Published in final edited form as:

Science. 2008 October 24; 322(5901): 553–557. doi:10.1126/science.1163433.

The Structure of a Transcribing T7 RNA Polymerase Complex Captured During Its Transition from Initiation to Elongation

Kimberly J. Durniak¹, Scott Bailey^{1,2,4}, and Thomas A. Steitz^{1,2,3,*}

¹Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA.

²Howard Hughes Medical Institute, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA.

³Department of Chemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA.

Structural studies of the T7 bacteriophage DNA-dependent RNA polymerase (T7 RNAP) have shown large structural differences between its initiation and elongation phases of transcription, but the mechanism by which this transition is achieved remains unclear. Crystal structures of T7 RNAP bound to promoter DNA containing either a seven or an eight nucleotide (nt) RNA transcript, now illuminate intermediate states along the transition pathway. These intermediate complex structures reveal that the six helix promoter-binding subdomain and the bound promoter rotate by approximately 45° upon synthesis of an 8nt RNA transcript, thereby allowing the maintenance of promoter contacts while expanding the active site to accommodate a growing heteroduplex. Although modest movements in the C-helix subdomain towards its elongation conformation are also observed, subdomain H undergoes little change.

RNA polymerases (RNAPs) exhibit three phases of transcription – initiation, elongation and termination. The initiation and elongation phases, have been studied extensively in the T7 RNAP system by biochemical and structural approaches [reviewed in (1,2)]. During the initiation phase, the RNAP binds to a specific promoter DNA sequence, opens the DNA duplex and feeds the template strand into the active site (3,4). The structure of a T7 RNAP initiation complex identified a six-helix bundle sub-domain (residues 72-150, 191-267), here termed the promoter binding domain (PBD), that is responsible for many of the interactions with the 17 base-pair (bp) promoter and, in part, for melting the DNA duplex (3,4) (Fig 1A). While remaining bound to the promoter, the polymerase produces many short transcripts from 2-12 nucleotides (nt) in length (5,6), a process often referred to as abortive synthesis. After the transition to the elongation phase and release of the promoter, the polymerase proceeds processively down the DNA template producing a full length RNA transcript. Comparison of the structures of the T7 RNAP initiation and elongation complexes revealed extensive conformational changes within the N-terminal 267 residues (N-terminal domain) and little change in the rest of the RNAP (Fig 1B) (7,8). A rigid body rotation of the PBD as well as the refolding of the N-terminal C-helix (residues 28-71) and H (residues 151-190) subdomains are responsible for abolishing the promoter binding site, enlarging the active site and creating an exit tunnel for the RNA transcript.

*To whom correspondence should be addressed. thomas.steitz@yale.edu.

⁴Present address: Department of Biochemistry and Molecular Biology, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, MD 21205, USA

Supporting Online Material

www.sciencemag.org Materials and Methods Figs. S1, S2, S3 Movie S1

The structural changes within the N-terminal domain account for the increased stability and the processivity of the elongation complex, yet provide little insight into how the polymerase is able to accommodate a growing transcript while maintaining its promoter contacts during the abortive synthesis phase. The structure of the initiation complex contains a 3nt RNA transcript in the active site DNA:RNA heteroduplex. Modeling the elongation of the RNA transcript by the addition of a single RNA nucleotide produced a steric clash with the PBD and incorporation of an additional NTP destroyed the crystal (4). Covalent cross-linking of the template strand to the RNA transcript established that the heteroduplex can be as long as 8bp (9,10) thereby eliminating the possibility that the heteroduplex might be only 3bp long (4). The enlarged active site observed in the structure of the elongation complex (Fig 1B) can accommodate a 7-8bp heteroduplex, however the new orientation of the PBD abolishes the promoter binding site. Since biochemical data suggest that the polymerase remains bound to the promoter with transcripts as long as 8 to 10nt, the question of how the polymerase accommodates a growing RNA transcript while maintaining its promoter contacts during the transition from initiation to elongation is raised.

Cross-linking, mutagenesis and proteolytic digestion experiments have provided information relating the length of the RNA transcript to changes within the N-terminal domain (11-13), and suggest the existence of at least one intermediate structure. A common conclusion from the biochemical data is that the transition mechanism consists of two stages. Cross-linking and mutagenesis experiments that prevent major refolding of the T7 RNAP structure (11-15) suggest that the first stage allows synthesis of up to 8nt of RNA with minimal changes to the N-terminal sub-domains. The second stage is presumed to include the major refolding events that occur during the synthesis of 9-14nt and allow the stable elongation complex to form (16). A transition mechanism consisting of two stages has also been proposed for the larger, multi-subunit eukaryotic RNAPs (17).

Models of the transition from initiation to elongation proposed previously have suggested that the N-terminal domain undergoes a gradual structural rearrangement in order to accommodate an 8bp heteroduplex(7,8). One model proposes a 10Å translation of the PBD and subdomain H away from the active site (18) , while another suggests the PBD can maintain its promoter contacts after rotating into the position observed in the elongation complex (7). None of these models are entirely consistent with the biochemical data (13,19). However, while this manuscript was in preparation, Tang *et al.* posited that a rotation of the N-terminal domain could accommodate their FRET data as well as previous biochemical data (2). Although these new FRET data predict a 20° rotation, they do not provide the direction of the rotation or which components are rotating.

The structures of T7 RNAP bound to promoter DNA containing either a 7nt or 8nt RNA transcript presented here represent complexes that are between the early initiation and the elongation phases. These structures exhibit the domain movements that are necessary to accommodate transcript elongation during the abortive synthesis phase, and provide a model that is consistent with all of the biochemical data as well as the two-stage transition mechanism.

Structures of Intermediate Transcription Complexes

The structure of a mutant T7 RNAP captured in an intermediate transcription initiation state complexed with promoter DNA and a 7nt RNA transcript was solved at 3.0Å resolution (Fig 1C). Our initial attempts to capture the structure of an intermediate complex using wild type T7 RNAP resulted in structures of complexes with DNA that lacked the added RNA transcript, presumably due to the instability of an intermediate complex relative to an aborted complex. To address this problem, we used a mutant T7 RNAP that produces fewer abortive products and dissociates more slowly than wild type when stalled at 6nt of RNA (14); this proved

invaluable in capturing a structure that represents an intermediate between the initiation and elongation phases of transcription. The polymerase with a proline-to-leucine point mutation at residue 266 (P266L) was co-crystallized with a 33-bp duplex DNA containing a 17bp conserved promoter region, a transcription bubble of 6 non-complementary bases followed by a 10bp downstream complementary region, and a 7nt long synthetic RNA that was complementary to the template in the bubble region (Fig 2B). Complexes employing smaller RNA transcripts either failed to crystallize or resulted in aborted complexes as happened in our previous attempts using the wild type polymerase. The substrate template DNA, non-template DNA and nascent RNA were annealed, and then mixed with the P266L mutant polymerase to form the intermediate complex (20). The structure was solved by molecular replacement using the protein component of the initiation complex (1QLN) as the search model (4). A difference electron density map calculated using 2Fo-Fc as coefficients and phases derived from coordinates that never included the nucleic acid is shown in Fig 2A. The refined model has an R_{factor} of 25.0% (R_{free} of 29.2%). Data collection and refinement statistics are shown in Table 1.

We have also solved a structure of an initiation complex bound to promoter DNA and an 8nt RNA transcript at 6.7Å resolution. The experimental design was similar to that of the 7nt RNA transcript complex. The 32bp DNA substrate contained the 17bp promoter, followed by an 8nt transcription bubble and 7bp of downstream duplex DNA (20) (Fig S1A). This structure was solved using molecular replacement with the C-terminal residues of the polymerase, followed by an additional search for the PBD and 17bp promoter DNA which was necessary to include in the search for the PBD due to its large scattering mass. Phases calculated from a model excluding the active site nucleic acid contained electron density for the active site DNA and RNA in a 2Fo-Fc electron density map calculated to 6.7Å resolution (Fig S1B). When a central helix in the PBD (residues 204-223) was left out during calculation of another electron density map, electron density appears for this helix in a different position from the 7nt RNA model. Comparison of the 8nt intermediate complex structure with that of the 7nt intermediate complex shows the PBD rotates about the same axis by an additional 6° and moves 3-5Å further from the catalytic site.

In our structure containing the 7nt RNA, the polymerase is observed bound to both upstream and downstream DNA, and has formed an 11nt transcription bubble (Fig 2A, B). The single-stranded non-template strand in the opened transcription bubble makes interactions with the finger and thumb domains that are similar to those seen in the elongation complex (Fig S2A, B). The RNA backbone of nucleotides 4 and 5 contacts the thumb domain. The 5' end of the RNA is located towards the specificity loop and its 3' end is situated in the active site in a pre-translocated position with the fingers domain in the expected closed conformation. The Cα backbones of the palm domains (residues 411-552) of the initiation and the intermediate complexes can be superimposed with a root mean squared deviation (rmsd) of 0.40Å.

Movements of the N-terminal Sub-domains, Specificity Loop and the Promoter DNA

The structure of the 7nt intermediate complex reveals a left-handed 40° rotation of the PBD, the specificity loop, and the bound promoter about an axis passing through residues 198-204, consistent with biochemical data indicating that T7 RNAP maintains promoter contacts during abortive synthesis (Fig 3A, B). The specificity loop (residues 739-769), which is an insertion into the fingers domain, is responsible for making sequence specific contacts with base pairs in the promoter region (4, 21). Superimposing the Cα backbone of the PBD and the specificity loop of the initiation complex on that of the intermediate complex gives an rmsd of 2.0Å, in agreement with a largely rigid body motion of the PBD and the specificity loop. The contacts between the PBD, the specificity loop and the promoter DNA in the structure of the

intermediate complex remain the same as in the initiation complex, but their rotational movement enlarges the active site to accommodate the 7nt of elongated RNA. Previous footprinting analyses of the promoter DNA shows protection by T7 RNAP during synthesis of up to 9nt of RNA (22). Experiments utilizing ultraviolet-laser cross-linking or mutagenesis of the promoter sequence indicate that the specificity loop loses its promoter contacts after the synthesis of 8-9nt (16, 23). We have established that the PBD can rotate as far as 45° during the transcription of 8nt of RNA (Movie S1), which is a structural change in the polymerase that is consistent with these biochemical data.

In the initiation to elongation structural change, a loop between the C1 helix (residues 28-41) and the C2 helix (residues 46-55) refolds creating an elongated C-helix subdomain important for elongation complex stability (13). Superposition of the polymerase palm domains of the initiation and elongation complexes reveals that the C1 helix does not change its position, rather the C2 helix stacks onto the C1 helix thereby lengthening it from 20Å to 50 Å (8) (Fig 4A). In the structure of the intermediate complex, the C2 helix is observed in between the conformations observed in the structures of the initiation and elongation complexes. As the PBD rotates away from the active site during the transition, the loop between the C1 and C2 helices starts to refold as the C2 helix rotates by 40° towards the C1 helix. The C2 helix must rotate by an additional 50° to form the continuous elongated C-helix subdomain observed in the elongation complex. A comparison of all three structures reveals the refolding of this loop is important for formation of the C-helix subdomain (Fig 4A). This is consistent with the mutagenesis of residues within this loop to proline, which disrupts the refolding event, increases the accumulation of transcripts 6-7nt long and prevents the formation of stable elongation complexes (13).

Although subdomain H is an insertion into the six helix bundle of the PBD, the flexible loop connecting these regions allows the PBD to rotate with minimal changes in the conformation of subdomain H. Part of subdomain H is a loop (residues 165-190) during the initiation phase, and ten residues within this loop (residues 167-177) are disordered in the structure of the intermediate complex. However, crystal packing interactions help to order the loop such that electron density for the entire loop is observed in the structure of the initiation complex. The other part of subdomain H is a helix of 14 residues (151-164), and a superposition of the initiation and intermediate complexes along the palm domain reveals little movement in subdomain H during the first part of the transition from initiation to elongation (Fig S3).

Role of the P266L Mutation

Comparison of the conformation of residue 266 in the structures of the initiation, the intermediate and the elongation complexes reveals that this residue plays a key role during the transition. The P266L mutation is located at the C-terminus of the PBD in a loop connecting the N-terminal domain with the C-terminal thumb, palm and fingers domains (Fig 1C). A rotation of 180° around the peptide bond between residue 267 and 268 allows the conformational changes in the N-terminal domain to occur without any change in the C-terminal domain (Fig 4B, C). A proline at residue 266, while not at the center of rotation, presumably limits the flexibility of the protein near this point because the covalent bond between the Cδ and the backbone nitrogen restricts the peptide backbone φ angle of proline to about -70°. However, the weak electron density in this region (residues 258-264) precludes a more detailed analysis of the backbone angles of the P266L mutant. Perhaps the mutation allows formation of an intermediate structure that is more energetically favorable relative to the initiation structure than is allowed by the proline. Consistent with this hypothesis, we were unable to produce crystals of this intermediate complex with wild type T7 RNAP.

Movements within the Thumb Domain and the Downstream DNA

The position of the thumb domain in the intermediate complex is between its position in the initiation complex and in the elongation complexes (Fig 4D), as is observed with the C2 helix. In the structure of the elongation complex, the thumb domain has moved by 8Å from its position in the initiation complex, creating a binding cleft for the upstream DNA (8). Although, the binding cleft has not yet formed in the structure of the intermediate complex, the thumb domain has moved by 4Å from its position in the initiation complex and makes interactions with the backbone of the non-template strand as well as with the RNA chain at residue 4 and 5. These interactions are similar to those seen in the structure of the elongation complex. This structure is consistent with mutagenesis experiments indicating the role the thumb domain plays in stabilizing the transcription complex (24,25).

The structure of the T7 RNAP intermediate complex now shows the polymerase bound to the promoter as well as to downstream DNA. Comparison of the downstream DNA in the structures of the intermediate complex with those of the elongation complexes shows that the downstream DNA is rotated by 30° toward the N-terminal domain in the intermediate complex (Fig 5A). The angle between the promoter and the downstream DNA is about 40°, bringing the upstream and downstream duplex DNAs to within 6Å of each other (Fig 5B). The difference in the position of the downstream DNA may be correlated to the position of subdomain H. A superposition of the Cα backbone of the palm domains of the intermediate and the elongation complexes (1MSW) (8) reveals that the position of subdomain H in the elongation complex would clash with the position of the downstream DNA in the intermediate complex (Fig 5C). Perhaps upon refolding into its position on top of the polymerase in the elongation complex, subdomain H is responsible for not only interacting with the nascent RNA chain and the specificity loop (8,26), but also for moving the downstream DNA into the position observed in the structure of the elongation complex. In our low resolution structure of the 8nt intermediate complex the position of the downstream DNA is different in each of the two copies in the asymmetric unit. The unbiased electron density for the downstream DNA agrees with the position of the downstream DNA seen in the 7nt intermediate complex in one copy, while in the other copy, the density agrees better with the position seen in the elongation complex. This provides further evidence that there is movement in the downstream DNA during the transition from initiation to elongation. The results of FRET (27) and chemically-tethered nuclease (28) experiments that were conducted to determine the relationship of the downstream DNA to the fingers domain, are consistent with either of the positions observed in the crystal structures.

A Model for the Transition from Initiation to Elongation

These structures of intermediate complexes resolve the ambiguity in the current models of the transition mechanism (7,8,18). The model proposed by Theis *et al.*, based on FRET data, suggested that the initial step of the transition from initiation to elongation involves a 10Å translation of the PBD and subdomain H away from the C-terminal active site (18,19). However, the FRET data also correlate well with these crystal structures of the intermediate complex. The left-handed rotation of the PBD moves the FRET probe on the non-template strand ~10Å from its original position, in agreement with the 10Å change observed from the FRET data, but achieved through a different mechanism. Our structures also reveal minimal changes in subdomain H during the initial stage of the transition. The rotation of the C2 helix towards the C1 helix observed in the structure of the 7nt RNA intermediate complex indicates formation of the C-helix subdomain is underway before the loss of promoter contacts and is consistent with mutagenesis experiments that disrupt the formation of the C-helix subdomain and result in the accumulation of RNA transcripts of 6-7nt in length (13). Experimental FRET data have disproved another model which proposed that the PBD would be able to maintain

its contacts with the promoter DNA after the conformational changes and refolding of the N-terminal subdomains that occur during the transition from initiation to elongation (7,19).

Recently, Tang *et al.* (2) proposed based on FRET data that the promoter and the N-terminal domain rotate by 20° about an axis passing through the -4 position on the non-template (NT) strand, during the synthesis of a 3-7nt RNA transcript. However, the direction of the rotation, as well as the specific changes of the individual subdomains of the N-terminal domain during the rotation, were not determined. Our structures show that the synthesis of an 8nt RNA transcript results in a left-handed rotation of 45° about an axis which passes through the end of a central helix within the middle of the PBD (Movie S1) rather than at one end of the PBD (-4NT strand) as proposed from the FRET measurements (Fig 3).

It is possible that during the initial stages of the transition, additional stable intermediate conformations occur. For example, as the RNA elongates, the PBD may undergo a stepwise rotation to reach the position observed in the structure of the 7nt intermediate complex. Attempts to capture additional intermediates using shorter RNAs have thus far failed to yield crystals or have resulted in structures without an RNA transcript in the active site. The difference between the 7nt RNA and a shorter 6nt RNA would be one G-C base-pair (Fig 2B). It is plausible that the -3kcal/mole associated with the G-C base pair provides enough energy to stabilize an intermediate state and that reducing the length of the RNA transcript shifts the equilibrium towards the more stable T7 RNAP initiation state conformation. Consistent with multiple rotation states, the PBD is observed to rotate by an additional 6° when the structures of the 7nt and 8nt intermediate complexes are compared. The recent FRET data of Tang *et al.* (2) show an increase in probe distances as the transcript increases in length, also indicating a gradual rotation of the PBD. The movement of the PBD from its position in the initiation complex to that in the intermediate can be interpolated in order to view the transition (Movie S1).

The second stage of the transition, after the RNA is extended beyond 8nt, involves the loss of promoter contacts and larger structural changes in the specificity loop, the PBD and subdomain H (12,13,15,29). Modeling the longer RNA transcript observed in the structure of the elongation complex onto the intermediate structure shows that an RNA chain extended beyond 8nt will clash with the specificity loop (Fig 5D). This is consistent both with mutagenesis (16) and cross-linking studies (9) which indicate that the specificity loop loses its contacts with the promoter and begins to interact with the growing RNA chain upon synthesis of 8-9nt. The PBD must also be released from its contacts with the promoter and rotate into the position it occupies in the elongation phase. During the first stage of the transition that we observe here, the PBD undergoes a left-handed 45° rotation. However, further examination suggests that the PBD can only achieve its final position in the elongation complex through a 260° right-handed rotation, as proposed by Theis *et al.* (18), which is in the opposite direction to that observed during the transition from the initiation to the 8nt intermediate complex. This larger rotation of the PBD is likely to be associated with changes in subdomain H, consistent with mutant T7 RNAPs which suggest that the hinge region between subdomain H and the PBD is important for the transition to the elongation phase (26).

Relation to the multi-subunit RNA Polymerases

The abortive synthesis phase of transcription initiation raises a common problem among the structurally-unrelated RNAPs of bacteriophage T7, bacteria and eukaryotes: how does the polymerase maintain contact with the promoter while accommodating an increase in the size of the elongating heteroduplex and progress down the template DNA. The sequence-specific recognition of the promoter DNA is carried out by the σ factor in *E.coli* RNA polymerase and TFIIB (with other cofactors) in the eukaryotic RNA polymerase II (30-32). The “inch-

worming” model posited that the polymerase domain moved down the template away from the promoter recognition domain, while the “scrunching” model suggested that the product accumulated in the enzyme. The emerging answer appears to exhibit aspects of both models. Structural studies of the three RNAP families have revealed that extension of the RNA transcript requires displacement of a steric block that prevents the RNA from being extended during the transition from initiation to elongation (4,33-35). The steric block is caused by a region of the polymerase or associated protein that is important for promoter recognition – the PBD in T7 RNAP, the σ factor in bacterial RNAP or TFIIB in eukaryotic RNAP. The intermediate structure of T7 RNAP reveals that as the RNA chain lengthens during abortive synthesis, the obstacle is rotated out of the way, which allows the polymerase to maintain promoter contacts while enlarging the product site. Likewise, the extension of the transcript is proposed to push a domain of the σ protein from the exit tunnel of bacterial RNAP (35), leading to promoter release, and the extension of the heteroduplex by yeast Pol II must displace TFIIB from the product binding site (34).

A similar mechanism of progressive product displacement of protein has also been proposed for the initiation to elongation transition in the DNA polymerase of bacteriophage ϕ 29 where DNA synthesis is primed at the 3' end of the template strand by a protein primer called terminal protein (TP). It has been proposed that as the primer strand is elongated by the polymerase, the priming domain of TP rotates around a hinge that connects this domain to the rest of TP, until the priming domain is completely pushed out of the active site and the terminal protein dissociates (36).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Steitz TA. Current Opinion in Structural Biology 2004;14:4. [PubMed: 15102443]
2. Tang GQ, Roy R, Ha T, Patel SS. Mol Cell 2008;30:567. [PubMed: 18538655]
3. Ikeda RA, Richardson CC. J. Biol. Chem 1987;262:3790. [PubMed: 3546319]
4. Cheetham GMT, Steitz TA. Science 1999;286:2305. [PubMed: 10600732]
5. Martin CT, Muller DK, Coleman JE. Biochemistry 1988;27:3966. [PubMed: 3415967]
6. Briebe LG, Sousa R. EMBO J 2001;20:6826. [PubMed: 11726518]
7. Tahirou TH, et al. Nature 2002;420:43. [PubMed: 12422209]
8. Yin YW, Steitz TA. Science 2002;298:1387. [PubMed: 12242451]
9. Temiakov D, et al. Proc. Natl. Acad. Sci. U.S.A 2000;97:14109. [PubMed: 11095736]
10. Huang J, Sousa R. J. Mol. Biol 2000;303:347. [PubMed: 11031112]
11. Ma K, Temiakov D, Anikin M, McAllister WT. Proc. Natl. Acad. Sci. U.S.A 2005;102:17612. [PubMed: 16301518]
12. Guo Q, Nayak D, Briebe LG, Sousa R. J. Mol. Biol 2005;353:256. [PubMed: 16169559]
13. Bandwar RP, et al. J. Biol. Chem 2007;282:22879. [PubMed: 17548349]
14. Guillerez J, Lopez PJ, Proux F, Launay H, Dreyfus M. Proc. Natl. Acad. Sci. U.S.A 2005;102:5958. [PubMed: 15831591]
15. Mukherjee S, Briebe LG, Sousa R. Cell 2002;110:81. [PubMed: 12150999]
16. Bandwar RP, Tang GQ, Patel SS. J. Mol. Biol 2006;360:466. [PubMed: 16780876]
17. Pal M, Luse DS. Proc. Natl. Acad. Sci. U.S.A 2003;100:5700. [PubMed: 12719526]
18. Theis K, Gong P, Martin CT. Biochemistry 2004;43:12709. [PubMed: 15461442]
19. Turingan RS, Theis K, Martin CT. Biochemistry May 29;2007 46:6165. [PubMed: 17472344]
20. Materials and methods are available as supporting material on Science Online.
21. Cheetham GMT, Jeruzalmi D, Steitz TA. Nature 1999;399:80. [PubMed: 10331394]

22. Ikeda RA, Richardson CC. *Proc. Natl. Acad. Sci. U.S.A* 1986;83:3614. [PubMed: 3459146]
23. Place C, Oddos J, Buc H, McAllister WT, Buckle M. *Biochemistry* 1999;38:4948. [PubMed: 10213596]
24. Briebe LG, Gopal V, Sousa R. *J. Biol. Chem* 2001;276:10306. [PubMed: 11124963]
25. Montesana PE, Chin-Bow ST, Sousa R, McAllister WT. *J. Mol. Biol* 2000;302:1049. [PubMed: 11183774]
26. He B, Rong M, Durbin RK, McAllister WT. *J. Mol. Biol* 1997;265:275. [PubMed: 9018042]
27. Turingan RS, Liu C, Hawkins ME, Martin CT. *Biochemistry* 2007;46:1714. [PubMed: 17253774]
28. Nayak D, Guo Q, Sousa R. *J. Mol. Biol* 2007;371:490. [PubMed: 17580086]
29. Ma K, Temiakov D, Jianh M, Anikin M, McAllister WT. *Journal of Biological Chemistry* 2002;277:43206. [PubMed: 12186873]
30. Muller DK, Martin CT, Coleman JE. *Biochemistry* Apr 18;1989 28:3306. [PubMed: 2545254]
31. Murakami KS, Darst SA. *Curr Opin Struct Biol* Feb;2003 13:31. [PubMed: 12581657]
32. Hahn S. *Nat Struct Mol Biol* 2004;11:394. [PubMed: 15114340]
33. Chen HT, Hahn S. *Cell* 2004;119:169. [PubMed: 15479635]
34. Bushnell DA, Westover KD, Davis RE, Kornberg RD. *Science* 2004;303:983. [PubMed: 14963322]
35. Murakami KS, Masuda S, Darst SA. *Science* 2002;296:1280. [PubMed: 12016306]
36. Kamtekar S, et al. *EMBO J* Mar 22;2006 25:1335. [PubMed: 16511564]
37. We thank the staff of at the APS beamline 24-ID and at the NSLS beamline X25. We also thank Tyler Jorgensen for help with the supplemental movie and the staff of the CSB core facility at Yale University. This work was supported by NIH grant GM57510 to T.A.Steitz.

Figure 1.

A Comparison of the structures of T7 RNAP Initiation, Intermediate and Elongation Complexes. The molecules have been similarly oriented by superposition of their palm domains. The C-terminal domain is shown as a surface with the thumb domain removed to allow views of the DNA and RNA. The non-template strand is shown in light green, the template in blue, and the RNA transcript in red. Subdomains of the N-terminal domain are colored as follows: the C-helix is yellow, subdomain H is green and the PBD is purple. **A.** The initiation complex bound to the promoter has a 3nt transcript with its 5' end in contact with the PBD. **B.** The elongation complex exhibits a 220° right hand rotation of the PBD, a refolding of subdomain H onto the top of the polymerase, and the formation of an elongated C-helix

subdomain, when compared to the initiation complex. **C. Top:** The T7 RNAP in the 7nt RNA intermediate complex is bound to both promoter and downstream DNA. The PBD has rotated by 40° away from the C-terminal domain, avoiding a steric clash with the transcript and allowing for 7bp of heteroduplex to form in the active site. **Bottom:** A schematic drawing of the sequences constituting the subdomains of the N-terminal domain.

Figure 2.

Nucleic Acid Substrate. **A.** A 3Å resolution electron density omit map of the nucleic acid component of the active site calculated using 2Fo-Fc as coefficients and contoured at 1.0σ. A model of the template (blue) and the RNA (red) is superimposed. **B.** The sequence of the substrate observed in the 7nt intermediate complex structure. The polymerase has opened a 11nt transcription bubble. Shaded nucleotides are not seen in the electron density map.

Figure 3. Promoter and PBD movements during the transition. **A.** A view looking down onto the promoter bound to the PBD. A 40° rotation of the PBD away from the active site occurs around an axis that passes through a flexible loop of residues 198-204. The catalytic aspartic acid residues (D812, D537) represent the active site. **B.** The same view as in (A), without the PBD but showing the specificity loop, which also rotates.

Figure 4.

Other conformational changes associated with the transition **A**. A comparison of the C-helix subdomain conformations between the initiation (pink), intermediate (yellow) and elongation (blue) complexes. During the transition, the C2 helix rotates by 90°, resulting in its stacking on the C1 helix in the elongation phase, increasing the length of the C-helix by an additional 30Å. **B**. A 180° rotation around the peptide bond between M267 and F268 allows for the conformational changes in the N-terminal domain with little effect on the C-terminal domain. The elongation complex is shown in light blue, the initiation complex is in pink, residue 266 is highlighted in red. **C**. The same region shown in (B) but with the intermediate complex (dark blue) and the initiation complex (pink) superimposed. **D**. A superposition of the thumb domain from the initiation (pink), the intermediate (pale orange), and elongation (blue). The thumb backs away from the active site by 4Å during the transition and by another 5Å in the elongation complex, creating a binding cleft for the upstream DNA in the elongation complex.

Figure 5.

Arrangement of the downstream DNA. **A.** The downstream duplex is rotated by 30° towards the N-terminal domain compared to its position in the elongation complex (shown in grey). **B.** The angle between the upstream and downstream duplex DNA is at about 40° , bringing the phosphate backbones within $6\text{\AA}$ of each other. **C.** The refolded subdomain H from the structure of the elongation complex (grey) creates a clash with the position of the downstream DNA as observed in the intermediate complex. Residues involved in the clash are shown as spheres. **D.** A close-up view of the 5' end of the 7nt RNA reveals that a modeled extension of the RNA by 3 additional nts would clash with the specificity loop in the position observed in the intermediate complex.

Table 1

Crystallographic Data and Refinement Statistics

Space Group	P6 ₁	P2 ₁ 2 ₁ 2 ₁
Unit Cell dimensions (Å)	81.0 × 81.0 × 358.8	99.2 × 180.6 × 198.6
Resolution (Å)	50-2.9	50-6.7
R _{merge} ^{a,b}	8.1 (62.0)	15.2 (60.1)
I/σ ^c	17.7 (1.6)	7.1 (1.1)
Completeness(%)	93.6 (49.4) (3.0Å 88.9)	87.8 (39.7)
Unique reflections	27,490	6,016
Redundancy	6.4 (2.5)	3.5 (1.7)
Copies in AU	1	2
Phasing resolution (Å)	50-3.0	50-6.7
Rmsd bond length (Å)	0.006	
Rmsd bond angle (°)	1.205	
R _{cryst} (%)	25.0	40.0
R _{free} (%)	29.2	44.2
PDB ID	To be submitted	To be submitted

^aNumbers in parentheses correspond to the highest resolution shell.

^bR_{merge} is $\sum_j |I_j - \langle I \rangle|$, where I_j is the intensity of an individual reflection and $\langle I \rangle$ is the mean intensity for multiple observations of symmetry-related reflections.

^cI/σ is the mean Intensity divided by the standard deviation.