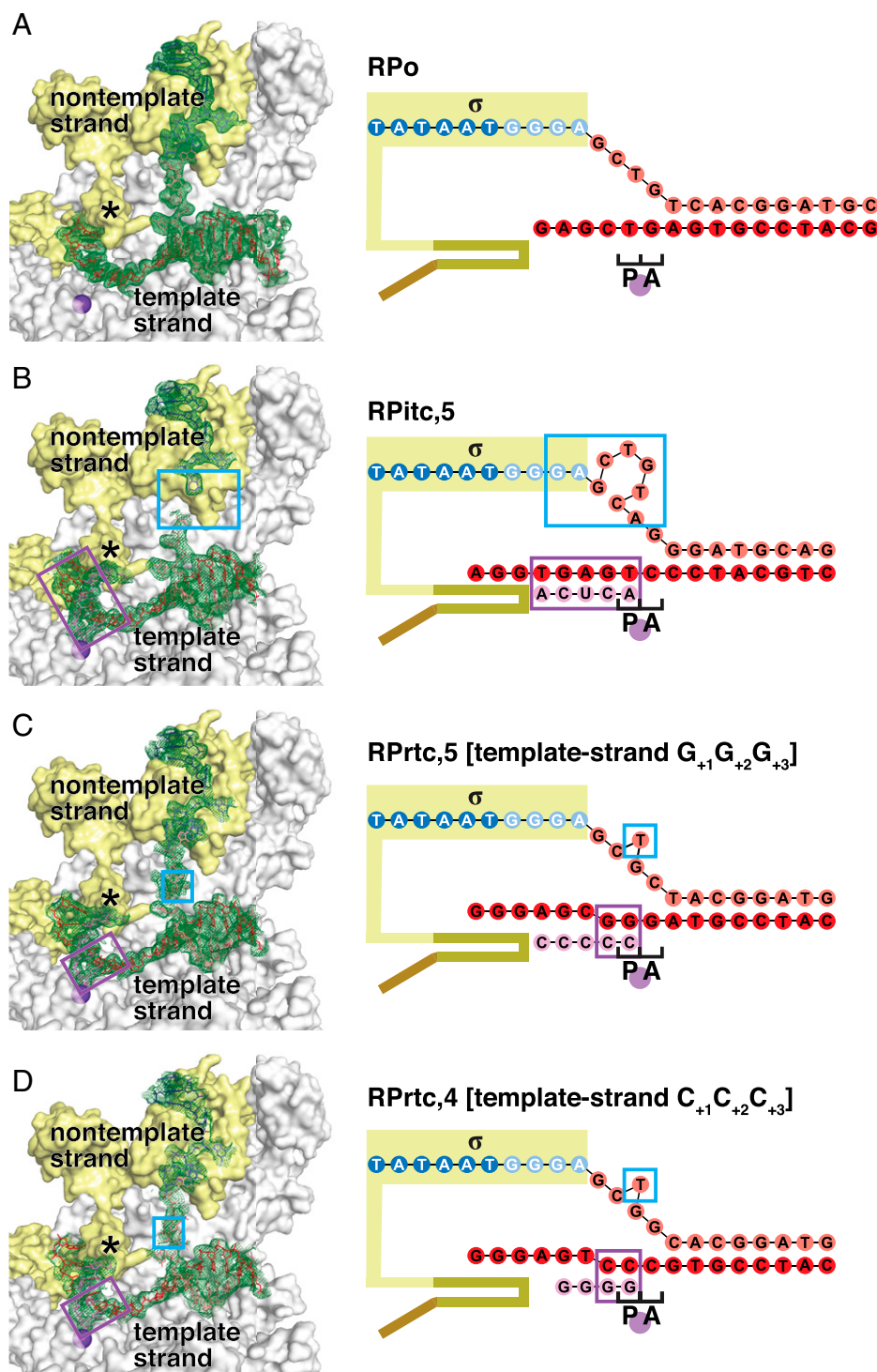
Fig. 1. Crystal structures of RPrtc,4 and RPrtc,5: RNA extension through RNA slipping without DNA scrunching. (*A–D*) Crystal structures of transcription initiation complexes engaged in standard transcription initiation and reiterative transcription initiation. *Left*, experimental electron density (mFo-DFc; contoured at 2.0σ in *A* and 1.5σ in *B–D*) and atomic model, showing interactions of RNAP and σ with transcription-bubble nontemplate strand, transcription-bubble template strand, and downstream dsDNA (RNAP β subunit and β′ nonconserved domain omitted for clarity). *Right*, nucleic-acid scaffold. RNAP, gray; RNAP active-center catalytic Mg^{2+}(I) ion, violet; σ, yellow; σ finger, asterisk in *Left* subpanel and yellow-brown in *Right* subpanel; σR3–σR4 linker in RNA exit channel, brown; −10 element of DNA nontemplate strand, dark blue; discriminator element of DNA nontemplate strand, light blue; rest of DNA nontemplate strand, pink; DNA template strand, red; RNA product, magenta. Cyan rectangles in *B* indicate disordered regions containing scrunched nucleotides. Cyan rectangles in *C* and *D* indicate ordered scrunched nucleotides. Bulged-out nucleotides in *B–D*, *Right*, indicate bulged-out scrunched nucleotides. Violet rectangles indicate RNA–DNA hybrids. Raised template-strand nucleotides in *C* and *D* indicate non–base-paired nucleotides. (*A*) RPo (PDB ID: 4G7H; ref. 34). (*B*) RPitc containing 5-nt RNA product generated by in crystallo standard transcription initiation (RPitc,5). (*C*) RPrtc containing 5-nt RNA product generated by in crystallo reiterative transcription initiation on nucleic-acid scaffold having a template-strand $G_{+1}G_{+2}G_{+3}$ homopolymeric sequence (RPrtc,5 [$G_{+1}G_{+2}G_{+3}$]). (*D*) RPrtc containing 4-nt RNA product generated by in crystallo reiterative transcription initiation on nucleic-acid scaffold having a template-strand $C_{+1}C_{+2}C_{+3}$ homopolymeric sequence (RPrtc,4 [$C_{+1}C_{+2}C_{+3}$]).

Fig. S2 *B*, *Right*, flipped and/or bulged nucleotides in cyan box). The disorder of the 8-nt segment of disordered non–template-strand DNA indicates that the segment adopts an ensemble of distinct conformations (Fig. 1*B* and *SI Appendix*, Fig. S2*B*, cyan boxes). We conclude that formation of RPitc,5 involves ~4 bp of DNA scrunching.

The crystal structure of RPrtc,5 [$G_{+1}G_{+2}G_{+3}$] shows a reiteratively transcribing complex with a 5-nt RNA product and an RNAP active-center post-translocated state (Fig. 1*C*; *SI Appendix*, Fig. S2*C*). Only part of the RNA product—the part comprising the 3′ nucleotide and the adjacent nucleotide—is

base paired to the DNA template strand, as a 2-bp RNA–DNA hybrid, with the RNA 3′ nucleotide and the corresponding DNA template-strand nucleotide located in the RNAP active-center P site and the next DNA template-strand nucleotide in the RNAP active-center A site available for base pairing with an incoming NTP. The positions of the RNA and DNA relative to the RNAP active center indicate that, as compared to those in RPo, 1 bp of downstream dsDNA has been unwound, 1 nt of each strand of DNA has been translocated relative to the RNAP active center, and the 5′ end of the RNA product has been translocated by 4 nt relative to the RNAP active center,
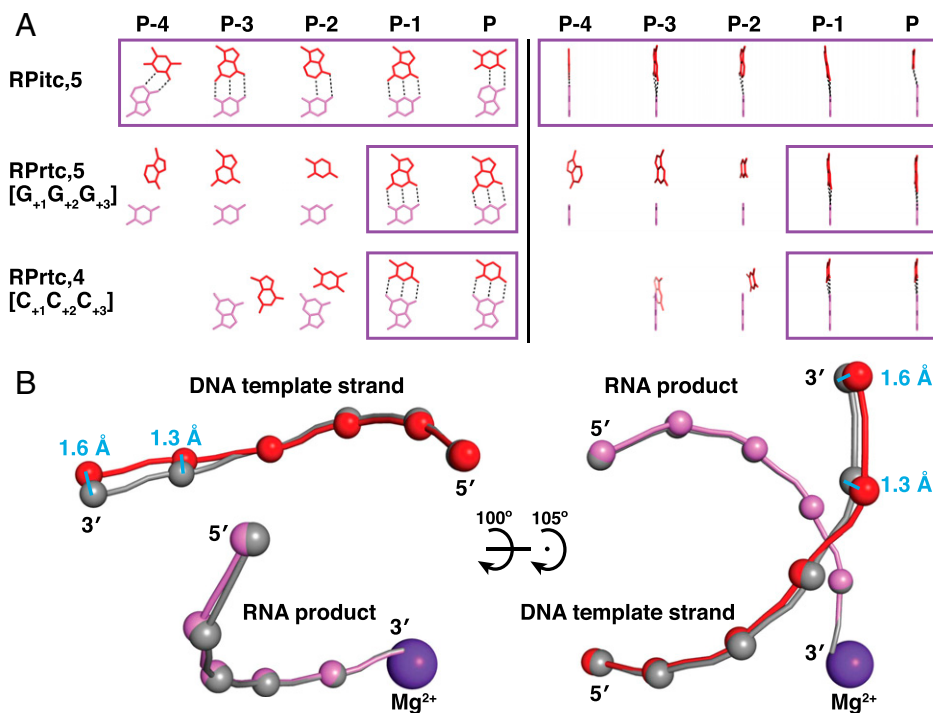
Liu et al.
Structural and mechanistic basis of reiterative transcription initiation

PNAS | 3 of 12
https://doi.org/10.1073/pnas.2115746119

Fig. 2. Crystal structures of RPrtc,4 and RPrtc,5: short RNA–DNA hybrid. (*A*) RNA–DNA base pairing in crystal structures of transcription initiation complexes engaged in standard transcription initiation (RPitc,5) and reiterative transcription initiation (RPrtc,5 [$G_{+1}G_{+2}G_{+3}$] and RPrtc,4 [$C_{+1}C_{+2}C_{+3}$]). *Left*, template-strand DNA bases (red) and corresponding RNA bases (magenta) in view orientation parallel to RNA–DNA hybrid helix axis. *Right*, template-strand DNA bases (red) and corresponding RNA bases (magenta) in view orientation perpendicular to RNA–DNA hybrid helix axis. Positions are numbered relative to the RNAP active-center P site. Dashed lines indicate Watson–Crick H-bonds. Violet rectangles indicate RNA–DNA hybrids. At positions P-4, P-3, and P-2 of RPrtc,5 [$G_{+1}G_{+2}G_{+3}$], and at positions P-3 and P-2 of RPrtc,4 [$C_{+1}C_{+2}C_{+3}$], template-strand DNA bases are displaced relative to their locations in RPitc,5, and no base pairing occurs. (*B*) Superimposition of DNA template strand and RNA of RPrtc,5 [$G_{+1}G_{+2}G_{+3}$] (red spheres, DNA phosphates; magenta spheres, RNA phosphates; violet sphere, RNAP active-center catalytic Mg^{2+} ion) on DNA template strand and RNA of RPitc,5 (gray spheres, DNA and RNA phosphates). *Left*, view orientation parallel to RNA–DNA hybrid helix axis; *Right*, view orientation perpendicular to RNA–DNA hybrid helix axis. Distances in cyan, displacement of template-strand DNA nucleotides at positions P-4 and P-3 of RPrtc,5 [$G_{+1}G_{+2}G_{+3}$] relative to their locations in RPitc,5.

translocating 1 nt in register with template-strand DNA and 3 nt out of register with template-strand DNA. The crystal structure of RPrtc,5 [$G_{+1}G_{+2}G_{+3}$] shows that the 1 nt of non–template-strand DNA that is translocated relative to the RNAP active center is accommodated through DNA scrunching, with unstacking and flipping of 1 nt of non–template-strand DNA between the discriminator element and downstream dsDNA (Fig. 1*C* and *SI Appendix*, Fig. S2*C*, cyan boxes). The crystal structure of RPrtc,5 [$G_{+1}G_{+2}G_{+3}$] shows ordered density for all transcription-bubble non–template-strand nucleotides, including the scrunched, unstacked, flipped nucleotide: 5 nt in the −10 element, 4 nt in the discriminator element, and 5 nt between the discriminator element and downstream dsDNA (Fig. 1*C*; *SI Appendix*, Fig. S2*C*). The crystal structure shows graphically that formation of RPrtc,5 [$G_{+1}G_{+2}G_{+3}$] involves 1 bp of DNA scrunching and 3 nt of RNA slipping.

The crystal structure of RPrtc,4 [$C_{+1}C_{+2}C_{+3}$] shows a reiteratively transcribing complex with a 4-nt RNA product and an RNAP active-center post-translocated state, a 2-bp RNA–DNA hybrid, 1 bp of unwinding of downstream dsDNA, 1 nt of translocation of each DNA strand, 1 nt of translocation of the 5′ end of the RNA product in register with template-strand DNA, and 2 nt of translocation of the 5′ end of the RNA product out of register with template-strand DNA (Fig. 1*D*; *SI Appendix*, Fig. S2*D*). The crystal structure shows graphically that formation of RPrtc,4 [$C_{+1}C_{+2}C_{+3}$] involves 1 bp of DNA scrunching and 2 nt of RNA slipping.

We infer, based on the structures in Fig. 1 and *SI Appendix*, Fig. S2, that standard transcription initiation to generate a post-translocated state of RPitc,x involves x − 1 bp of DNA scrunching, whereas reiterative transcription initiation to form a post-translocated state of RPrtc,x involves 1 bp of DNA scrunching and x − 2 nt of RNA slipping. Expressing these inferences in terms of mechanism, we infer that, in standard transcription initiation, following synthesis of a 2-nt initial RNA product, 1 bp of DNA scrunching occurs for each 1 nt of RNA extension (7), and we infer that, in contrast, in reiterative transcription initiation, following synthesis of a 2-nt initial RNA product, 1 bp of DNA scrunching occurs to position the 3′ end of the initial RNA product in the RNAP active-center P site and no further DNA scrunching—just RNA slipping—occurs in RNA extension (see *Discussion*).

Crystal Structures of RPrtc,4 and RPrtc,5: Short RNA–DNA Hybrid. In the crystal structure of RPitc,5, all nucleotides of the RNA product are complementary and base paired to the DNA template strand, yielding a 5-bp RNA–DNA hybrid (Figs. 1*B* and 2*A*). In contrast, in the crystal structures of RPrtc,5 [$G_{+1}G_{+2}G_{+3}$] and RPrtc,4 [$C_{+1}C_{+2}C_{+3}$], only the 2 nt at the 3′ end of the RNA product are complementary and base paired to the DNA template strand, yielding a 2-bp RNA–DNA hybrid (Figs. 1 *C* and *D* and 2*A*). We conclude that, unlike standard transcription initiation, reiterative transcription initiation involves a short RNA–DNA hybrid: a hybrid that is only 2 bp in length in the RNAP active-center post-translocated state of the crystal structures ("postslipped state") and that would be only 3 bp in length upon NTP binding and phosphodiester bond formation to yield the RNAP active-center pretranslocated state ("preslipped state").

In RPrtc, the conformation and interactions with RNAP of the RNA product—even the part of the RNA product that is not complementary to and not base paired to template-strand DNA—are the same as in RPitc (Fig. 2*B*). In contrast, in RPrtc, the conformation and interactions of the part of the DNA template strand that is not complementary to and not base paired to RNA differ from those in RPitc (Fig. 2*B*). Inspection of structures of RPrtc and RPitc indicates that RNAP has numerous interactions with RNA, and few interactions with DNA, in the RNAP hybrid binding region, accounting for the observation that RNA conformation, rather than DNA conformation, is maintained upon loss of RNA–DNA complementarity and base pairing.

In our crystal structures of RPitc and RPrtc, the 5′ end of the RNA product is in contact with the "σ finger" (also referred to as σ region 3.2), which enters the RNAP active-center cleft and obstructs the path of the RNA product (Fig. 1; see refs. 2, 3, 32–35, 37, and 38). In solution and in some crystal forms under some conditions, extension of the RNA product beyond a length of 5 nt can drive stepwise displacement of the σ finger (37). In contrast, with the crystal forms and conditions of this work, RNA extension does not drive stepwise displacement of the σ finger—presumably because of crystal-lattice constraints on conformational change in σ—and thus, RNA products are limited to a length of 5 nt.

Cryo-EM Structure of RPrtc,≥11: RNA Extension through RNA Slipping without DNA Scrunching. Reiterative transcription initiation can generate RNA products up to at least 50 nt in length (*SI Appendix*, Fig. S1; refs. 9, 11). Because the volume of the RNAP active-center cleft in RPrtc cannot accommodate more than ~10 nt of RNA, long RNA products generated by reiterative transcription initiation must exit from and extend outside the RNAP active-center cleft (28). A key unresolved question is where long RNA products generated in reiterative transcription initiation exit the RNAP active-center cleft. In the crystal structures of this work, as in the crystal structures of refs. 27 and 28, RNA products generated by reiterative transcription initiation were limited in length because further RNA extension was blocked by the presence of the σ finger in the RNAP active-center cleft and by crystal-lattice constraints that prevented displacement of the σ finger from the RNAP active-center cleft (35, 37), opening of the RNAP clamp (39–41), or any other conformational change that could open a path for further extension of RNA and for extrusion of RNA from the RNAP active-center cleft. One hypothesis is that, in solution, complete displacement of the σ finger from the RNAP active-center cleft channel could allow long RNA products generated in reiterative transcription initiation to exit the RNAP active-center cleft through the RNAP RNA-exit channel, the same exit route used by RNA in standard transcription (28). Another hypothesis is that, in solution, a smaller conformational change in σ and/or RNAP could allow long RNA products generated in reiterative transcription initiation to exit the RNAP active-center cleft through a different route (28). Consistent with the first hypothesis, our crystal structures show that the 3′ region of RNA products generated in reiterative transcription can follow the same path relative to RNAP as in standard transcription initiation (Fig. 2). However, arguing against the first hypothesis, it is unclear how, with only 1 bp of DNA scrunching (Fig. 1), the system could acquire the energy needed to drive complete displacement of the σ finger from the RNAP hybrid binding site and displacement of the σ region-3/region-4 linker from the RNAP RNA-exit channel (free energy that, in standard transcription initiation, is thought to be provided by ~8 to 10 bp of DNA scrunching; refs. 7 and 8), and it is unclear how the σ region-3/region-4 linker could be completely displaced from the RNAP RNA-exit channel without triggering promoter escape (which, in standard transcription initiation, is thought to be triggered by displacement of the region-3/region-4 linker; refs. 32, 33, 37, and 38).

To resolve these questions, we performed cryo-EM structure determination, analyzing *T. thermophilus* reiteratively transcribing complexes prepared in solution. We incubated a nucleic-acid scaffold containing a full transcription bubble and a $C_{+1}C_{+2}C_{+3}$ template-strand homopolymeric sequence with RNAP holoenzyme and GTP, and we applied samples to glow-discharged graphene-oxide-coated grids (42, 43), flash-froze samples, and performed single-particle–reconstruction cryo-EM [RPrtc,≥11 ($C_{+1}C_{+2}C_{+3}$); Fig. 3 and *SI Appendix*, Figs. S1A and S3–S5]. This approach avoided the limitations imposed by crystal-lattice

constraints (Fig. 3 versus Figs. 1 and 2 and refs. 27 and 28). In addition, by employing a nucleic-acid scaffold that contained a full transcription bubble, this approach avoided possible limitations imposed by use of a nucleic-acid scaffold that lacked an upstream duplex (Fig. 3 versus Figs. 1 and 2 and refs. 27 and 28). Use of glow-discharged graphene-oxide-coated grids was essential in order to obtain a satisfactory distribution of particle orientations on grids (*SI Appendix*, Fig. S3D).

The cryo-EM structure of RPrtc,≥11 [$C_{+1}C_{+2}C_{+3}$] has an overall resolution of 3.0 Å, with higher local resolution for regions of interest, including the transcription-bubble nontemplate and template DNA strands and the RNA product (*SI Appendix*, Fig. S3 D and E). Map quality is high, with ordered, traceable density for 7 bp of upstream dsDNA, all nucleotides of the nontemplate and template strands of the transcription bubble, 10 bp of downstream dsDNA, and 11 nt of the RNA product corresponding to the 11-nt segment containing the RNA 3′ end (Fig. 3; *SI Appendix*, Figs. S3E and S5).

The cryo-EM structure of RPrtc,≥11 [$C_{+1}C_{+2}C_{+3}$] shows a reiteratively transcribing complex with a long, ≥11 nt RNA product and an RNAP active-center post-translocated state (Fig. 3 A and B; *SI Appendix*, Fig. S4). As observed in our crystal structures of RPrtc with short RNA products (Fig. 2), the 2 nt at the 3′ end of the RNA product are base paired to the DNA template strand and the next 3 nt of the RNA product are close to, but not Watson-Crick base paired with, the DNA template strand (Fig. 3 A and B). The next 6 nt of the RNA product follow a previously unobserved path that diverges from the DNA template strand because of a collision with the σ finger, crosses the transcription bubble, crosses the DNA nontemplate strand, and exits the RNAP active-center cleft at a position on the face of RNAP opposite the standard RNA exit ("alternative RNA exit"; Fig. 3 A and B and *SI Appendix*, Fig. S4B). Additional nucleotides of the long (tens to hundreds of nucleotides) RNA products generated by reiterative transcription under these conditions (*SI Appendix*, Fig. S1A) would be located outside the alternative RNA exit and would extend into bulk solvent; these nucleotides are not observed in the structure, presumably because of segmental disorder (Fig. 3 A and B), analogous to the segmental disorder previously observed for nucleotides of long RNA products located outside the RNAP standard RNA exit (44). The structure contains σ and exhibits the same σ-DNA and σ-RNAP interactions, except for those made by the σ-finger tip (see below), as in RPo (Fig. 3 A and B), indicating that production of long RNAs by reiterative transcription initiation does not involve substantial disruption of σ-DNA and σ-RNAP interactions and does not involve promoter escape.

Analogously in our crystal structures of RPrtc with short RNA products (Fig. 1 C and D), the positions of the RNA and DNA relative to the RNAP active center indicate that, as compared to those in RPo, 1 bp of downstream dsDNA has been unwound, 1 nt of each strand of DNA has been translocated relative to the RNAP active center, and the 5′ end of the RNA product has been translocated by ≥11 nt relative to the RNAP active center, translocating 1 nt in register with template-strand DNA and ≥10 nt out of register with template-strand DNA. The cryo-EM structure of RPrtc,≥11 [$C_{+1}C_{+2}C_{+3}$] shows that the 1 nt of non–template-strand DNA that is translocated relative to the RNAP active center is accommodated through 1 bp of DNA scrunching, with unstacking and flipping of non–template-strand position +1 (i.e., the position 7 bp downstream of the −10 element), and with changes in conformation of the nucleotides flanking non–template-strand position +1 (Fig. 3 B and C and *SI Appendix*, Fig. S4 C, cyan boxes). The cryo-EM structure of RPrtc,≥11 [$C_{+1}C_{+2}C_{+3}$] further shows that the 1 nt of template-strand DNA that is translocated relative to the RNAP active center likewise is accommodated through 1 bp of DNA scrunching, with unstacking and flipping of template-strand position −9 (i.e., the fourth position within the −10 element), and with small
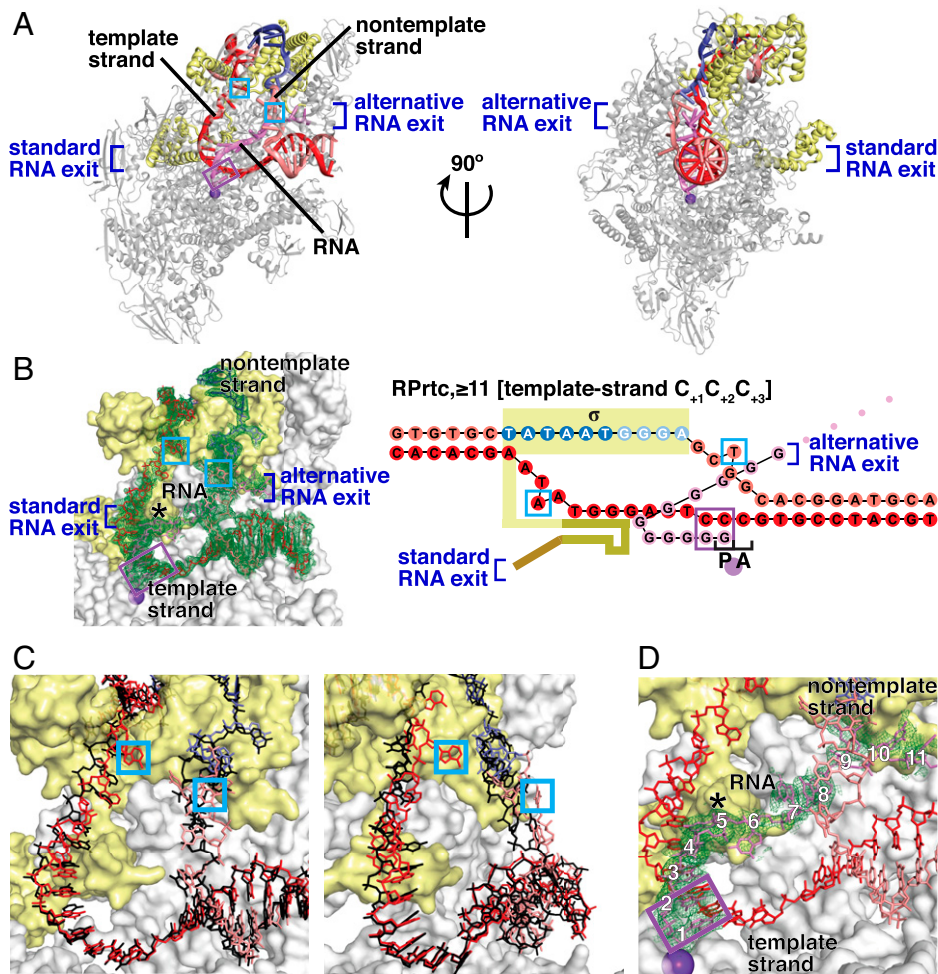
Liu et al.
Structural and mechanistic basis of reiterative transcription initiation

PNAS | 5 of 12
https://doi.org/10.1073/pnas.2115746119

Fig. 3. Cryo-EM structure of RPrtc,≥11: RNA extension through RNA slipping without DNA scrunching. (*A*) Overall structure (β' nonconserved region omitted for clarity; two orthogonal view orientations). Dark blue brackets indicate the standard RNA exit and alternative RNA exit. Cyan rectangles indicate scrunched nucleotides. Violet rectangles indicate RNA–DNA hybrids. Other symbols and colors in panels *A* to *D* are as in Fig. 1. (*B*) *Left*, cryo-EM density and atomic model, showing interactions of RNAP and σ with transcription-bubble nontemplate strand, transcription-bubble template strand, and downstream dsDNA. *Right*, nucleic-acid scaffold. Yellow-brown, σ finger (note displacement of σ-finger tip); magenta dots, RNA outside RNAP active-center cleft (nucleotides rN ≥ 11). (*C*) Superimposition of DNA in RPrtc,≥11 [C_{+1}C_{+2}C_{+3}] (pink and red) on DNA in RPo (black; PDB ID: 512D; ref. 30) (two view orientations). (*D*) Close-up of cryo-EM density and atomic model for RNA (nucleotides rN1 to rN11 numbered in white).

changes in the conformation of template-strand nucleotides downstream of position −9 (Fig. 3 *B* and *C* and *SI Appendix*, Fig. S4*D*, cyan boxes). The cryo-EM structure shows graphically that formation of RPrtc,≥11 [C_{+1}C_{+2}C_{+3}] involves 1 bp of DNA scrunching and ≥10 nt of RNA slipping.

Cryo-EM Structure of RPrtc,≥11: Short RNA–DNA Hybrid. As observed in our crystal structures of RPrtc with short RNA products (Fig. 2*A*), in the cryo-EM structure of RPrtc,≥11 [C_{+1}C_{+2}C_{+3}], only the 2 nt at the 3′ end of the RNA product are complementary and base paired to the DNA template strand, yielding a 2-bp RNA–DNA hybrid (Figs. 3 *A* and *B*; *SI Appendix*, Fig. S5*A*). The structure thus supports our conclusion that, unlike standard transcription initiation, reiterative transcription initiation involves a short RNA–DNA hybrid: a 2-bp hybrid in the post-slipped state and a 3-bp hybrid in the preslipped state.

In our crystal structures of RPrtc with short RNA products, the conformation of the RNA nucleotides at positions P-4, P-3, and P-2 relative to the active-center P site—the RNA nucleotides not complementary to and not Watson-Crick base paired to template-strand DNA—is the same as in a standard transcription initiation complex, and the conformation of the corresponding part of template-strand DNA is different (Fig. 2*B*). In contrast, in the cryo-EM structure of RPrtc,≥11 [C_{+1}C_{+2}C_{+3}], the opposite is true: the conformation of the RNA nucleotides at positions P-4, P-3, and P-2 relative to the active-center P site differs from the conformation in standard transcription initiation and elongation complexes, and the conformation of the corresponding part of template-strand DNA is the same as in standard transcription initiation and elongation complexes (*SI Appendix*, Fig. S5*B*). The conformations of the

RNA product and the DNA template strand at positions P-4, P-3, and P-2 in RPrtc,≥11 [C_{+1}C_{+2}C_{+3}] allow the formation of non–Watson-Crick, wobble, or wobble-like H-bonds at positions P-3 and P-2 (*SI Appendix*, Fig. S5*A*). The difference in the conformations of RNA and DNA at positions P-4, P-3, and P-2 in our crystal structures of RPrtc with short RNA products and our cryo-EM structure of RPrtc,≥11 [C_{+1}C_{+2}C_{+3}] likely is attributable to the absence in the former, and the presence in the latter, of additional RNA nucleotides 5′ to this RNA segment and additional template-strand DNA nucleotides in the full-transcription–bubble nucleic-acid scaffold (Figs. 1 *C* and *D* and 3*B*).

Cryo-EM Structure of RPrtc,≥11: RNA Exit through Nontemplate-Strand Scrunching Portal. In the cryo-EM structure of RPrtc,≥11 [C_{+1}C_{+2}C_{+3}], the RNA segment 6 to 11 nt from the RNA 3′ end (nucleotides rN6 to rN11) follows a path that differs by ~130° from the path of the RNA product in standard transcription initiation and elongation complexes and that differs by ~30° from the paths of the shorter RNA products in the crystal structures of RPrtc in refs. 27 and 28 (Fig. 3 *A*, *B*, and *D*; *SI Appendix*, Fig. S5*C*). This RNA segment, nucleotides rN6 to rN11, diverges from the path of the RNA product in standard transcription initiation and elongation complexes because of a collision with the σ-finger tip involving nucleotide rN6 (Figs. 3 *B* and *D*; *SI Appendix*, Fig. S5 *C*, *Left*). This RNA segment, nucleotides rN6 to rN11, then crosses the transcription bubble spanning ~20 Å, crosses the DNA nontemplate strand spanning an additional ~10 Å, and exits from the RNAP active-center cleft at a position on the face of RNAP opposite the RNA exit used in

standard transcription elongation complexes (Figs. 3 *B* and *D*; *SI Appendix*, Fig. S5 *C*, *Center* and *Right*).

Three residues of the σ-finger tip make van der Waals interactions with nucleotide rN6, causing the path of nucleotides rN6 to rN11 to diverge by ~130° from the path of the RNA product in a standard transcription initiation or elongation complex (σ residues D323, E324, and D326; residues numbered here and below as in *T. thermophilus* RNAP holoenzyme; *SI Appendix*, Fig. 5 *C*, *Left*). Four residues of the RNAP β subunit and one residue of σ make H-bonded or salt-bridged interactions with the sugar–phosphate backbone of nucleotides rN6 to rN11 (β residues K188, R189, T419, and R420, and σ residue K325; *SI Appendix*, Fig. S5*C*). Two residues of the RNAP β subunit and one residue of σ make single H-bonds with RNA bases of nucleotides rN6 to rN11 (β residues N187 and G417, and σ residue T77; *SI Appendix*, Fig. S5*C*). The observations that most protein–RNA interactions with nucleotides rN6 to rN11 involve the sugar–phosphate backbone and that interactions with bases involve single H-bonds that, with wobble, could be made with any base suggest that the RNA-exit pathway observed in this structure may be compatible with any RNA sequence. Residues of the RNAP β subunit that make protein–RNA interactions with nucleotides rN6 to rN11 are residues located immediately N-terminal to β conserved region βa5 (β residues 187 to 189) and residues located in β conserved region βa7, also known as fork-loop 2 (β residues 417 to 420) (*SI Appendix*, Fig. S6, *Left*; β conserved regions defined as in ref. 45). All six residues of the RNAP β subunit that interact with nucleotides rN6 to rN11 are invariant or highly conserved across gram-negative, gram-positive, and *Thermus–Deinococcus*-clade bacteria (*SI Appendix*, Fig. S6, *Left*). Residues of σ that make protein–RNA interactions with nucleotides rN6 to rN11 are residues at the N terminus of σR1.2 (σ residue 77) and residues of the part of σR3.2 that forms the σ-finger tip (σ residues 323 to 326) (*SI Appendix*, Fig. S6, *Right*; σ conserved regions defined as in ref. 36). Three of five residues of σ that interact with nucleotides rN6 to rN11 are invariant or highly conserved across gram-negative, gram-positive, and *Thermus–Deinococcus*-clade bacteria (*SI Appendix*, Fig. S6, *Right*). The observation that residues of RNAP β and σ that make protein–RNA interactions with nucleotides rN6 to rN11 are highly conserved across bacterial species suggests that the RNA-exit pathway observed in this structure may mediate the production of long RNA products by reiterative transcription initiation across bacterial species.

At the point where nucleotides rN6 to rN11 cross the DNA nontemplate strand, direct DNA–RNA interactions occur, involving the stacking of the base of the nucleotide at non–template-strand position +2 (8 nt downstream of the −10 element) on the base of nucleotide rN8 (Figs. 3 *B* and *D*; *SI Appendix*, Fig. S5 *C*, *Center*). This DNA–RNA base-stacking interaction is facilitated by the 1 nt of scrunching of the nontemplate strand, which, by unstacking and flipping out the nucleotide at non–template-strand position +1, frees the nucleotide at non–template-strand position +2 for DNA–RNA base stacking (Figs. 3 *B* and *D*; *SI Appendix*, Fig. S5 *C*, *Center*). The ability of nucleotides rN6 to rN11 to cross the DNA nontemplate strand is further facilitated by the 1 nt of scrunching of the nontemplate strand in that the unstacking and flipping of the nucleotide at nontemplate strand position +1 opens space for, and removes a steric barrier to, the passage of nucleotides rN8 and rN9 past the DNA nontemplate strand.

Nucleotides rN8 to rN10 are accommodated within the same cavity within the RNAP active-center cleft that accommodates scrunched non–template-strand DNA in standard transcription initiation (Fig. 3 *A*, *B*, and *D*; refs. 5, 8, 35, and 46). Nucleotide rN11 exits the RNAP active-center cleft and interacts with bulk solvent. Nucleotide rN11 exits the RNAP active-center cleft

through the same opening through which long, ≥6 to 8 nt, segments of scrunched non–template-strand DNA exit the RNAP active-center cleft (alternative RNA exit in Fig. 3 *A* and *B*; non–template-strand scrunching portal in refs. 5, 8, and 46). We infer that production of long RNAs by reiterative transcription initiation exploits the same cavity within the RNAP active-center cleft (providing an alternative RNA path) and the same opening from the RNAP active-center cleft (providing an alternative RNA exit) that exist to accommodate and extrude scrunched non–template-strand DNA in standard transcription initiation.

Two changes in RNAP conformation present in the structure obtained following reiterative transcription initiation in solution (Fig. 3; *SI Appendix*, Fig. S5), but not in structures obtained following reiterative transcription initiation in crystallo (Figs. 1 and 2; refs. 27 and 28), account for the ability to produce long RNAs in the former, but not in the latter. First, in the structure of RPrtc,≥11 [$C_{+1}C_{+2}C_{+3}$], the tip of the σ finger folds back on itself, moving ~5 Å farther away from the RNAP active center (Fig. 3 *B*, *Right*; *SI Appendix*, Fig. S5*D*), in a manner similar to, but less marked than, the folding back of the tip of the σ finger driven by collision with the RNA 5′ end that occurs in standard transcription initiation upon extension of the RNA product to a length of 6 nt (*SI Appendix*, Fig. S5*D*; ref. 37). This change in local conformation of the σ-finger tip enables reiteratively transcribed RNA to enter into the alternative RNA pathway and to be extended beyond a length of 5 nt. Second, in the structure of RPrtc,≥11 [$C_{+1}C_{+2}C_{+3}$], the RNAP clamp (39–41) opens by ~3°, increasing the width of the RNAP active-center cleft by ~2 Å (*SI Appendix*, Fig. S5*E*). This change in RNAP clamp conformation enables reiteratively transcribed RNA to cross non–template-strand DNA in order to access the alternative RNA exit and leave the RNAP active-center cleft. In our crystal structures of complexes obtained following reiterative transcription initiation in crystallo, neither of these two conformational changes occurred, and RNA products therefore were limited to lengths of 4 to 5 nt (Fig. 1 *C* and *D*). In other crystal structures obtained following reiterative transcription initiation in crystallo, using different sequences and different conditions, the first, but not the second, of these two conformational changes occurred, and RNA products therefore were limited to lengths of 6 to 8 nt (27, 28).

Mapping of RNAP Leading-Edge and Trailing-Edge Positions in RPrtc: RNA Extension without DNA Scrunching. DNA scrunching by RNAP has two biochemically detectable hallmarks: 1) downstream movement of the RNAP leading edge, but not the RNAP trailing edge, relative to DNA (8, 47–50), and 2) expansion of the transcription bubble (7, 47). In a preceding section, we proposed that RNA extension in reiterative transcription initiation does not involve DNA scrunching, except for 1 bp of DNA scrunching to position the 3′ end of the initial RNA product in the RNAP active-center P site. As a first approach to test this proposal, we assessed positions of the RNAP leading edge and trailing edge relative to DNA during reiterative transcription initiation in solution by *E. coli* RNAP holoenzyme (Fig. 4).

We used unnatural amino acid mutagenesis to incorporate the photoactivatable amino acid p-benzoyl-L-phenylalanine (Bpa) at the RNAP leading edge and trailing edge, and we used protein–DNA photocrosslinking to define positions of the RNAP leading edge and trailing edge relative to DNA (methods essentially as in refs. 47 and 48). We analyzed reiterative transcription initiation at derivatives of the bacteriophage N25 promoter containing either a template-strand $G_{+1}G_{+2}G_{+3}$ homopolymeric sequence (RPo [$G_{+1}G_{+2}G_{+3}$] and RPrtc [$G_{+1}G_{+2}G_{+3}$]) or a template-strand $C_{+1}C_{+2}C_{+3}$ homopolymeric sequence (RPo [$C_{+1}C_{+2}C_{+3}$] and RPrtc [$C_{+1}C_{+2}C_{+3}$]). For reference, we analyzed standard transcription initiation at the wild-type N25 promoter (RPo WT and RPitc WT). In vitro transcription experiments, carried out using the same reaction conditions, show

Liu et al.
Structural and mechanistic basis of reiterative transcription initiation

PNAS | **7 of 12**
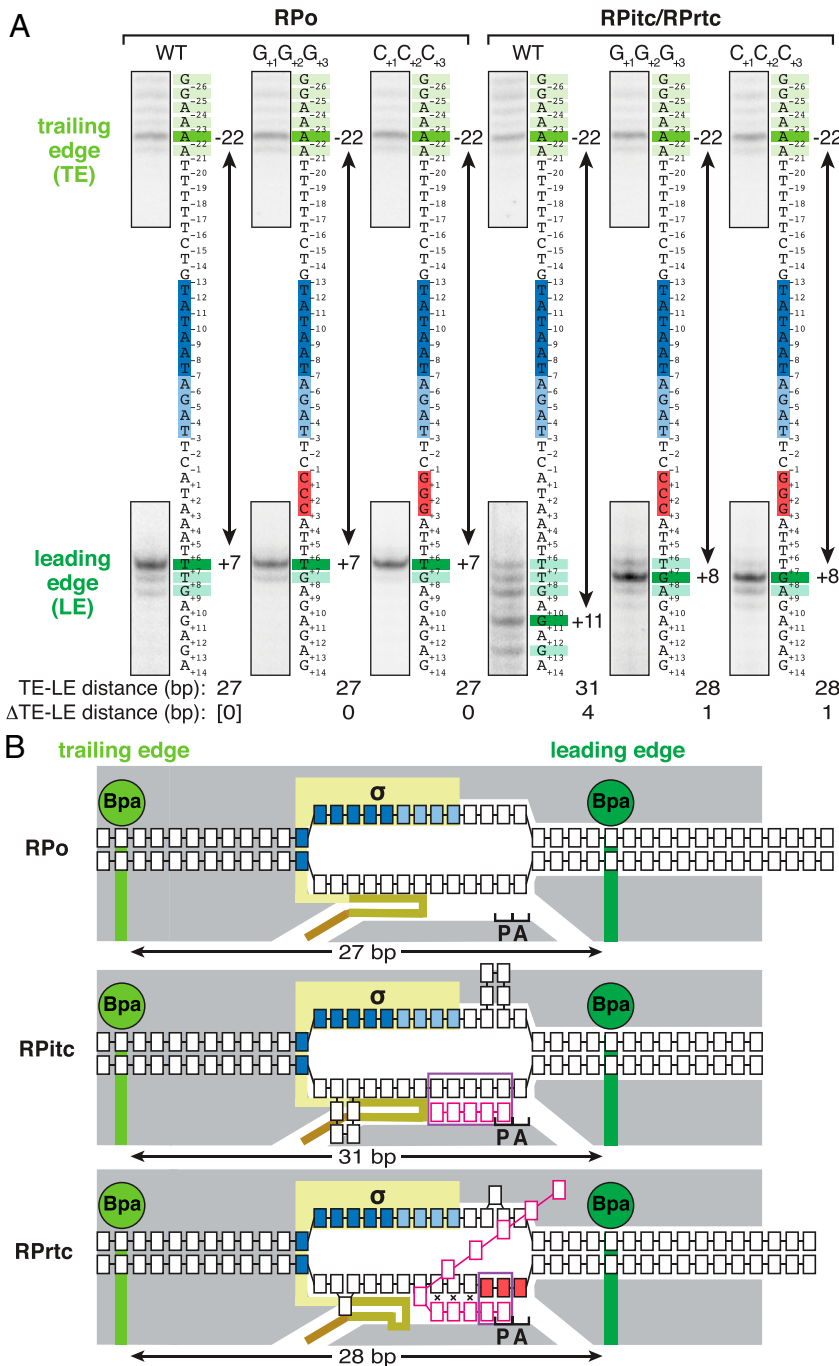https://doi.org/10.1073/pnas.2115746119

Fig. 4. Mapping of RNAP leading-edge and trailing-edge positions in RPrtc by use of protein–DNA photocrosslinking: RNA extension without DNA scrunching. (*A*) RNAP trailing-edge and leading-edge positions in transcription initiation complexes at the N25 promoter (WT) and derivatives of the N25 promoter containing template-strand $G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$ homopolymeric sequences ($G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$). First bracketed subpanel, protein–DNA photocrosslinking data for RPo; second bracketed subpanel, protein–DNA photocrosslinking data for transcription initiation complexes engaged in standard transcription initiation (RPitc) and reiterative transcription initiation (RPrtc). Promoter sequences are shown with positions numbered relative to the transcription start site and with positions of the −10 element, the discriminator element, and the homopolymeric sequence highlighted in blue, light blue, and red. Dark and light olive-green bars indicate strong and weak RNAP trailing-edge crosslinks, and dark and light forest green bars indicate strong and weak RNAP leading-edge crosslinks. *Bottom*, observed modal TE-LE distances and differences in modal TE-LE distance relative to modal TE-LE distance in RPo at WT N25 promoter [Δ(TE-LE distance)]. (*B*) Mechanistic interpretation of data in *A*. Three states are shown: RPo, RPitc [specifically, RPitc having a 5-nt RNA product in a posttranslocated state (RPitc,5 post), corresponding to the major crosslink in *A*], and RPrtc. Gray, RNAP; yellow, σ; yellow-brown, σ finger (note displacement of σ-finger tip in RPrtc); brown, σ region-3/region-4 linker; light green, trailing-edge Bpa and crosslinking site for trailing-edge Bpa; dark green, leading-edge Bpa and crosslinking site for leading-edge Bpa; black boxes with blue fill, −10 element nucleotides; black boxes with light blue fill, discriminator-element nucleotides; black boxes with red fill, template-strand homopolymeric-sequence nucleotides; other black boxes, other DNA nucleotides (nontemplate-strand nucleotides above template-strand nucleotides); magenta boxes, RNA nucleotides; violet rectangles, RNA–DNA hybrids; P and A, RNAP active-center product and addition sites. Raised template-strand nucleotides and black x's indicate non–base-paired nucleotides. Scrunching of nontemplate and template DNA strands is indicated by bulged-out nucleotides. Initial product formation in both standard transcription initiation and reiterative transcription initiation involves one step of DNA scrunching. RNA extension in standard transcription initiation involves additional DNA scrunching, but RNA extension in reiterative transcription does not.

production of long—up to at least 50 nt—reiterative transcripts with the $G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$ promoters in the presence of CTP and GTP, respectively, and show production of short (up to ~8 nt) standard transcripts with the WT promoter in the presence of ATP and UTP (*SI Appendix,* Fig. S1; refs. 7 and 51).

The photocrosslinking results indicate that RPrtc exhibits an RNAP trailing-edge position that is unchanged as compared to RPo, exhibits an RNAP leading-edge position that is shifted downstream by 1 bp as compared to RPo, and thus exhibits an RNAP trailing-edge/leading-edge distance (TE-LE distance) that is increased by 1 bp as compared to RPo (Fig. 4*A*). In contrast, RPitc exhibits an RNAP trailing-edge position that is unchanged as compared to RPo, exhibits an RNAP leading-

edge position that is shifted downstream by up to 6 bp (range = 1 to 6 bp; mode = 4 bp) as compared to RPo, and thus exhibits an RNAP TE-LE distance that is increased by up to 6 bp as compared to RPo (range of increase = 1 to 6 bp; modal increase = 4 bp; Fig. 4*A*). An increase in RNAP TE-LE distance is a defining hallmark of DNA scrunching (8, 47, 48). Thus, the results indicate that RPrtc engaged in synthesis of RNA products up to at least 50 nt in length exhibits only 1 bp of DNA scrunching, whereas RPitc engaged in synthesis of RNA up to 8 nt in length exhibits up to 6 bp of DNA scrunching. We conclude, consistent with the crystal and cryo-EM structures of the preceding sections, that in reiterative transcription—following synthesis of a 2-nt initial RNA product and 1 bp of DNA scrunching to
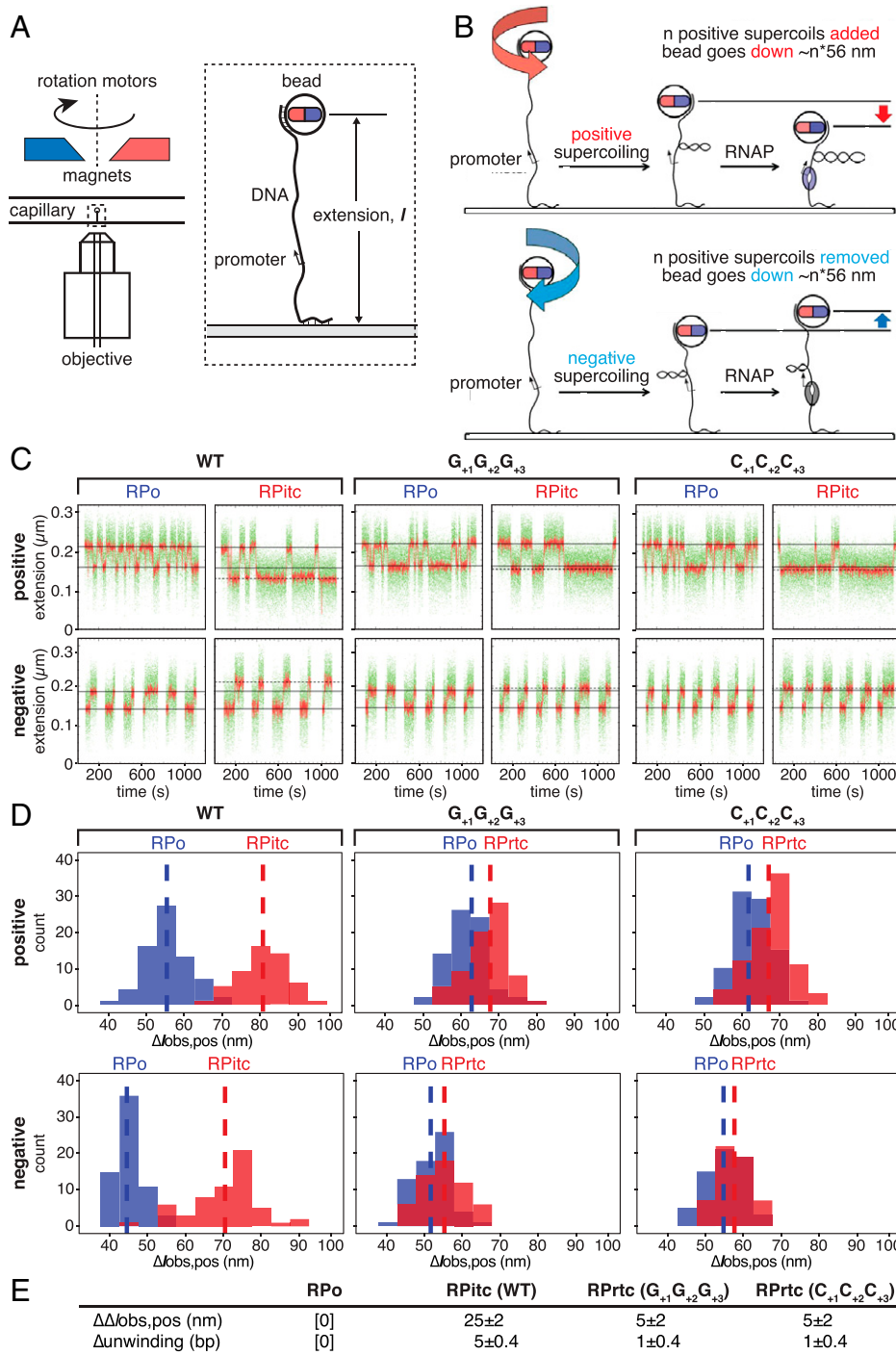
Fig. 5. Measurement of transcription-bubble size in RPrtc by use of single-molecule DNA nanomanipulation: RNA extension without DNA scrunching. (*A*, *B*) Experimental approach (7, 47, 56). (*A*) Apparatus. (*B*) End-to-end extension (*l*) of a mechanically stretched, positively supercoiled (*Top*) or negatively supercoiled (*Bottom*) DNA molecule is monitored. Unwinding of n turns of DNA by RNAP results in compensatory gain of n positive supercoils or loss of n negative supercoils and movement of the bead by n*56 nm. (*C*) Single-molecule time traces for RPo and RPitc at the N25 promoter (WT; *Left*), and for RPo and RPrtc at derivatives of the N25 promoter containing template-strand $G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$ homopolymeric sequences ($G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$; *Middle* and *Right*). Upper subpanels, positively supercoiled DNA; lower subpanels, negatively supercoiled DNA. Green points, raw data (30 frames/s); red points, averaged data (1 s window); horizontal black lines, unbound and RPo states; dashed horizontal black lines, RPitc and RPrtc states (with the difference in Δl_{obs} between RPo and RPitc being substantially greater than the difference in Δl_{obs} between RPo and RPrtc). (*D*) Single-molecule transition-amplitude histograms for RPo and RPitc at the N25 promoter (WT; *Left*), and for RPo and RPrtc at derivatives of the N25 promoter containing template-strand $G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$ homopolymeric sequences ($G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$; *Middle* and *Right*). Upper subpanels, positively supercoiled DNA; *Lower* subpanels, negatively supercoiled DNA. Vertical dashed lines, means; $\Delta l_{obs,pos}$, transition amplitudes with positively supercoiled DNA; $\Delta l_{obs,neg}$, transition amplitudes with negatively supercoiled DNA. (*E*) Differences in $\Delta l_{obs,pos}$ and DNA unwinding relative to those in RPo at WT N25 promoter (means ±2 SEM).

	RPo	RPitc (WT)	RPrtc ($G_{+1}G_{+2}G_{+3}$)	RPrtc ($C_{+1}C_{+2}C_{+3}$)
$\Delta\Delta l_{obs,pos}$ (nm)	[0]	25±2	5±2	5±2
Δunwinding (bp)	[0]	5±0.4	1±0.4	1±0.4

position the 3′ end of the initial RNA product in the RNAP active-center P site—RNA extension does not involve DNA scrunching (Fig. 4*B*).

Measurement of Transcription-Bubble Size in RPrtc: RNA Extension without DNA Scrunching. As a second approach to test the proposal that RNA extension in reiterative transcription does not involve DNA scrunching—except for 1 bp of DNA scrunching to position the 3′ end of the initial RNA product in the RNAP active-center P site—we assessed transcription-bubble expansion during reiterative transcription initiation in solution by *E. coli* RNAP holoenzyme (Fig. 5).

We used magnetic-tweezers single-molecule DNA nanomanipulation (7, 47) to assess reiterative transcription (Fig. 5 *A* and *B*), analyzing the same promoter derivatives as in the preceding section (Fig. 4; SI Appendix, Fig. S1B). The resulting transition amplitudes, transition-amplitude histograms, and RNAP-dependent DNA unwinding values show that RNAP-dependent DNA unwinding is greater by 1 ± 0.4 bp in RPrtc engaged in synthesis of RNA products up to at least 50 nt in length than it is in RPo (Fig. 5 *C*–*E*). In contrast, RNAP-dependent DNA unwinding is greater by 5 ± 0.4 bp in RPitc engaged in synthesis of RNA products up to ~8 nt than it is in RPo (Fig. 5 *C*–*E*). An increase in transcription-bubble size is a defining hallmark of DNA scrunching (7, 47). Thus, the results

from single-molecule DNA nanomanipulation indicate that RPrtc engaged in synthesis of RNA products up to at least 50 nt in length exhibits only 1 bp of DNA scrunching. We conclude, consistent with the conclusions of the preceding sections, that in reiterative transcription, following synthesis of a 2-nt initial RNA product and 1 bp of DNA scrunching to position the 3′ end of the initial RNA product in the RNAP active-center P site, RNA extension does not involve DNA scrunching.

Discussion

Mechanism of Reiterative Transcription. Taken together, our results from X-ray crystallography (Figs. 1 and 2), cryo-EM (Fig. 3; *SI Appendix*, Fig. S5), site-specific protein–DNA photocrosslinking (Fig. 4), and DNA single-molecule nanomanipulation (Fig. 5) establish that, whereas standard transcription initiation involves 1 bp of DNA scrunching for each step of RNA extension, reiterative transcription initiation at promoters containing template-strand $G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$ homopolymeric sequences involves only 1 bp of DNA scrunching, irrespective of the number of steps of RNA extension. We conclude that. in reiterative transcription initiation at promoters containing template-strand $G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$ homopolymeric sequences, following synthesis of the initial 2 nt RNA product (Fig. 6, *Left*, lines 1 to 2), 1 bp of DNA scrunching occurs to place the RNA 3′ end in the RNAP active-center P site (Fig. 6, *Left*, lines 2 to 3), and no additional DNA scrunching occurs in RNA extension (Fig. 6, *Right*). We infer that, in reiterative transcription initiation, RNA extension does not involve movement of DNA relative to the RNAP active center, but instead involves RNA slipping, in which, in each step of RNA extension, RNA slips upstream by 1 nt relative to both template-strand DNA and the RNAP active center to place the RNA 3′ end in the RNAP active-center P site (Fig. 6, *Right*). Thus, we infer that, in RNA extension in reiterative transcription initiation, the nucleotide-addition cycle consists of RNA slipping to convert a pretranslocated state having a 3-bp RNA–DNA hybrid and having the RNA 3′ end in the RNAP active-center A site ("preslipped" state; Fig. 6, *Left*, RPitc,3 pre and Fig. 6, *Right*, RPrtc,n pre) into a posttranslocated state having a 2-bp RNA–DNA hybrid and having the RNA 3′ end in the RNAP active-center P site ("postslipped" state; Fig. 6, *Right*, RPrtc,n post), followed by NTP binding, phosphodiester-bond formation, and pyrophosphate release (Fig. 6, *Right*). Thus, in contrast to standard transcription initiation, in which the RNA–DNA hybrid increases in length by 1 bp in each step of RNA extension up to a length of ~10 bp, at which point promoter escape ensues (Fig. 6, *Left*; refs. 1, 2, 3, and 51), in reiterative transcription initiation, the RNA–DNA hybrid does not extend beyond a 3-bp state and instead alternates between a 3-bp preslipped state and a 2-bp postslipped state in each step of RNA extension (Fig. 6, *Right*). RNA slipping by 1 nt breaks 1 bp of the RNA–DNA hybrid—specifically the upstream-most base pair of the RNA–DNA hybrid (Fig. 6, *Right*); breakage of the upstream-most base pair of the RNA–DNA hybrid occurs because slipping moves the RNA nucleotide that had been base paired to the upstream-most nucleotide of the template-strand homopolymeric sequence into alignment with a noncomplementary DNA nucleotide upstream of the template-strand homopolymeric sequence (Fig. 6, *Right*). According to this mechanism, the "branching point," or "decision point," between the standard transcription initiation and reiterative transcription initiation pathways occurs upon formation of RPitc,3 pre (Fig. 6, lines 4 to 5); at this decision point, DNA scrunching followed by RNA extension yields standard transcription initiation, whereas RNA slipping followed by RNA extension yields reiterative transcription initiation (Fig. 6, lines 4 to 5).

The mechanism for reiterative transcription initiation defined here and set forth in Fig. 6, requires a template-strand homopolymeric sequence at least 3 nt in length. This aspect of the mechanism is consistent with, and supported by, the observation from previous work that reiterative transcription initiation is efficient only at promoters containing template-strand homopolymeric sequences at least 3 nt in length (9, 11, 20, 24, 52).

Our results from cryo-EM structure determination of RPrtc,≥11 $[C_{+1}C_{+2}C_{+3}]$ (Fig. 3; *SI Appendix*, Fig. S5) further show that production of long, ≥11 nt RNAs by reiterative transcription initiation involves an alternative RNA path for nucleotides rN6 to rN11 and an alternative RNA exit for nucleotides rN ≥ 11 (Fig. 3 *A, B*, and *D*; *SI Appendix*, Fig. S5C). The alternative RNA path exploits the cavity within the RNAP active-center cleft that accommodates scrunched non–template-strand DNA in standard transcription initiation (Fig. 1B; refs. 5, 8, 35, and 46), and the alternative RNA exit exploits the portal between the RNAP active-center cleft and bulk solvent that mediates the extrusion of long segments of scrunched nontemplate-strand DNA in standard transcription initiation (5, 8, 46). Because reiterative transcription initiation involves only 1 bp of DNA scrunching (Figs. 1 *C* and *D*, 3 *A–C*, and Figs. 4–6), this cavity within the RNAP active-center cleft and this portal between the RNAP active-center cleft and bulk solvent—both of which would be occupied by scrunched nontemplate-strand DNA in standard transcription initiation—are not occupied by scrunched nontemplate-strand DNA in reiterative transcription and instead are available to be occupied by RNA.

A small change in local conformation of the σ-finger tip—a folding back of the σ-finger tip upon itself—enables reiteratively transcribed RNA to enter the alternative RNA path, thereby enabling the extension of reiteratively transcribed RNA beyond a length of 5 nt (*SI Appendix*, Fig. S5D). A small change in RNAP clamp conformation—an opening of the RNAP clamp by ~3°—enables reiteratively transcribed RNA to cross the nontemplate DNA strand and to access the alternative RNA exit, thereby enabling the extension of reiteratively transcribed RNA beyond a length of 8 nt and the extrusion of reiteratively transcribed RNA from the RNAP active-center cleft (*SI Appendix*, Fig. S5E).

Because the alternative RNA path and the alternative RNA exit enable reiteratively transcribed RNA to be extruded from the RNAP active-center cleft without substantially disrupting σ-DNA and σ-RNAP interactions (Fig. 3 *A* and *B*), these features enable reiterative transcription initiation to generate long RNAs without promoter escape.

Because most protein–RNA interactions in the alternative RNA path and alternative RNA exit involve interactions with the RNA sugar–phosphate backbone (*SI Appendix*, Fig. S5C), and because interactions with RNA bases involve only single H-bonds that likely could be made, with wobble, by any RNA nucleotide (*SI Appendix*, Fig. S5C), the alternative RNA path and alternative RNA exit are likely to be compatible with any RNA sequence.

Because the RNAP and σ residues that make protein–RNA interactions with RNA in the alternative RNA path and alternative RNA exit are invariant or highly conserved across gram-negative, gram-positive, and *Thermus–Deinococcus*-clade bacteria (*SI Appendix*, Fig. S6), the alternative RNA path and alternative RNA exit and the mechanism of reiterative transcription initiation set forth here and in Fig. 6 are likely to be conserved across bacterial species.

Prospect. The results and mechanism of this work, pertain to reiterative transcription initiation at promoters containing template-strand $G_{+1}G_{+2}G_{+3}$ and $C_{+1}C_{+2}C_{+3}$ homopolymeric sequences (Figs. 1–6). Previous work shows that reiterative
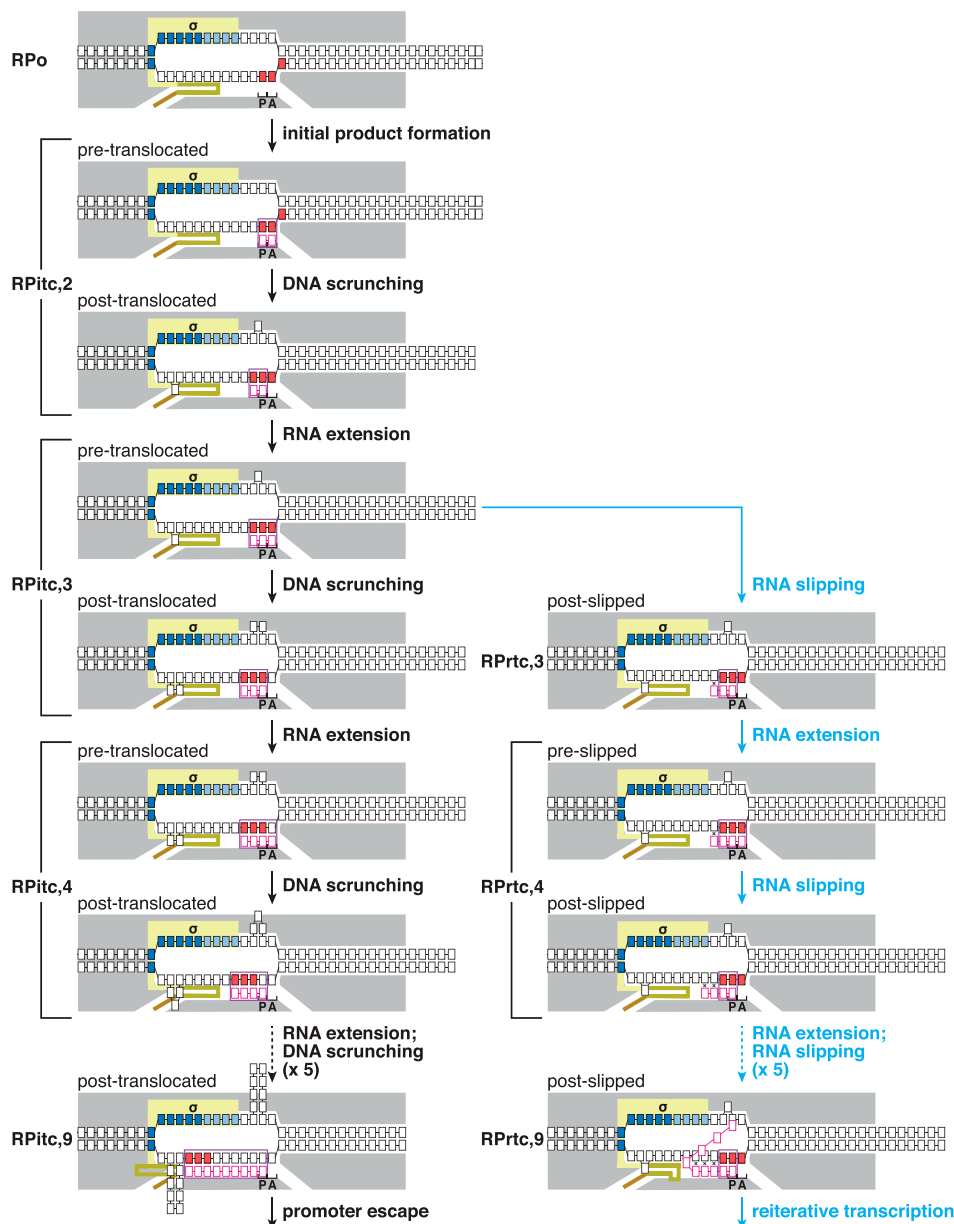
Fig. 6. Mechanisms of standard transcription initiation and reiterative transcription initiation. Standard transcription initiation (*Left* column) and reiterative transcription initiation (first four panels of *Left* column followed by panels of *Right* column). Cyan, reactions present only in reiterative transcription: i.e., cycles of RNA extension and slippage. Other colors and symbols are as in Fig. 4. Scrunching is indicated by bulged-out nucleotides (~8 to 10 scrunched bp prior to promoter escape in the standard transcription initiation pathway; 1 scrunched bp in the reiterative transcription initiation pathway). Scrunched nucleotides of nontemplate and template DNA strands during initial transcription are accommodated as bulges within the unwound transcription bubble.

transcription initiation also occurs efficiently at promoters containing template-strand $A_{+1}A_{+2}A_{+3}$ and $T_{+1}T_{+2}T_{+3}$ homopolymeric sequences (9, 11, 24) and at promoters containing ≥4-nt homopolymeric sequences (9, 11, 24). We consider it likely that the mechanism of Fig. 6 also applies to these promoters. Previous work also shows that reiterative transcription initiation can occur at promoters containing more complex, nonhomopolymeric repeat sequences (9, 11, 24). We consider it likely that a mechanism related to the mechanism of Fig. 6, but with different extents of DNA scrunching and RNA slipping, applies at these promoters. We note that the crystal structure determination, cryo-EM structure determination, site-specific protein–DNA photo–cross-linking, and single-molecule nanomanipulation procedures of this report could be used to define mechanisms of reiterative transcription at any promoter.

Previous work has shown that complexes engaged in reiterative transcription initiation synthesizing RNA products up to at least 8 nt in length can switch from reiterative transcription initiation to standard transcription initiation, yielding productive complexes that escape the promoter and synthesize full-length RNA products

comprising nontemplated, reiterative-transcription–dependent nucleotides at their 5′ ends followed by templated, standard-transcription–dependent nucleotides (9, 11, 24). Key unresolved questions include how this switching occurs and where the RNA products that result from switching exit the RNAP active-center cleft. We hypothesize that RNA products that result from switching exit the RNAP active-center cleft through the RNAP RNA-exit channel, with the templated RNA segment generated after switching to standard transcription initiation proceeding into and through the RNAP RNA-exit channel as in standard transcription, and pulling the 5′ nontemplated RNA segment behind it. According to this hypothesis, the RNA product that results from switching would proceed into and through the RNAP RNA-exit channel as an RNA loop and would trigger promoter escape by displacing the σ region-3/region-4 linker from the RNA-exit channel as in standard transcription (*SI Appendix*, Fig. S7, *Right*). Cryo-EM structures of transcription complexes containing double-stranded RNA in the RNAP RNA-exit channel verify that the dimensions of the RNAP RNA-exit channel could accommodate such an RNA loop (53, 54). Cryo-EM structure determination, analyzing transcription

Liu et al.
Structural and mechanistic basis of reiterative transcription initiation

PNAS | 11 of 12
https://doi.org/10.1073/pnas.2115746119

BIOCHEMISTRY

BIOPHYSICS AND COMPUTATIONAL BIOLOGY

elongation complexes generated by switching from reiterative transcription initiation to standard transcription initiation, potentially could provide a means to test this hypothesis.

Materials and Methods

Crystal structures were determined using molecular replacement. The cryo-EM structure was determined using single-particle reconstruction. Protein–DNA photo–cross-linking was performed as in refs. 47, 48, and 55. Magnetic-tweezers single-molecule DNA nanomanipulation was performed as in ref. 47.

Full details of methods are presented in *SI Appendix, Materials and Methods*.

1. E. F. Ruff, M. T. Record Jr., I. Artsimovitch, Initial events in bacterial transcription initiation. *Biomolecules* **5**, 1035–1062 (2015).
2. J. T. Winkelman, B. E. Nickels, R. H. Ebright, "The transition from transcription initiation to transcription elongation: Start-site selection, initial transcription, and promoter escape" in *RNA Polymerase as a Molecular Motor*, R. Landick, J. Wang, T. R. Strick, Eds. (RSC Publishing, Cambridge, UK, ed. 2, 2021), pp. 1–24.
3. A. Mazumder, A. N. Kapanidis, Recent advances in understanding σ^{70}-dependent transcription initiation mechanisms. *J. Mol. Biol.* **431**, 3947–3959 (2019).
4. J. Chen, H. Boyaci, E. A. Campbell, Diverse and unified mechanisms of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **19**, 95–109 (2021).
5. J. T. Winkelman *et al.*, Crosslink mapping at amino acid-base resolution reveals the path of scrunched DNA in initial transcribing complexes. *Mol. Cell* **59**, 768–780 (2015).
6. E. Margeat *et al.*, Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophys. J.* **90**, 1419–1431 (2006).
7. A. Revyakin, C. Liu, R. H. Ebright, T. R. Strick, Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* **314**, 1139–1143 (2006).
8. A. N. Kapanidis *et al.*, Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314**, 1144–1147 (2006).
9. C. L. Turnbough Jr., R. L. Switzer, Regulation of pyrimidine biosynthetic gene expression in bacteria: Repression without repressors. *Microbiol. Mol. Biol. Rev.* **72**, 266–300 (2008).
10. J. P. Jacques, D. Kolakofsky, Pseudo-templated transcription in prokaryotic and eukaryotic organisms. *Genes Dev.* **5**, 707–713 (1991).
11. C. L. Turnbough Jr., Regulation of gene expression by reiterative transcription. *Curr. Opin. Microbiol.* **14**, 142–147 (2011).
12. M. Chamberlin, P. Berg, Mechanism of RNA polymerase action: Characterization of the DNA-dependent synthesis of polyadenylic acid. *J. Mol. Biol.* **8**, 708–726 (1964).
13. M. Chamberlin, P. Berg, Deoxyribo ucleic acid-directed synthesis of ribonucleic acid by an enzyme from *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **48**, 81–94 (1962).
14. Q. Meng, C. L. Turnbough Jr., R. L. Switzer, Attenuation control of *pyrG* expression in Bacillus subtilis is mediated by CTP-sensitive reiterative transcription. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10943–10948 (2004).
15. A. H. Tu, C. L. Turnbough Jr., Regulation of *upp* expression in *Escherichia coli* by UTP-sensitive selection of transcriptional start sites coupled with UTP-dependent reiterative transcription. *J. Bacteriol.* **179**, 6665–6673 (1997).
16. F. Qi, C. L. Turnbough Jr., Regulation of codBA operon expression in *Escherichia coli* by UTP-dependent reiterative transcription and UTP-sensitive transcriptional start site switching. *J. Mol. Biol.* **254**, 552–565 (1995).
17. K. Severinov, A. Goldfarb, Topology of the product binding site in RNA polymerase revealed by transcript slippage at the phage lambda PL promoter. *J. Biol. Chem.* **269**, 31701–31705 (1994).
18. C. Liu, L. S. Heath, C. L. Turnbough Jr., Regulation of *pyrBI* operon expression in Escherichia coli by UTP-sensitive reiterative RNA synthesis during transcriptional initiation. *Genes Dev.* **8**, 2904–2912 (1994).
19. H. C. Guo, J. W. Roberts, Heterogeneous initiation due to slippage at the bacteriophage 82 late gene promoter in vitro. *Biochemistry* **29**, 10702–10709 (1990).
20. Y. Cheng, S. M. Dylla, C. L. Turnbough Jr., A long T. A tract in the *upp* initially transcribed region is required for regulation of *upp* expression by UTP-dependent reiterative transcription in *Escherichia coli*. *J. Bacteriol.* **183**, 221–228 (2001).
21. L. A. Wagner, R. B. Weiss, R. Driscoll, D. S. Dunn, R. F. Gesteland, Transcriptional slippage occurs during elongation at runs of adenine or thymine in *Escherichia coli*. *Nucleic Acids Res.* **18**, 3529–3535 (1990).
22. X. Han, C. L. Turnbough Jr., Regulation of *carAB* expression in *Escherichia coli* occurs in part through UTP-sensitive reiterative transcription. *J. Bacteriol.* **180**, 705–713 (1998).
23. I. E. Jensen-MacAllister, Q. Meng, R. L. Switzer, Regulation of *pyrG* expression in *Bacillus subtilis*: CTP-regulated antitermination and reiterative transcription with *pyrG* templates in vitro. *Mol. Microbiol.* **63**, 1440–1452 (2007).
24. I. O. Vvedenskaya *et al.*, Massively systematic transcript end readout, "MASTER": Transcription start site selection, transcriptional slippage, and transcript yields. *Mol. Cell* **60**, 953–965 (2015).
25. C. L. Turnbough Jr., Regulation of bacterial gene expression by transcription attenuation. *Microbiol. Mol. Biol. Rev.* **83**, e00019-19 (2019).
26. L. E. Macdonald, Y. Zhou, W. T. McAllister, Termination and slippage by bacteriophage T7 RNA polymerase. *J. Mol. Biol.* **232**, 1030–1047 (1993).
27. Y. Shin, M. Hedglin, K. S. Murakami, Structural basis of reiterative transcription from the *pyrG* and *pyrBI* promoters by bacterial RNA polymerase. *Nucleic Acids Res.* **48**, 2144–2155 (2020).
28. K. S. Murakami, Y. Shin, C. L. Turnbough Jr., V. Molodtsov, X-ray crystal structure of a reiterative transcription complex reveals an atypical RNA extension pathway. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 8211–8216 (2017).
29. G. Zhang *et al.*, Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. *Cell* **98**, 811–824 (1999).
30. Y. Feng, Y. Zhang, R. H. Ebright, Structural basis of transcription activation. *Science* **352**, 1330–1333 (2016).
31. K. S. Murakami, S. Masuda, E. A. Campbell, O. Muzzin, S. A. Darst, Structural basis of transcription initiation: An RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285–1290 (2002).
32. K. S. Murakami, S. Masuda, S. A. Darst, Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science* **296**, 1280–1284 (2002).
33. D. G. Vassylyev *et al.*, Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* **417**, 712–719 (2002).
34. Y. Zhang *et al.*, Structural basis of transcription initiation. *Science* **338**, 1076–1080 (2012).
35. R. S. Basu *et al.*, Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme. *J. Biol. Chem.* **289**, 24549–24559 (2014).
36. A. Feklístov, B. D. Sharon, S. A. Darst, C. A. Gross, Bacterial sigma factors: A historical, structural, and genomic perspective. *Annu. Rev. Microbiol.* **68**, 357–376 (2014).
37. L. Li, V. Molodtsov, W. Lin, R. H. Ebright, Y. Zhang, RNA extension drives a stepwise displacement of an initiation-factor structural module in initial transcription. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 5801–5809 (2020).
38. V. Mekler *et al.*, Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* **108**, 599–614 (2002).
39. A. Mazumder, M. Lin, A. N. Kapanidis, R. H. Ebright, Closing and opening of the RNA polymerase trigger loop. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15642–15649 (2020).
40. D. Duchi, A. Mazumder, A. M. Malinen, R. H. Ebright, A. N. Kapanidis, The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp. *Nucleic Acids Res.* **46**, 7284–7295 (2018).
41. A. Chakraborty *et al.*, Opening and closing of the bacterial RNA polymerase clamp. *Science* **337**, 591–595 (2012).
42. M. Bokori-Brown *et al.*, Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. *Nat. Commun.* **7**, 11293 (2016).
43. T. G. Martin, A. Boland, A. W. P. Fitzpatrick, S. H. W. Scheres, Graphene oxide grid preparation. https://doi.org/10.6084/m9.figshare.3178669.v1. Accessed 9 January 2022.
44. Z. Yin, J. T. Kaelber, R. H. Ebright, Structural basis of Q-dependent antitermination. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 18384–18390 (2019).
45. W. J. Lane, S. A. Darst, Molecular evolution of multisubunit RNA polymerases: Sequence analysis. *J. Mol. Biol.* **395**, 671–685 (2010).
46. A. Hasemeyer, "Mechanism of DNA scrunching during initial transcription," PhD thesis, Rutgers University, New Brunswick, NJ (2018).
47. L. Yu *et al.*, The mechanism of variability in transcription start site selection. *eLife* **6**, e32038 (2017).
48. J. T. Winkelman *et al.*, Multiplexed protein-DNA cross-linking: Scrunching in transcription start site selection. *Science* **351**, 1090–1093 (2016).
49. J. T. Winkelman, R. L. Gourse, Open complex DNA scrunching: A key to transcription start site selection and promoter escape. *BioEssays* **39**, 1600193 (2017).
50. J. T. Winkelman, P. Chandrangsu, W. Ross, R. L. Gourse, Open complex scrunching before nucleotide addition accounts for the unusual transcription start site of *E. coli* ribosomal RNA promoters. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E1787–E1795 (2016).
51. L. M. Hsu, N. V. Vo, C. M. Kane, M. J. Chamberlin, In vitro studies of transcript initiation by *Escherichia coli* RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. *Biochemistry* **42**, 3777–3786 (2003).
52. X. F. Xiong, W. S. Reznikoff, Transcriptional slippage during the transcription initiation process at a mutant *lac* promoter in vivo. *J. Mol. Biol.* **231**, 569–580 (1993).
53. J. Y. Kang *et al.*, RNA polymerase accommodates a pause RNA hairpin by global conformational rearrangements that prolong pausing. *Mol. Cell* **69**, 802–815.e5 (2018).
54. X. Guo *et al.*, Structural basis for NusA stabilized transcriptional pausing. *Mol. Cell* **69**, 816–827.e4 (2018).
55. J. T. Winkelman *et al.*, XACT-seq comprehensively defines the promoter-position and promoter-sequence determinants for initial-transcription pausing. *Mol. Cell* **79**, 797–811.e8 (2020).
56. A. Revyakin, R. H. Ebright, T. R. Strick, Single-molecule DNA nanomanipulation: Improved resolution through use of shorter DNA fragments. *Nat. Methods* **2**, 127–138 (2005).

12 of 12 | **PNAS**
https://doi.org/10.1073/pnas.2115746119

Liu et al.
Structural and mechanistic basis of reiterative transcription initiation