determine the magnitude of the error in k_f determination. The magnitude may again vary from promoter to promoter. This raises the serious possibility that in some cases the kinetic constants derived from abortive initiation experiments may require significant correction.

A second problem comes from the fact that in most preparations of RNA polymerase, only a fraction is capable of binding to the promoter. This is generally true of DNA-binding proteins. The inactive fraction for RNA polymerase varies. Schlax and co-workers reported an active fraction of about 40%. Some commercial preparations showed about 20% active fractions (S. Roy, unpublished observation). Clearly, in such cases the kinetic parameters, particularly K_B, will be affected directly and need to be corrected for active fractions.

[47] Probing the Role of Region 2 of *Escherichia coli* σ^{70} in Nucleation and Maintenance of the Single-Stranded DNA Bubble in RNA Polymerase-Promoter Open Complexes

By PIETER L. DEHASETH and LAURA TSUJIKAWA

The bacterial RNA polymerase (RNAP), referred to as the ''holoenzyme,'' or Eσ, is a complex of the ''core enzyme'' (also core or E) and sigma factor (σ), a type of bacterial initiation factor.[1–4] The multisubunit core enzyme shows similarities in both sequence and structure to the multisubunit eukaryotic RNAP.[5–8] The role of the sigma factor is to recognize promoter DNA and to initiate the process of RNAP-induced strand separation to form the transcription-competent open promoter complex (see Refs. 3, 9, and 10 for reviews). The latter is the focus of this article.

[1] R. Burgess, A. Travers, J. J. Dunn, and E. K. F. Bautz, *Nature* **221,** 43 (1969).
[2] W. R. McClure, *Annu. Rev. Biochem.* **54,** 171 (1985).
[3] C. Gross, C. Chan, A. Dombroski, T. Gruber, M. Sharp, J. Tupy, and B. A. Young, *Cold Spring Harb. Symp. Quant. Biol.* **63,** 141 (1998).
[4] P. L. deHaseth, M. Zupancic, and M. T. Record, Jr., *J. Bacteriol.* **180,** 3019 (1998).
[5] L. A. Allison, M. Moyle, M. Shales, and C. J. Ingles, *Cell* **42,** 599 (1985).
[6] D. Sweetser, M. Nonet, and R. A. Young, *Proc. Natl. Acad. Sci. USA* **84,** 1192 (1987).
[7] G. Zhang, E. A. Campbell, L. Minakhin, C. Richter, K. Severinov, and S. A. Darst, *Cell* **98,** 811 (1999).
[8] E. Ebright, *J. Mol. Biol.* **304,** 687 (2000).

0076-6879/03 $35.00

A major advance in our understanding of the transcriptional apparatus was achieved with the determination of the structure of the bacterial sigma factor,[11] core enzyme,[7] holoenzyme,[12,13] and a cocomplex of holoenzyme with a forked DNA[14] (see later) encompassing most of the sequence recognized by RNAP. These structures are of great importance to our understanding of the interaction of RNAP with promoter DNA[14,15] (see also Young *et al.*[16]). Relevant features of these structures include (1) the large total area of contact between core and σ^{70}, (2) DNA-binding groups of sigma factor that stick out from the body of the holoenzyme, and (3) the existence of separate tunnels for each strand of promoter DNA, leading the template strand deep into the hydrophobic core of the RNAP to the active site of the enzyme while keeping the nontemplate strand much closer to the surface.

RNAP–Promoter Interactions

In addition to the "primary" σ factor responsible for the recognition of promoters for housekeeping genes (σ^{70} in *Escherichia coli*), most bacteria have multiple minor sigma factors that help the cell deal with adverse conditions. Across eubacteria, the promoters recognized by holoenzymes containing the primary sigma factors are similar to the consensus σ^{70} promoter of *E. coli*. The −35 region (consensus TTGACA) is recognized by amino acids in σ^{70} region 4.2,[3,9] the "TG" ("extended −10") sequence by amino acids in region 3.0 (also referred to as region 2.5[17]), and the −12 T and −11 A of the −10 region (consensus TATAAT) by amino acids in region 2.4.[3,9] While it is likely that the TAAT sequence in the −10 region from −10 to −7 is recognized as single-stranded (SS) DNA,[18–21] it is not yet clear which amino acids are responsible.

[9] M. Lonetto, M. Gribskov, and C. A. Gross, *J. Bacteriol.* **174,** 3843 (1992).
[10] J. D. Helmann and J. Chamberlin, *Annu. Rev. Biochem.* **57,** 839 (1988).
[11] E. A. Campbell, O. Muzzin, M. Chlenov, J. L. Sun, C. A. Olson, O. Weinman, M. L. Trester-Zedlitz, and S. A. Darst, *Mol. Cell* **9,** 527 (2002).
[12] K. S. Murakami, S. Masuda, and S. A. Darst, *Science* **296,** 1280 (2002).
[13] D. G. Vassylyev, S.-I. Sekine, O. Laptenko, J. Lee, M. N. Vassylyeva, S. Borukhov, and S. Yokoyama, *Nature* **417,** 712 (2002).
[14] K. S. Murakami, S. Masuda, E. A. Campbell, O. Muzzin, and S. A. Darst, *Science* **296,** 1285 (2002).
[15] V. Mekler, E. Kortkhonjia, J. Mukhopadhyay, J. Knight, A. Revyakin, A. N. Kapanidis, W. Niu, Y. W. Ebright, R. Levy, and R. H. Ebright, *Cell* **108,** 599 (2002).
[16] B. A. Young, T. M. Gruber, and C. Gross, *Cell* **109,** 417 (2002).
[17] K. A. Barne, J. A. Bown, S. J. W. Busby, and S. D. Minchin, *EMBO J.* **16,** 4034 (1997).
[18] C. W. Roberts and J. W. Roberts, *Cell* **86,** 495 (1996).
[19] J. Qiu and J. D. Helmann, *Nucleic Acids Res.* **27,** 4541 (1999).

Formation of the open complex from free RNAP and promoter DNA requires conformational changes in both DNA and the RNAP.[22] McClure[23] showed that the RNAP concentration dependence of the kinetics of open complex formation allowed the process to be subdivided into a concentration-dependent step (closed complex formation) and a concentration-independent ''isomerization'' step (open complex formation). Subsequent work[24,25] revealed the presence of additional intermediates on the pathway to open complex formation. Steps in formation of an open complex are usually described as ''bind, nucleate, melt.''[26] A kinetic scheme with the currently recognized intermediates in the process is shown as[27,28]

$$R + P \rightleftharpoons I_0 \rightleftharpoons I_1 \rightleftharpoons I_2 \rightleftharpoons RP_o$$

SCHEME 1

where R is RNAP; P is promoter DNA; and I_0 (= RP_c) is the initial closed complex (RNAP–DNA contacts at the −35, TG and −10 sequences, but no strand separation). I_1 and I_2 are significant kinetic intermediates at the well-studied promoter P_R. I_2 but not I_1 was found to be resistant to a short challenge by the anionic competitor heparin.[27] RP_o is the open complex.

Nucleation (= initiation) of strand separation is likely driven by binding free energy. The challenge is to correlate this scheme with structural studies.[29] As proposed by Saecker *et al*., unfavorable conformational changes in I_1 have been introduced in both RNAP and promoter DNA. The distortion of DNA in I_1 by bending, twisting, or unwinding[30] would facilitate nucleation at −11A, possibly by flipping this base out of the helix[31–34] to generate the intermediate I_2. This would be similar to the structures

[20] D. L. Matlock and T. Heyduk, *Biochemistry* **39,** 12274 (2000).
[21] M. S. Fenton and J. D. Gralla, *Proc. Natl. Acad. Sci. USA* **98,** 9020 (2001).
[22] R. S. Spolar and M. T. Record, Jr., *Science* **263,** 777 (1994).
[23] W. R. McClure, *Proc. Natl. Acad. Sci. USA* **77,** 5634 (1980).
[24] H. Buc and W. R. McClure, *Biochemistry* **24,** 2712 (1985).
[25] J. H. Roe, R. R. Burgess, and M. T. Record, Jr., *J. Mol. Biol.* **176,** 495 (1984).
[26] M. T. Record, Jr., W. S. Reznikoff, M. L. Craig, K. L. McQuade, and P. J. Schlax, ''*Escherichia coli* and *Salmonella,* Cellular and Molecular Biology'' (F. C. Neidhardt, editor in chief), p. 792. ASM Press, Washington, DC, 1996.
[27] M. L. Craig, O. V. Tsodikov, K. L. McQuade, J. P. E. Schlax, M. W. Capp, R. M. Saecker, and M. T. Record, Jr., *J. Mol. Biol.* **283,** 741 (1998).
[28] R. M. Saecker, O. V. Tsodikov, K. L. McQuade, P. E. Schlax, M. W. Capp, and M. T. Record, Jr., *J. Mol. Biol.* **319,** 649 (2002).
[29] K. S. Murakami and S. A. Darst, *Curr. Opin. Struct. Biol.* **13,** 31 (2003).
[30] P. L. deHaseth and J. D. Helmann, *Mol. Microbiol.* **16,** 817 (1995).
[31] J. D. Helmann and P. L. deHaseth, *Biochemistry* **37,** 5959 (1999).
[32] M. S. Fenton, S. J. Lee, and J. D. Gralla, *EMBO J.* **19,** 1130 (2000).

characterized for, among others, the repair enzyme uracil *N*-glycosylase; accompanied by major changes in the conformation of the enzyme, the U is rotated completely out of the helix in a process called base flipping.[35–39] The nucleation step is likely rate limiting in open complex formation; for the P_R promoter, in both the forward and the reverse directions, the slowest step was found to be the interconversion between I_1 and I_2.[27] Several experimental approaches have established a crucial role for the −11A in formation of an open complex.[20,31,34,40,41] Propagation of strand separation to yield RP_o involves unzipping of DNA from the −11A in a downstream direction; concomitant conformational changes in the RNAP position each strand in its own tunnel.

Sigma Factor Region 2.3 and Nucleation of DNA Strand Separation

Sequence comparison of sigma factors has revealed the existence of four regions of significant sequence conservation.[9,42] This section focuses on region 2, which has been subdivided into 2.1, 2.2, 2.3, and 2.4. Some residues in region 2.1 and many in region 2.2 are involved in recognition of the core.[43] Region 2.4 carries out recognition of the −10 element (reviewed in Lonetto *et al.*[9]) and region 2.3 nucleation of strand separation, and likely also recognition of bases in the −10 element.[10,33,44–46]

In *E. coli*, region 2.3, extending from residues 417 to 434, has the sequence DKFEYRRGYKFSTYATWW. A comparison of the amino acid sequences of 53 primary sigma factors (supplementary material to

[33] M. Tomsic, L. Tsujikawa, G. Panaghie, Y. Wang, J. Azok, and P. L. deHaseth, *J. Biol. Chem*. **276,** 31891 (2001).
[34] L. Tsujikawa, M. G. Strainic, H. Watrob, M. D. Barkley, and P. L. deHaseth, *Biochemistry* **41,** 15334 (2002).
[35] Y. L. Jiang and J. T. Stivers, *Biochemistry* **41,** 11236 (2002).
[36] P. M. Patel, M. Suzuki, E. Adman, A. Shinkai, and L. A. Loeb, *J. Mol. Biol*. **308,** 823 (2001).
[37] G. Xiao, M. Tordova, J. Jagadeesh, A. C. Drohat, J. T. Stivers, and G. L. Gilliland, *Proteins* **35,** 13 (1999).
[38] T. Hollis, Y. Ichikawa, and T. Ellenberger, *EMBO J*. **19,** 758 (2000).
[39] D. R. Davies, I. Y. Goryshin, W. S. Reznikoff, and I. Rayment, *Science* **289,** 77 (2000).
[40] Y. Guo and J. D. Gralla, *Proc. Natl. Acad. Sci. USA* **95,** 11655 (1998).
[41] H. M. Lim, H. J. Lee, S. Roy, and S. Adhya, *Proc. Natl. Acad. Sci USA* **98,** 14849 (2001).
[42] M. Gribskov and R. R. Burgess, *Nucleic Acids Res* **14,** 6745 (1986).
[43] M. M. Sharp, C. L. Chan, C. Z. Lu, M. T. Marr, S. Nechaev, E. W. Merritt, K. Severinov, J. W. Roberts, and C. A. Gross, *Genes Dev*. **13,** 3015 (1999).
[44] Y.-L. Juang and J. D. Helmann, *J. Mol. Biol*. **235,** 1470 (1994).
[45] X. Huang, F. J. Lopez de Saro, and J. D. Helmann, *Nucleic Acids Res*. **25,** 2603 (1997).
[46] G. Panaghie, S. E. Aiyar, K. L. Bobb, R. S. Hayward, and P. L. deHaseth, *J. Mol. Biol*. **299,** 1217 (2000).

Campbell *et al.*[11]) confirmed the previously observed extensive sequence conservation in region 2.3.[9] Seven residues (F419, K426, F427, Y430, A431, W433, and W434 in the *E. coli* numbering used throughout this proposal) were invariant among these 53 sigma factors. Flanking the 2.3 region, in 2.2, residue K414, and in 2.4, residues Q446, R448, and R451 also are conserved for all sigma factors of the data set. The 2.3 regions of many minor sigma factors show much less conservation when compared to each other or to the primary sigma factors.

Amino acid residues Y430 and W433 are seen in the cocrystal structure[14] to be positioned close to the double-stranded (ds) – single-stranded junction of the forked DNA template, and thus are likely candidates involved in the nucleation of DNA melting. The following model[33] is supported by currently available data. In I_o (Scheme 1), both sequence-specific (H-bonding) and sequence-nonspecific (electrostatic) interactions serve to hold promoter DNA in a unique position with respect to aromatic amino acids Y430 and W433. These two residues likely cooperate to flip the −11A out of the DNA helix to nucleate the DNA-melting process. Several basic residues (K414, K418, R423, and K426) are likely involved in promoter DNA binding. The structure of the *E. coli* region 2.3[47] reveals that the side chains of these four residues and of Y425, Y430, W433, and W434 stick out on approximately the same face of the protein, where they can interact with promoter DNA.

Use of Forked DNAs to Dissect Steps in Formation of Open Complexes

Guo and Gralla[40] pioneered the use of forked DNAs as model templates for the study of DNA and protein sequences required for recognition by and stable complex formation with RNAP.[20,21,32–34,48] For these templates, the DNA from the −35 region through the first base pair of the −10 region is double stranded and the remainder is present as a single-stranded extension of the nontemplate strand. The interaction of RNAP with forked templates is a diagnostic tool for the ability to nucleate strand separation. Differentiation between binding and isomerization can be accomplished by subjecting RNAP–forked DNA complexes to a challenge with the poly anion, heparin (see later).

[47] A. Malhotra, E. Severinova, and S. A. Darst, *Cell* **87,** 127 (1996).
[48] L. Tsujikawa, O. V. Tsodikov, and P. L. deHaseth, *Proc. Natl. Acad. Sci USA* **99,** 3493 (2002).

Experimental Procedures

The experimental approach described here allows a quantitative evaluation of the extent to which a substitution in σ factor affects steps in the formation of an open complex or its stability. This can be achieved by determination of (1) the kinetics of open complex formation at promoters, (2) the affinity between RNAP and a forked DNA, (3) the stability of RNAP–forked DNA complexes, and (4) the extent of binding of ssDNA at a fixed concentration of RNAP. See Fig. 1 for sequences of the aforementioned DNAs.

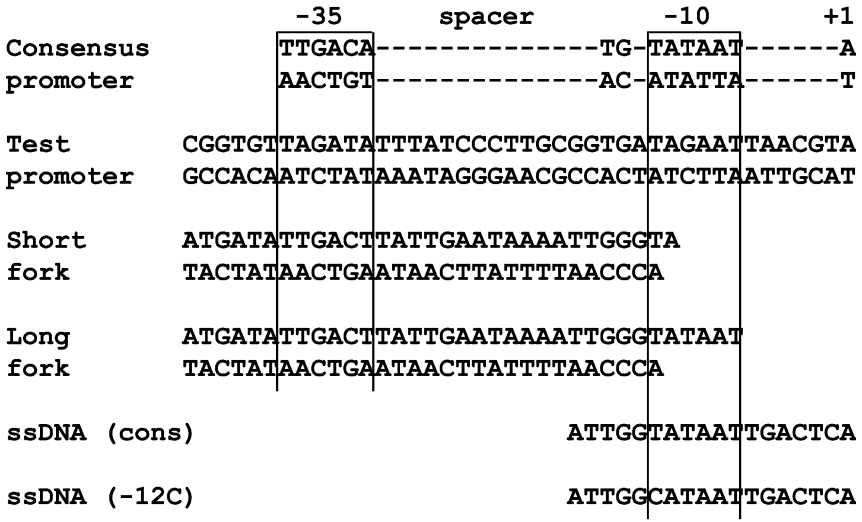

FIG. 1. Promoter, fork junction, and single-stranded DNA sequences. For the consensus *E. coli* σ^{70} promoter, only the sequence of the core promoter region, from the −35 region to the start site (+1), is shown. Lack of significant sequence conservation is indicated by hyphens. The test promoter is a variant of the P_{RM} promoter of bacteriophage λ. It forms open complexes at a rate that lends itself well to the use of manual methods. The two fork junction or forked templates have the identical bottom strand (template strand of the promoter from which the sequences were derived), but differ in the length of the 3′ extension of the top (nontemplate) strand. The short fork has an overhanging A, corresponding to the −11A of the −10 region of promoter DNA, whereas the overhang of the long fork extends to the 3′ end of the −10 region. The two single-stranded DNAs have the −10 sequence of the nontemplate strand of promoter DNA. The −12C oligodeoxyribonucleotide has a nonconsensus base at the −12 position, significantly weakening the specific interaction with the RNA polymerase.

Reagents and Techniques

RNAP. Purification of histidine-tagged σ factor is as described previously.[46] Reconstitution of holoenzyme from core polymerase (purchased from EpicentreTechnologies) and purified σ factor is accomplished by incubating the core (650 n*M*) and sigma factor (3.25 μM) in 24 μl storage buffer [0.01 *M* Tris–HCl, pH 7.9, 0.1 *M* NaCl, 0.1 m*M* EDTA, 0.1 m*M* dithiothreitol (DTT), and 50% glycerol for 30 min on ice. The reconstituted holoenzyme is used without further purification. In this article, all RNAP concentrations refer to the total amount of core enzyme in solution (i.e., without correction for the amount of RNAP active in promoter binding).

Oligodeoxyribonucleotide (oligo) Purification and End Labeling. Synthetic oligos comprising each strand of a forked template (see Fig. 1) are purified by gel electrophoresis on a polyacrylamide gel containing 15% acrylamide, 7 *M* urea, and TBE buffer (100 m*M* Tris, pH 8.3, 100 m*M* boric acid, and 2 m*M* EDTA).[49] DNA bands are visualized by ultraviolet shadowing on a fluorescent TLC silica plate (Kodak) background. DNA is extracted by a crush-and-soak technique,[50] recovered by loading onto a C18 column, and eluted with 40% methanol. The eluate is dried in a speed-vacuum concentrator, and DNA is redissolved in 50 μl of TE buffer (10 m*M* Tris, pH 7.9, and 1 m*M* EDTA). The concentration is determined by measuring absorbance at 260 nm. The extinction coefficient for each oligo is calculated using the biopolymer calculator program from the Schepartz Lab website (http://paris.chem.yale.edu/extinct.html). Oligos (10–30 pmol) are labeled at the 5′ end with ^{33}P in reactions containing T4 polynucleotide kinase (New England Biolabs) and [γ^{33}P]ATP (Dupont NEN).[49] The final concentration of labeled oligo is adjusted to 1 μM by the addition of water.

Annealing Reactions. 5′ end-labeled top (long) strand oligos (100 n*M*) are annealed to unlabeled bottom (short) strand oligos (150 n*M*) in 40 μl buffer containing 50 m*M* NaCl and 25 m*M* Tris (pH 8.0). The mix is incubated in a temperature block at 90° for 5 min, after which the heat is turned off to allow slow cooling to room temperature.

Preparation of Labeled Promoter DNA by Polymerase Chain Reaction (PCR). The test promoter is a P_{RM} variant with an improved −10 region (TAGAAT instead of TAGATT) contained on a fragment also harboring

[49] J. Sambrook, E. F. Fritsch, and T. Maniatis, Molecular Cloning, 2nd Ed., pp. 13–45, *Cold Spring Harbor Laboratory Press, Plainview, NY*, 1989.

[50] J. Sambrook and D. W. Russell, Molecular Cloning, 3rd Ed., pp. 10–14, *Cold Spring Harbor Laboratory Press, Plainview, NY*, 2001.

a P_R promoter inactivated by mutation. For this promoter, the formation of open complexes is on the minute time scale, for an RNAP concentration of 200 n*M*. The labeled promoter is obtained by PCR using plasmid DNA as the template. PCR primers are designed to yield a promoter fragment of about 470 bp. Each 200-μl reaction contains 100 ng (30 fmol) of plasmid DNA and 50 nmol of each primer, at least one of which is 5′ end labeled; amplification is for 30 cycles.

Buffer. All binding reactions are carried out in HEPES buffer: 30 m*M* HEPES, pH 7.5, 100 m*M* NaCl, 1 m*M* DTT, and 100 μg/ml bovine serum albumin.

Detection of RNAP–DNA Complexes by Electrophoretic Mobility Shift Assay (EMSA). Binding reactions are analyzed on 4 or 5% nondenaturing gels (Accugel 29:1, National Diagnostics) in 1 × TAE (40 mM Tris–acetate, 1 mM EDTA). The gels are run at 100 V for 2.5 h, dried, and exposed overnight to a phosphorimager screen (Molecular Dynamics). The percentages of bound and free DNA substrate are determined from band intensities using ImageQuant software (Molecular Dynamics).

KINETICS OF OPEN COMPLEX FORMATION. Open complex formation can be monitored by several methods, each exploiting different properties of the complex: appearance of $KMnO_4$ sensitivity due to strand separation, ability to make RNA, or stability as reflected by resistance to challenge by heparin. We have generally used the latter and have followed the kinetics of open complex formation, a sensitive method for detecting the effects of mutations in σ factor. Separate reactions were set up for each time point. Binding was initiated at zero time by the addition of RNAP (200 n*M* final) to promoter DNA (2 n*M* final, about 10^4 cpm) in 20 μl (final volume) of HEPES buffer prewarmed for 2 min at the desired temperature. Typically, eight time points are taken over the range of 30 s to 45 min. The reactions are challenged for 1 min with heparin (100 μg/ml final) prior to loading onto a running 4% nondenaturing gel. Band intensities are determined as indicated earlier. Binding data (percentage of DNA bound as a function of time) are fit to exponential functions, with the plateau level(s) and the rate constant(s) (k_{obs}) as variable parameters. A comparison of values of k_{obs} (normalized to $k_{obs} = 1$ for RNAP with wt σ^{70}) among the reconstituted mutant holoenzymes is shown in Fig. 3a.

Our fragment bears an additional promoter (the mutated P_R promoter retained some activity) at which open complex formation is much slower. In order to get a good fit, we use a double exponential equation: $y = A1\ [1\text{-}\exp(-k_{1obs}t)] + A2\ [1\text{-}\exp(-k_{2obs}t)]$, where y is the fraction of DNA bound, A1 and A2 are the plateaus for the two phases, and k_{1obs} and k_{2obs} are rate constants. The relevant rate constant is that for the faster of the two processes. With the promoter shown in Fig. 1, in HEPES

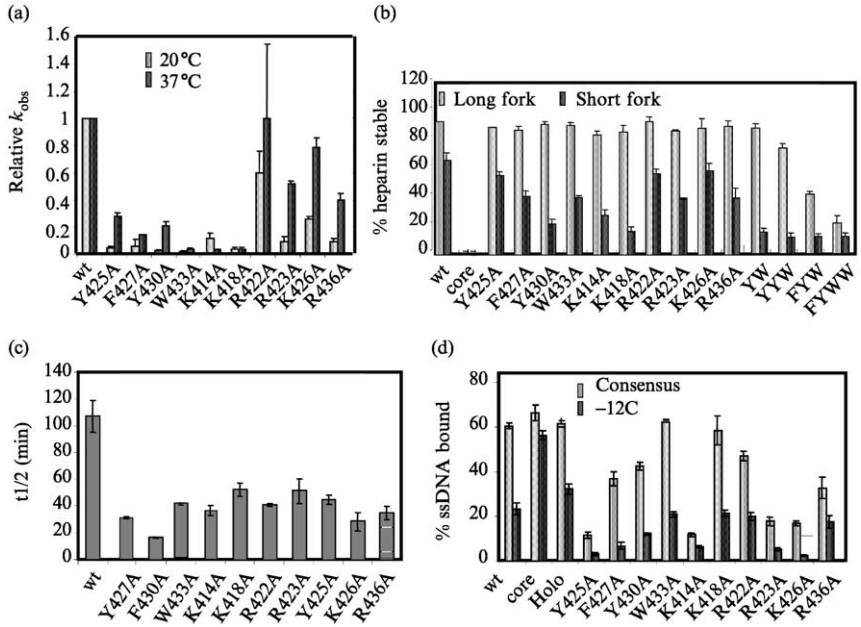

FIG. 2. Sample data on the interaction of RNAP bearing substitutions in σ^{70} with various DNAs. (a) Kinetics of open complex formation. Rate constants for the formation of stable complexes by mutant RNAP have been normalized to that for binding to wt RNAP. (b) Extent of formation of heparin-stable complexes with two different forked DNAs. YW: A substitutions at 430 and 433; YYW: A at 425, 430, and 433; FYW: A at 427, 430, and 433; FYWW: A at 427, 430, 433, and 434. (c) Stability of complexes with short fork DNA. Dissociation rate constants of heparin-stable complexes have been normalized to that for dissociation from wt RNAP. (d) Extent of complex formation with ssDNA. Two single-stranded oligodeoxyribonucleotides were used, differing in the sequence of the −12 base. See text for additional details.

buffer, formation of an open complex with 200 n*M* wt RNAP proceeds with a $k_{1\text{obs}} = 4 \quad 0.4\ \text{min}^{-1}$ at 37° (i.e., the reaction is half over in about 0.2 min).

The kinetics of open complex formation (Fig. 2a)[33,46] show that the alanine substitutions for basic and aromatic amino acids in region 2.3 of σ^{70} render the reconstituted RNAP cold sensitive[44,46,51,52] for open complex formation. This is as expected for residues important for the strand

[51] Y.-L. Juang and J. D. Helmann, *Biochemistry* **34,** 8465 (1995).
[52] S. E. Aiyar, Y.-L. Juang, J. D. Helmann, and P. L. deHaseth, *Biochemistry* **33,** 11501 (1994).

separation process: the effects of the substitutions become more prominent when less thermal energy is available to drive the melting process. Large effects of a single alanine substitution on the overall process of open complex formation are seen for Y425A, F427A, Y430A, W433A, K414A, and K418A. Runoff RNA synthesized after preincubation of promoter and (mutant) RNAP for a set amount of time gave similar results (shown elsewhere[46]). However, the effects of the Y430A and W433A substitutions at 37° were not evident; during the preincubation, the holoenzyme containing all but the most deleterious of σ^{70} substitutions is still able to bind a significant fraction of the promoter DNA.

BINDING OF FORKED DNAs. Labeled forked template (1 n*M*) is preincubated at room temperature for 1 min in HEPES buffer. RNAP is added to 65 n*M*, and the 10-μl reactions are incubated at 25° for 30 min, followed by a 10-min challenge with 100 μg/ml heparin. The mixtures are then subjected to EMSA, run at room temperature. Results are shown for both forked templates in Fig. 2b.[33] Single substitutions do not reduce the fraction of long fork DNA bound to RNAP, as the binding has ''topped out,'' but they can be differentiated very well with short fork DNA. However, use of the long fork allows differentiation among the various holoenzymes reconstituted with multiply substituted σ^{70}, for which the binding to short fork has ''bottomed out.'' Data in Fig. 2b[33] show that the single substitutions most deleterious to the formation of a heparin-resistant RNAP–forked DNA complex are Y430A, K414A, and K418A.

Additional information can be obtained by determining the extent of binding as a function of [RNAP], that is, a binding isotherm (Fig. 3). Individual binding reactions, with a final volume of 10 μl, are prepared, each with 1 n*M* forked DNA and RNA polymerase at concentrations ranging from 0 to 200 n*M*. The mixtures are incubated at room temperature for 30 min. To one set of reactions heparin is added to a final concentration of 100 μg/ml and to the other, the identical volume of ddH_2O. All reactions are incubated for an additional 10 min prior to the addition of 2 μl of a stock solution of loading dye (0.25% bromphenol blue, 0.25% xylene cyanol FF, and 30% glycerol). Fractions of DNA bound in the absence and presence of heparin, $\theta_{-\mathrm{hep}}$, and $\theta_{+\mathrm{hep}}$, respectively, are determined from band intensities, as described earlier.

Reactions, Equations, and Data Analysis. Complexes stable enough to survive a heparin challenge are assumed to contain RNAP in a conformation similar to that in an open promoter complex.[48] The interaction of RNAP with a forked template proceeds along the following scheme:

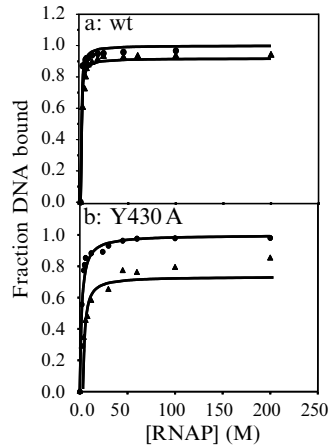

FIG. 3. RNAP–long fork interaction: Binding isotherms for formation of heparin-sensitive and -resistant complexes for wt holoenzyme and holoenzyme bearing the Y430A substitution in σ^{70}. Circles represent binding in the absence of a heparin challenge, and triangles represent binding following such a challenge. Data analysis and curve fitting were as described in the text. Best fits were obtained with $K_1 = 1.9 \times 10^8\ M^{-1}$ and $K_f = 11$ for RNAP containing wt sigma, and $K_1 = 1.8 \times 10^8\ M^{-1}$ and $K_f = 2.5$ for Y430A. See text and Table III.

$$R + F \underset{}{\overset{K_1}{\rightleftharpoons}} RF_1 \overset{K_f}{\rightleftharpoons} RF_{CR}$$

SCHEME 2

where F is the forked DNA and RF_{CR} is the competitor (i.e., heparin) –resistant complex.[48] The equilibrium constants for the two steps are defined as follows:

$$K_1 = \frac{k_1}{k_{-1}} = \frac{[RF_1]}{[R][F]} \tag{1}$$

$$K_f = \frac{k_f}{k_r} = \frac{[RF_{CR}]}{[RF_1]} \tag{2}$$

where K_1 provides information about the affinity between RNAP and the forked DNA and K_f about the relative amounts of heparin-sensitive (RF_1) and heparin-resistant (RF_{CR}) complexes present at equilibrium. The experimental goal is to determine the values of K_1 and K_f. The relevant expression for θ_{+hep} is[53]

[53] O. V. Tsodikov and M. T. Record, Jr., *Biophys. J.* **76,** 1320 (1999).

$$\theta_{+\text{hep}} = \frac{K_1 K_f [R_f]}{(K_1 + K_1 K_f)[R_f]} \tag{3}$$

At the lowest concentrations of RNAP, the concentration of free RNAP, $[R_f]$, cannot be approximated adequately by the total concentration of the enzyme. Therefore, the following explicit expression for $[R_f]$ was used at all RNAP concentrations:

$$[R_f] = \frac{b + \sqrt{(b^2 + 4K_{app}[R])}}{2K_{app}} \tag{4}$$

where

$$b = -1 - (K_{app}[\text{forked DNA}] + K_{app}[R]) \tag{5}$$

For data obtained in the absence of a heparin challenge,

$$\theta_{-\text{hep}} = \frac{K_{app}[R_f]}{1 + K_{app}[R_f]} \tag{6}$$

$$\text{with} \quad K_{app} = K_1 + K_1 K_f \tag{7}$$

Values for K_1 and K_f are obtained by simultaneous fitting of data to the equations for the RNAP concentration dependence of $\theta_{+\text{hep}}$ [Eq. (3–5)] and $\theta_{-\text{hep}}$ [Eq. (4–7)]. To achieve this, the [RNAP] are entered twice in the same column and both sets of binding data (+ and − heparin) are entered in another column of the Sigma Plot spread sheet. In two adjacent columns, 0 or 1 is entered as multipliers to direct the use of one or the other equation. Sample pages containing data and equations are found in Tables I and II, respectively. See Fig. 3 for data and fits.

Data collected in Table III represent examples of the use of forked DNAs for determining the effects of amino acid substitution in σ^{70} on binding (K_1) and the subsequent steps (K_f). The major effects of substitutions for two aromatic amino acids, Y430A and W433A, are seen to be similar: reduction of the values of K_f. However, the K414A substitution primarily causes a reduction in K_1. In this assay, K418 has only a relatively small effect, mostly on K_f. The triple substitution YYW causes a drastic reduction in K_f as seen for both short and long forks. For the short, but not the long fork, an additional effect on K_1 is observed for as yet unknown reasons.

Stability of RNAP–DNA Complexes. In the aforementioned protocols, the added heparin serves to distinguish complexes with a very short half-life from those that have acquired significant stability. Upon addition of the heparin, it is observed that a fraction of the complexes dissociates

TABLE I
SAMPLE DATA ENTRY[a]

	1 [RNAP] (*M*)	2 frx. bound − or + hep	3 − heparin multiplier	4 + heparin multiplier
1	0	0	1	0
2	2E-09	0.556	1	0
3	4E-09	0.774	1	0
4	5E-09	0.808	1	0
5	6E-09	0.852	1	0
6	0	0	0	1
7	2E-09	0.288	0	1
8	4E-09	0.343	0	1
9	5E-09	–	0	1
10	6E-09	0.452	0	1

[a] For a subset of Y430A data shown in Fig. 3b. Lines 1–5, data without heparin challenge; lines 6–10, data with heparin challenge. Typically, each set of data contains 12 or more RNAP concentrations (2E-09 = 2 n*M*, etc.). To allow simultaneous fitting, Eq. 4–7 (for fitting data without heparin challenge) have column 3 as the multiplier, and Eq. 3–6 (for fitting data with heparin challenge) have column 4.

readily, with the remaining fraction dissociating much more slowly.[48] By following the time course of dissociation of "heparin-stable" complexes, any differential effects of the substitutions on the stabilities of these complexes can be determined. Dissociation rate constants, or complex half-lives, are determined from data obtained 5 min or more after the addition of heparin, using nonlinear fits [$y = \exp(-k_{\text{diss}}t)$], or from the slopes of semilogarithmic plots of the fraction bound.

Binding data shown in Fig. 2c[33] indicate that of all single substitutions tested, Y430A causes the greatest decrease in the stability of the complex between a mutated RNAP and forked DNA: from a half-life of 100 min for complexes with wt RNAP[33] to 15 min.

INTERACTION OF RNAP WITH SINGLE-STRANDED DNA. End-labeled oligo (10 n*M*) is incubated with 65 n*M* RNAP for 30 min at 25°, followed by EMSA on a 5% nondenaturing gel run at 4°. Band intensities are determined as indicated earlier. If the gel is run at 25°, no bands other than that of the free DNA are observed.[32] It is likely that our protocol is not optimal, and that oligo binding mainly occurs upon loading the sample onto the gel at 4°.

Data in Fig. 2d[33] show that Y425, K414, R423, and K426 play a role in specific ssDNA binding. Specific recognition of the ssDNA is demonstrated

TABLE II
SAMPLE PAGE OF EQUATIONS USED FOR SIMULTANEOUS FITTING WITH SIGMA PLOT[a]

```
[Variables]
x = col(1)
y = col(2)

[Parameters]
k1 = 2e7
k2 = 1

[Equation]
dnat = 1e-9
k = k1 + k1*k2
b = −1 −k*dnat + k*x
rfree=(b + sqrt(b*b + 4*k*x))/(2*k)
f = col(3)*k*rfree/(1 + k*rfree) + col(4)*k1*k2*rfree/(1 + (k1 + k1*k2)*rfree)
fit f to y

[Constraints]
k1 > 0
k2 > 0

[Options]
tolerance = 0.0000100
stepsize = 100
iterations = 10000
```

[a] Here k1, K_1; k2, K_f; dnat, the total concentration of DNA (1 n*M* in this experiment); k, K_{app}; rfree, the concentration of free RNAP; x, column 1; and y, column 2, as in Table I. Locations of the multipliers in columns 3 and 4 of Table I are entered into the equation for *f*, under the header [Equation]. Values under the header [Parameters] are the initial estimates for k1 and k2.

by the observation that binding is reduced as a consequence of the T-12C substitution.[33]

Interpretation of Data

The following conclusions can be drawn from the aggregate of data: (1) K414 participates directly in double-stranded promoter DNA binding, while the role of K418 is not yet entirely clear. (2) Y430, crucial for interaction with the nucleation-modeling forked DNA, would be involved in the nucleation of DNA melting, possibly by stacking onto the −11 base after its flipping out of the helix. (3) K414, Y425, R423, and K426 interact with single-stranded DNA following nucleation, but perhaps concomitant with DNA strand separation. Apparently, K414 may engage in interactions with both dsDNA and ssDNA. W433 plays an important role, which was

TABLE III
COMPARISON OF BINDING EXPERIMENTS WITH SHORT AND LONG FORKS

	Short fork[a]				Long fork[a]			
RNAP[b]	K_1, 10^8 M^{-1}		K_f		K_1, 10^8 M^{-1}		K_f	
wt σ^{70}	2.8	0.1	1.1		1.9	0.1	11.1	0.1
Y430A	1.77	0.04[c]	0.41	0.05[c]	1.8	0.1[c]	2.5	0.3[c]
W433A	1.94	0.03	0.44	0.04	ND[d]		ND	
YYW[e]	0.34	0.05	0.21	0.01	2.2	0.1	1.7	0.1
K414A	0.27	0.01	0.65	0.05	0.15	0.03	3.8	0.6
K418A	1.41	0.02	0.35	0.02	1.8	0.1	3.1	0.2

[a] See Fig. 1.
[b] with wt σ^{70} or σ^{70} bearing indicated substitution(s).
[c] Data from Tsujikawa *et al.*[48]
[d] Not done.
[e] Triple substitution: Y425A, Y430A, and W433A.

not captured by our model templates. It may interact (stack) with the downstream helix after base flipping. From the σ^{70} and holoenzyme structures it is evident that the F427 side chain is buried in the sigma structure so any effects of the F427A substitution are likely indirect. John Helmann's results with *Bacillus subtilis* Eσ^A RNAP[44,54] and Jay Gralla's with *E.coli* Eσ^{70} [32] are consistent with the aforementioned conclusions, particularly with respect to the roles of aromatic amino acids.

Acknowledgments

Research in our laboratory was supported by NIH Grant GM31808 to PLdH. We thank Drs. David Auble, Phil Rather, and David Samols for helpful comments and Dr. Oleg Tsodikov for his extensive prior help with analysis of the binding data.

[54] J. C. Rong and J. D. Helmann, *J. Bacteriol.* **176,** 5218 (1994).