Cell, Vol. 109, 417–420, May 17, 2002, Copyright ©2002 by Cell Press

Views of Transcription Initiation

Minireview

Brian A. Young,[1] Tanja M. Gruber,[2] and Carol A. Gross[2,3]
[1]Department of Biochemistry and Biophysics
[2]Departments of Stomatology and Microbiology and Immunology
University of California, San Francisco
San Francisco, California 94143

Initiation of transcription is the first step in gene expression and a major point of regulation. Recent structural studies reveal the nature of the initiating complex and suggest new ways of accomplishing the processes required for initiation.

RNA polymerase (RNAP), the enzyme that carries out transcription, is a remarkable molecular machine. During initiation, it must recognize promoter DNA from the vast excess of non-promoter DNA, separate the duplex to expose the template strand, and initiate RNA synthesis using only mononucleotides. Before beginning processive elongation, it must transition to a non-sequence-specific DNA binding protein that moves forward (and in some cases backward) along the DNA. Four recent reports illuminate these processes. Three use crystallography to provide structural information about the prokaryotic initiation factor σ (Campbell et al., 2002), and the initiating form of prokaryotic RNAP without (Murakami et al., 2002b) and with promoter DNA (Murakami et al., 2002a). A fourth provides a distance constrained model of initiating RNAP and its interaction with promoter DNA based upon systematic measurements of fluorescence resonance energy transfer (FRET) of probes located throughout the initiating RNAP and in the DNA (Mekler et al., 2002). We discuss these results and the insights and speculations they provoke about how this machine accomplishes these complex processes.

Structure

Transcription initiation in prokaryotes is carried out by holoenzyme (*holo*), comprising core RNAP (*core*) plus the initiation specific subunit, σ. Core is an ~400 kDa complex of five subunits ($\alpha_2\beta\beta'\omega$), which shares considerable sequence and even more structural homology with its eukaryotic counterparts (e.g., RNAP II), whereas σ has little sequence homology to its eukaryotic counterparts, the general transcription factors. Holo first recognizes the two conserved hexamers in the promoter, located at −10 and −35 relative to the transcription startpoint of +1, then melts the DNA from −11 to +4 to form the *open complex*, and then begins synthesizing the nascent RNA. The three sections below summarize the structures of σ, holo, and the open complex.

σ *Subunit*. All bacteria have one primary σ factor, which directs the majority of transcription. σs have four conserved regions, which mediate binding to core and to DNA (Figure 1). As the structure of only one σ domain (σ_2: region 1.2–2.4 from *E. coli*) had been described (Malhotra et al., 1996), Campbell et al. (2002) tried to crystallized σ^A, the primary σ of the thermophile *Thermus aquaticus (Taq)*. This proved impossible but serendipitous protease contamination produced crystallizable fragments diffracting to ~2 Å.

σ^A has three stably folded domains, σ_2, σ_3 and σ_4, connected by flexible linkers. Each domain is predicted to bind both core and DNA (Figure 1). σ_2 is essentially identical to *E. coli* σ_2, with an exposed region 2.2 helix predicted to form a primary interface with core and the region 2.3–2.4 helix, which recognizes the −10 element and contains aromatic residues important for melting and recognition of the non-template strand of the −10 element. Both σ_3 and σ_4 are comprised of three helices. One helix in σ_3 is responsible for recognizing two conserved bases located upstream of the −10 region, present in "extended −10 promoters," which do not need a −35 promoter element. Two helices in σ_4 form an HTH motif; one of these helices recognizes the −35 region of the promoter. Campbell et al. were also able to obtain the structure of σ_4 complexed with a −35 element, allowing the first high-resolution view of promoter recognition. This pivotal work defined the domain structure of σ, provided assurance that the genetic inferences about how σ recognized the −35 element were generally correct, and produced high-resolution structures that allowed definitive placement of σ on the holo structure.

Holoenzyme. Two studies provide our first glimpse of the structure of an initiation competent multi-subunit RNAP. A 4 Å electron density map for *Taq* holo, combined with structures of core (Zhang et al., 1999) and portions of σ^A (Campbell et al., 2002) allowed Murakami et al. (2002b) to provide a structure of holo and a first view of a (nearly) intact σ. FRET experiments measuring distances between probes in σ^{70} and core allowed Mekler et al. (2002) to also place σ^{70} on core by assuming core probes are largely stationary upon holo formation. The crystallographic work gives high-resolution structural information, while the FRET analysis provides information on region 1.1, which was absent in the *Taq* holo crystals. Overall, the crystal structure and the FRET-based model agree with one another and with previous biochemical and genetic evidence (Gross et al., 1998). When describing these results, we use the term *upstream* to refer to DNA before the start site (−1 to −∞) and *downstream* to refer to DNA after the start site (+1 to +∞).

To put the holoenzyme structure in context, we first revisit the model of elongating prokaryotic RNAP (Figure 2A), derived by combining the structure of core with crosslinking studies that place nucleic acids on the structure (Korzheva et al., 2000; Zhang et al., 1999). RNAP can be crudely described as a crab claw whose active site is positioned at the base of its two pincers. Downstream DNA, located in an internal channel formed between the pincers (also called the *jaws*), separates into its two strands near the active site. The strands turn upward (relative to the plane of the page in Figure 2A), taking different paths through the polymerase and rean-

[3]Correspondence: cgross@cgl.ucsf.edu

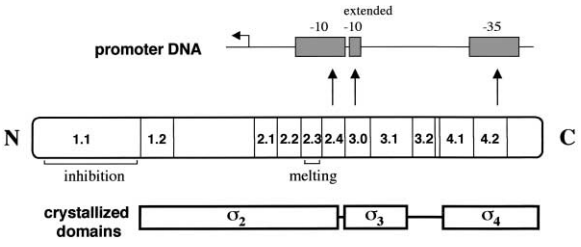

Figure 1. Conserved regions of σ

neal to form the upstream duplex, which is at a right angle to the downstream DNA. Nascent RNA follows the template strand for about 9 bases and then exits the polymerase underneath a flap that juts out from the bottom of the pincers. Studies of multiple crystals of polymerase indicate that the pincers and the flap are mobile. The particularly flexible top pincer, called the *clamp*, is derived primarily from a portion of the β′ subunit. The bottom pincer (with two independently mobile modules, β1 and β2) and the flap are derived from the β subunit. Flexibility is presumably required for the conformational changes necessary to accommodate steps in transcription (Darst et al., 2002 and references therein).

The Murakami et al. structure (2002b) provides evidence of the importance of these mobile features (Figure 2B). The three domains of σ are spread out across one face of core, each interacting with and altering the position of a mobile domain of ββ′ relative to its position in core. Since σ_2 interacts with the β′ clamp in the upper pincer and σ_3 interacts with β1 in the lower pincer, these interactions can modulate opening and closing of the downstream DNA channel. Likewise, interaction of σ_4 with the β flap can alter the RNA exit channel. Additionally, several regions disordered in core become ordered upon interaction with σ, including structures homologous to the zipper and the lid of RNAP II, which may be important in guiding nucleic acids through RNAP (Gnatt et al., 2001). If the zipper and lid structures are maintained in elongating bacterial RNAP, as they are in elongating RNAP II, they will have different interacting partners than they do in the holoenzyme or in the initiation complex, and may therefore be key focal points for the transition between initiating and elongating RNAP.

σ is properly positioned in holo to bind promoter DNA. Its DNA binding determinants are solvent exposed, with a spacing roughly consistent with that expected from the position of the DNA elements to which they bind. For example, σ_2 (−10 recognition) and σ_4 (−35 recognition) are separated by about 76 Å, roughly the distance between the middle of −10 region and the −35 region in B-form DNA (Murakami et al., 2002b). One remarkable feature of σ that allows such separation is the σ_3–σ_4 linker. Its 33 amino acids, derived mostly from region 3.2, traverse the 45 Å that separates σ_3 and σ_4, skirting the active site and passing through core's RNA exit channel before connecting with σ_4.

FRET analysis locates σ region 1.1 in the downstream DNA channel (Mekler et al., 2002); a similar conclusion is reached by the Murakami et al. (2002b) more indirectly. Region 1.1 must move from this position to form an open complex, and the next set of structures described indicates that this is the case.

Clearly, the interface between σ and core is crucial for reconfiguring both partners for initiation; some notion of its importance comes from the realization that ~8500 Å^2 are buried, almost twice that of the largest reversible protein/protein interface known, and actually more comparable to interaction surfaces in oligomeric proteins (~10,000 Å^2) (Murakami et al., 2002b).

Open Complex. The same two groups investigate the structure of the open complex. Mekler et al. (2002) used FRET analysis to determine the positions of σ and down-

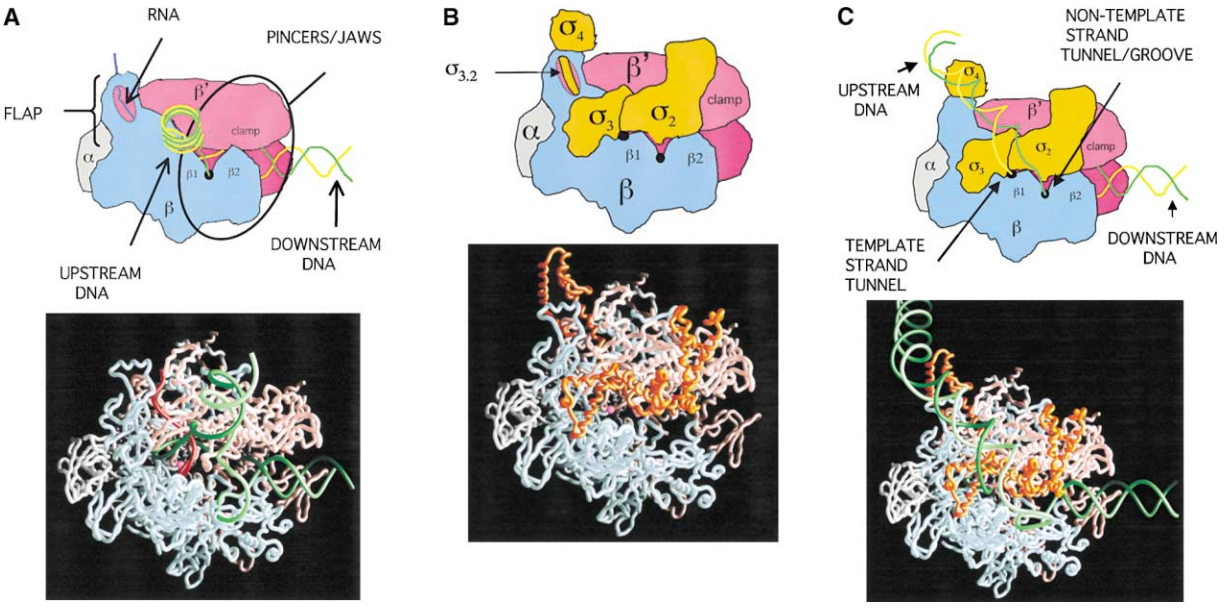

Figure 2. RNA Polymerase Structures

(A) Elongating core. (B) Holo. (C) Open complex. (Lower panels kindly provided by S. Darst.)

stream DNA in the open complex. Their major finding is that region 1.1 of σ moves from inside to outside the active site channel.

Murakami et al. (2002a) crystallized holo bound to a well-characterized open complex mimic—the fork junction, a fragment of upstream promoter DNA that is double stranded from −41 to −12, with the non-template strand continuing in single stranded form to −7. With the exception of region 1.1 noted above, they find the positions of σ are similar in holo and the open complex. Fork junction DNA lies across the upstream face of holo, with essentially all sequence-specific recognition carried out by σ (Figure 2C). Predicted contacts in the −10 region validate previous genetic assignments: the −12 position is very close to the region 2.4 amino acids believed to contact this base pair; the non-template strand is draped over the region 2.3 aromatic residues implicated in their recognition; and the extended −10 recognition determinants are very close to the region 3.0 amino acids believed to contact them. In contrast, most likely as a result of crystal packing artifacts, σ_4 is slightly mispositioned from the −35 region placement that is predicted genetically and observed in the structure of the σ_4 −35 promoter region complex (Campbell et al., 2002).

Both groups use the current information and previous data (Gnatt et al., 2001; Korzheva et al., 2000; Naryshkin et al., 2000) to generate similar models of RNAP open complexes (Figure 2C). The upstream DNA drapes over the domains of σ. At −11 the template and non-template strands separate. The template strand turns sharply down into the active site, becoming completely enclosed in a positively charged protein tunnel formed of portions of σ, β, and β′, and framed by universally conserved basic residues of σ. The non-template stand continues to interact with σ until about −7 and then bends down between two lobes of the β subunit (β1 and β2). The strands reanneal at about +5, and this downstream duplex continues through the main channel in the tunnel formed by the jaws, just as in the structure of the elongation complex. Overall, a sharp bend is introduced between the upstream DNA and the downstream DNA.

Function

Promoter Recognition and Spacer Accommodation. The work presented here goes some way toward explaining how binding to core relieves region 1.1 autoinhibition of DNA binding in free σ. In holo, region 1.1 is removed from its location in σ and placed in the active site channel of polymerase, possibly because its high negative charge allows it to act as a downstream DNA mimic. This same idea could explain autoinhibition: region 1.1 might bind DNA recognition determinants in free σ because of its negative charge, thereby out-competing promoter DNA.

The structures also suggest a mechanism by which RNAP binds promoters with as few as 16 or as many as 18 nucleotides in the spacer region between the −35 and −10 elements, which can change the distance between these elements by as much as 10 Å. σ_4 (−35 recognition), is perched on the end of the flexible flap of β. Shifting the angle at which this flexible flap juts out of the core enzyme alters the distance between σ_4, and σ_2 (−10 recognition) somewhat. Larger lengths of DNA can be accommodated by stretching (or "kinking") them over a bulge in β′ that intervenes between the domains in σ that recognize the −10 and −35 hexamers. The enhanced DNaseI hypersensitivity in the spacer region of promoters with longer spacers is consistent with this explanation, as such "kinked" DNA would be expected to be more susceptible to DNase I cleavage (Murakami et al., 2002a).

Region 1.1 and Open Complex Formation. Region 1.1 also plays a positive role in transcription initiation: holo having σ^{70} lacking region 1.1 forms open complexes very slowly at several promoters (Gross et al., 1998 and references therein). The jaws are closed in the holo structure (which was obtained with σ^A lacking region 1.1) but must be open in wild-type holo to permit downstream DNA to enter. Thus, Murakami et al. (2002b) propose that opening the jaws is a rate-limiting step at some promoters, and that region 1.1 accelerates this opening by binding between them, thereby accelerating open complex formation. At one weak promoter, holo lacking region 1.1 forms melted complexes more readily than intact holo (Vuthoori et al., 2001); perhaps, as suggested by Mekler et al. (2002), this is because the rate limiting step at this promoter is ejecting region 1.1 from the jaws.

A great deal about the relationship between σ and jaw opening remains to be worked out. Region 1.1 is conserved only in primary or housekeeping σ factors; how are the jaws of the polymerase opened during initiation with alternate σ factors, which lack region 1.1? Additionally, the notion that σ opens the jaws of polymerase currently lacks experimental backing. The most recent analysis of *E. coli* core RNAP indicates that its jaws are wide open (Darst et al., 2002). Although this could be artifactual, it is worthwhile recalling that σ (and TFIIF in eukaryotes) decreases binding to non-promoter DNA. Do they do this by partially closing the jaws? Even after solution measurements of the placement of the jaws in core and holo resolve this particular question, additional work is clearly needed to understand how downstream DNA is efficiently placed in the jaws during open complex formation.

De Novo RNA Synthesis. In contrast to DNA polymerases, which can only extend existing nucleic acid chains, RNAP is able to create a nucleic acid polymer de novo—using only mononucleotides. The initiation step is difficult: both incoming nucleotide and the initiating nucleotide, which attacks the incoming NTP, must be stabilized in the correct geometry. This is especially difficult because RNAP prefers to begin RNA chains with an ATP (which base pairs with the template strand more weakly than would CTP or GTP).

How does RNAP accomplish de novo synthesis? Specific interactions with the initiating nucleotide must hold it rigidly in place, facilitating chemical attack on the incoming nucleotide. Such specific interactions would explain why polymerase prefers to start transcripts with ATP (followed by GTP, UTP, and then CTP). Indeed, a subcomplex of core polymerase, $\alpha_2\beta$, and possibly even the isolated β subunit has a site for the initiating nucleotide (Naryshkina et al., 2001). Darst now suggests that a disordered loop of σ near the beginning of the σ_3–σ_4 linker, pointing toward the active site, assists in binding the initiating nucleotide. They tested this idea, using

an extended −10 promoter, which does not use −35 recognition determinants and can therefore be transcribed by holo ending at σ_3, which lacks the disordered loop. Holo with this truncated σ requires a much higher concentration of initiating dinucleotide to reach maximal levels of transcription than does holo with full-length σ (Campbell et al., 2002). Using σ to provide specificity is appealing. A marked preference for a particular nucleotide in the attacking site might have disagreeable side effects when elongating a transcript. If σ performed this function, the selectivity required for de novo synthesis would be present only at initiation.

We suggest another possible role for this disordered loop of σ-stabilization of the melted state of the promoter by binding to the template strand near the start site in single stranded form (thus keeping it from reannealing). This idea comes from a consideration of the ribosomal RNA (rRNA) promoters. rRNA promoters cannot form stable open complexes; they require high levels of initiating nucleotide to stabilize the melted state required for efficient transcription. This requirement is an important regulation mechanism in the cell, and is mediated in part by the presence of a stretch of CG bases, called the discriminator region, near the start site of rRNA promoters. The unique nucleotides in the discriminator may prevent the stabilizing interactions between the disordered loop in σ and the template strand, thereby preventing stable open complex formation. In this context, an alternate single-strand-specific interaction may be necessary to achieve a stable open complex. The bridge created between the template strand and RNAP by the initiating nucleotide may provide this interaction—thus the requirement for high initiating nucleotides. Likewise, removal of the disordered loop in σ should prevent stable open complex formation and create a requirement for high initiating nucleotides.

Promoter Clearance and Abortive Initiation. All RNAPs reiteratively synthesize and release short RNA transcripts called abortives, ($\sim$2 to 9 nucleotides in length). Based on the structure of elongating eukaroytic RNAP II, Kornberg proposed that shorter nascent RNAs dissociate because they make fewer contacts with polymerase than do longer RNAs (Gnatt et al., 2001). Another explanation for these transcripts has emerged from the Murakami et al. structure (2002b). Region 3.2 occupies the RNA exit channel, leading to the speculation that nascent RNAs must successfully compete with region 3.2 to be retained in elongating polymerase. When RNA transcripts lose the competition, they are ejected as abortive transcripts; when they win, region 3.2 is ejected and the transcript is successfully elongated. Consistent with this idea, holo having σ lacking region 3.2 produced fewer abortives relative to full-length transcript than does holo with wt σ (Murakami et al., 2002b). If this is a universal explanation for abortive transcription, a general transcription factor must play this role in eukaryotic initiation.

Why might polymerase place region 3.2 of σ right where the RNA must go, thus wasting valuable NTP energy synthesizing abortive RNAs? Maybe the competition between σ and RNA is an important part of the promoter clearance process. During promoter clearance, the polymerase must extricate itself from promoter-specific contacts so it can processively elongate trancripts. Interestingly, promoter clearance tends to coincide with the end of abortive synthesis and could be set in motion by release of region 3.2 from core by the successfully elongating RNA chain. Release of region 3.2 could cause promoter clearance because it weakens the σ/core interface, thus allowing core to dissociate from σ. Alternatively, either movement of region 3.2 out of the channel or of RNA into the channel could alter the position of σ_4 (which is perched on the flap surrounding the channel), making correct interactions with the −35 element impossible and promoting promoter dissociation. In either case, the NTP energy utilized during abortive initiation may be a small price to pay in order to switch from specific interaction with the promoter to processive elongation.

Conclusion

These first glimpses of the initiation competent polymerase provide extraordinary insight into the functions the machine performs. Although these structures do not provide final answers for how these processes are accomplished, they do allow us to conceptualize concrete models. No doubt these first peeks into the structure of initiating RNAP will motivate an enormous number of future experiments to test these ideas.

Selected Reading

Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Mol. Cell *9*, 527–539.

Darst, S.A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G., and Wriggers, W. (2002). Proc. Natl. Acad. Sci. USA *99*, 4296–4301.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Science *292*, 1876–1882.

Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., and Young, B. (1998). Cold Spring Harb. Symp. Quant. Biol. *63*, 141–155.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S.A. (2000). Science *289*, 619–625.

Malhotra, A., Severinova, E., and Darst, S.A. (1996). Cell *87*, 127–136.

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R., and Ebright, R.H. (2002). Cell *108*, 599–614.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002a). Science, in press.

Murakami, K.S., Masuda, S., and Darst, S.A. (2002b). Science, in press.

Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., and Ebright, R.H. (2000). Cell *101*, 601–611.

Naryshkina, T., Mustaev, A., Darst, S.A., and Severinov, K. (2001). J. Biol. Chem. *276*, 13308–13313.

Vuthoori, S., Bowers, C.W., McCracken, A., Dombroski, A.J., and Hinton, D.M. (2001). J. Mol. Biol. *309*, 561–572.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Cell *98*, 811–824.