Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase α subunit

Shawn T. Estrem,[1] Wilma Ross,[1] Tamas Gaal,[1] Z. W. Susan Chen,[1] Wei Niu,[2] Richard H. Ebright,[2] and Richard L. Gourse[1,3]

[1]Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53706 USA; [2]Howard Hughes Medical Institute, Waksman Institute and Department of Chemistry, Rutgers University, Piscataway, New Jersey 08854 USA

We demonstrate here that the previously described bacterial promoter upstream element (UP element) consists of two distinct subsites, each of which, by itself, can bind the RNA polymerase holoenzyme α subunit carboxy-terminal domain (RNAP αCTD) and stimulate transcription. Using binding-site-selection experiments, we identify the consensus sequence for each subsite. The selected proximal subsites (positions –46 to –38; consensus 5′-AAAAAARNR-3′) stimulate transcription up to 170-fold, and the selected distal subsites (positions –57 to –47; consensus 5′-AWWWWWTTTTT-3′) stimulate transcription up to 16-fold. RNAP has subunit composition $\alpha_2\beta\beta'\sigma$ and thus contains two copies of αCTD. Experiments with RNAP derivatives containing only one copy of αCTD indicate, in contrast to a previous report, that the two αCTDs function interchangeably with respect to UP element recognition. Furthermore, function of the consensus proximal subsite requires only one copy of αCTD, whereas function of the consensus distal subsite requires both copies of αCTD. We propose that each subsite constitutes a binding site for a copy of αCTD, and that binding of an αCTD to the proximal subsite region (through specific interactions with a consensus proximal subsite or through nonspecific interactions with a nonconsensus proximal subsite) is a prerequisite for binding of the other αCTD to the distal subsite.

[*Key Words*: Promoter; RNA polymerase; α subunit; UP element; transcription initiation]

Received May 18, 1999; revised version accepted July 6, 1999.

Bacterial promoters consist of at least three RNA polymerase (RNAP) recognition sequences: The –10 element, the –35 element, and the UP element (Hawley and McClure 1983; Ross et al. 1993). The –10 and –35 elements are recognized by the RNAP σ subunit (Dombroski et al. 1992), and the UP element, located upstream of the –35 element, is recognized by the RNAP α subunit (Ross et al. 1993; Blatter et al. 1994). The best-characterized UP element is in the *rrnB* P1 promoter, in which the sequence determinants are located between positions –40 and –60 with respect to the transcription start site (Rao et al. 1994), and UP element-α interactions facilitate initial binding of RNAP and subsequent step(s) in transcription initiation (Rao et al. 1994; Strainic et al. 1998). A consensus UP element sequence (referred to here as the consensus full UP element), derived from binding-site-selection experiments, consists almost exclusively of A and T residues and increases promoter activity >300-fold (Estrem et al. 1998). UP elements have been identified upstream of many bacterial and phage promoters and can function with RNAPs containing different σ factors (e.g., Newlands et al. 1993; Ross et al. 1993, 1998; Fredrick et al. 1995).

Each RNAP α subunit consists of two domains connected by a long unstructured and/or flexible linker (Blatter et al. 1994; Jeon et al. 1997). The 28-kD amino-terminal domain (αNTD) is responsible for dimerization of α and for interaction with the remainder of RNAP (Igarashi and Ishihama 1991; Busby and Ebright 1994). The 8-kD carboxy-terminal domain (αCTD) is responsible for interaction with the UP element (Blatter et al. 1994) and with a number of transcriptional activators (Igarashi and Ishihama 1991; Busby and Ebright 1994; Savery et al. 1998). The αCTD residues most crucial for DNA interaction are nearly invariant in bacteria (Gaal et al. 1996; Murakami et al. 1996), and therefore the DNA sequences recognized by α are also likely to be highly conserved. The interdomain linker presumably accounts for the ability of αCTD to interact with DNA and/or activator molecules at different locations upstream of the –35 element (Newlands et al. 1992; Blatter et al.

[3]**Corresponding author.**
E-MAIL rgourse@bact.wisc.edu; FAX (608) 262-9865.

1994; Murakami et al. 1997b; Belyaeva et al. 1998; Hochschild and Dove 1998; Law et al. 1999). Despite the importance of the αCTD–DNA interaction for bacterial transcription, the details of DNA recognition by α remain to be elucidated.

Several lines of evidence suggest that the consensus full and *rrnB* P1 UP elements each contains two parts, that is, a proximal subsite, centered at about position –42, and a distal subsite, centered at about position –52 (Ross et al. 1993; Estrem et al. 1998). First, the *rrnB* P1 proximal subsite, in the absence of the *rrnB* P1 distal subsite, is protected by RNAP in hydroxyl radical DNA footprinting experiments and exhibits partial ability to stimulate transcription (Leirmo and Gourse 1991; Newlands et al. 1991; Rao et al. 1994). Second, the proximal subsite of the *rrnB* P1 UP element, by itself, is able to cooperate with CAP (catabolite activator protein) in CAP-dependent transcription (Czarniecki et al. 1997; Noel and Reznikoff, 1998; Law et al. 1999). Third, the proximal and distal subsites of the *rrnB* P1 UP element can be separated by insertion of 11 bp without loss of protection of either subsite by RNAP and without loss of the ability to stimulate transcription (Newlands et al. 1992). Fourth, in DNA affinity-cleaving experiments with an RNAP derivative containing Fe-EDTA incorporated into αCTD, two sets of cleavages are observed in the *rrnB* P1 UP element—one in the proximal subsite and one in the distal subsite (Murakami et al. 1997a).

Here we define consensus sequences for individual UP element subsites and determine the number of copies of αCTD required to interact with and respond to full UP elements and individual UP element subsites. The results have important implications for UP element structure/function and for promoter architecture.

Results

Identification of optimal proximal subsite sequences

To confirm that the proximal UP element subsite can function without a distal subsite and to define the optimal sequence for the proximal subsite, we performed binding-site-selection experiments analogous to those used to define the consensus full UP element (Estrem et al. 1998). We constructed a library of DNA fragments containing the *rrnB* P1 core promoter, randomized DNA sequences in the proximal subsite region (–46 to –38), and a sequence shown previously to lack UP element function in the distal subsite region (Fig. 1A). [Position –37 was not randomized, because it was shown previously that cytosine is critical at this position in *rrnB* P1 (Josaitis et al. 1990).]

We incubated RNAP with the DNA fragment library for a time limiting for RNAP-promoter complex formation, blocked further RNAP–promoter complex formation by addition of heparin, isolated RNAP–promoter complexes by nondenaturing PAGE, and amplified promoter DNA from RNAP–promoter complexes by PCR. After 13 cycles of selection and amplification by increasingly stringent conditions (see Materials and Methods), promoters were cloned as phage λ-borne *lacZ* fusions, and transcription activities were assessed by plating on MacConkey-lactose indicator agar. On the basis of plaque color, at least 90% of the selected promoters were more active than the control promoter lacking an UP element, and remarkably, ~30% were even more active than the wild-type *rrnB* P1 promoter.

Nineteen clones with the darkest red plaque color were analyzed by DNA sequencing, and eight different proximal subsite sequences were identified (Figs. 1A–C). Six of the eight sequences contained a perfect A tract from –46 to –41, and the remaining two contained near-perfect A tracts (interrupted only by a T at position –42 or by a C at –46). There also was a bias for purines at positions –38 and –40. We quantified promoter activities by measuring β-galactosidase activities of strains monolysogenic for phages containing the promoter–*lacZ* fusions. The proximal subsites stimulated transcription 82- to 170-fold (Fig. 1B), which is more than the stimulation observed with the full UP element from *rrnB* P1 (69-fold), but less than the stimulation observed with the consensus full UP element (330-fold; Estrem et al. 1998).

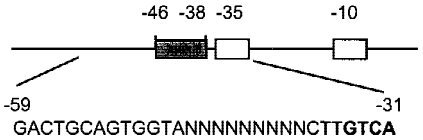
To provide information about the relative importance of individual positions for function, we introduced single transversions into a representative selected proximal subsite (promoter 4547; 130-fold stimulation; Fig. 2). Each substitution decreased transcription: substitutions at –41, –42, or –43 decreased proximal subsite function strongly (to 6- to 10 fold stimulation; 5%–8% the effect of the parent proximal subsite); substitutions at –44 and –45 decreased transcription moderately (to 34- to 37-fold stimulation; 26%–28% of the parent); and substitutions at positions –38, –39, –40, and –46 decreased transcription modestly (to 68-to 96-fold stimulation; 52%–74% of the parent). Taken together, the nucleotide frequencies from the binding-site-selection experiment (Fig. 1C) and the mutational analysis of a consensus proximal subsite (Fig. 2) suggest that positions –41 to –43 are most crucial for proximal subsite function.

Like the single transversion mutants, the selected sequence 4542 also contains a single base pair change from the sequence in proximal subsite 4547 (Fig. 1). In this case, however, the subsite has a T at position –42, yet exhibited full function in stimulating transcription. Furthermore, the *rrnB* P1 proximal subsite contains T at each of the three critical positions –41, –42, and –43, yet still stimulated transcription moderately (20-fold; Fig. 1B). We conclude that UP elements with T substitutions at these positions retain substantial function (see Discussion).

Identification of optimal distal subsite sequences

To determine whether the distal UP element subsite can function without a proximal subsite and to define the optimal sequence for the distal subsite, we constructed a library of DNA fragments containing the *rrnB* P1 core promoter, randomized sequences in the distal region (–59 to –46), and a sequence shown previously to lack UP element function in the proximal region (Fig. 3A). We

A



B

Name and No. of isolates	Relative Activity	Proximal subsite Sequence
4549 (3)	170	aaaaaaaga
4546 (3)	160	aaaaaaaca
4544 (3)	160	aaaaaaatg
4547 (2)	130	aaaaaagta
4548 (5)	130	aaaaaagtg
4542* (1)	130	aaaatagta
4545 (1)	110	caaaaaaca
4543* (1)	82	aaaaaaata
rrnB P1 Proximal	20	aaatttcct
No UP	1	cctaggaat

C

	-46	-44	-42	-40	-38
A	88	100	100	100	88	100	63	0	75
T	0	0	0	0	12	0	0	63	0
G	0	0	0	0	0	0	37	12	25
C	12	0	0	0	0	0	0	25	0

Figure 1. Proximal subsites. Sequences and relative activities of eight promoters selected for binding of RNAP in vitro and screened for high transcription in vivo. (*A*) DNA fragments contained a wild-type *rrnB* P1 core promoter sequence (open boxes indicate the –10 (TATAAT) and –35 (TTGTCA) elements), a distal region (–59 to –47) that does not influence transcription (SUB sequence; Rao et al. 1994), and different proximal region sequences (–46 to –38, shaded rectangle) derived from a random sequence library. The randomized residues are indicated as N and shown in context (nontemplate strand) below the schematic. The –35 hexamer is in boldface type. (*B*) Proximal subsite sequences are shown for the eight promoters, for an *rrnB* P1 construct containing only the proximal subsite of the *rrnB* P1 UP element (*rrnB* P1 proximal; RLG3098), and for a construct lacking an UP element (No UP; RLG3097). The promoter with the *rrnB* P1 proximal UP element contains the SUB sequence from –59 to –47. The No UP promoter is *rrnB* P1 with SUB sequence from –59 to –39 (Estrem et al. 1998). The names refer to the strain numbers of λ lysogens containing the indicated UP element sequence. (*) Promoters with single base pair mutations downstream of the transcription start site (most likely introduced during the last round of PCR following isolation of gel-shifted fragments). Sequence variation in this region does not affect *rrnB* P1 promoter activity (Bartlett and Gourse 1994). The number of isolates obtained for each sequence is indicated in parentheses. Promoter activities are expressed relative to the activity of the No UP promoter (activity = 1) and were determined from β-galactosidase measurements of λ lysogens containing promoter–*lacZ* fusions. Promoter activities differed by <4% in at least two experiments. The approximately twofold difference between the effect of the *rrnB* P1 proximal subsite reported here and the effect reported previously (20- vs. 8-fold; Rao et al. 1994) most likely derives from minor differences in the flanking sequences upstream and at positions –38 to –40 in the constructs used in the two studies. (*C*) Nucleotide frequencies (percentage of eight sequences shown in B) for each proximal subsite position (–46 to –38).

then performed binding-site-selection experiments and in vivo assays analogous to those used for the proximal subsite selection described above. On the basis of plaque color, ~50% of the resulting selected promoters exhibited activities greater than that of the control promoter lacking an UP element, but none of these promoters were as active as *rrnB* P1. From 21 clones producing the darkest red plaques, 19 different distal subsite sequences were identified (Fig. 3B,C). The sequences had a high frequency of A residues at –57, A or T from –56 to –53, and T from –52 to –47 (Fig. 3C), and stimulated transcription 4- to 16-fold (Fig. 3B). This level of transcription stimulation is less than that observed with the consensus full UP element, the consensus proximal subsite, or even the *rrnB* P1 proximal subsite.

The *rrnB* P1 distal subsite closely matches the bind-

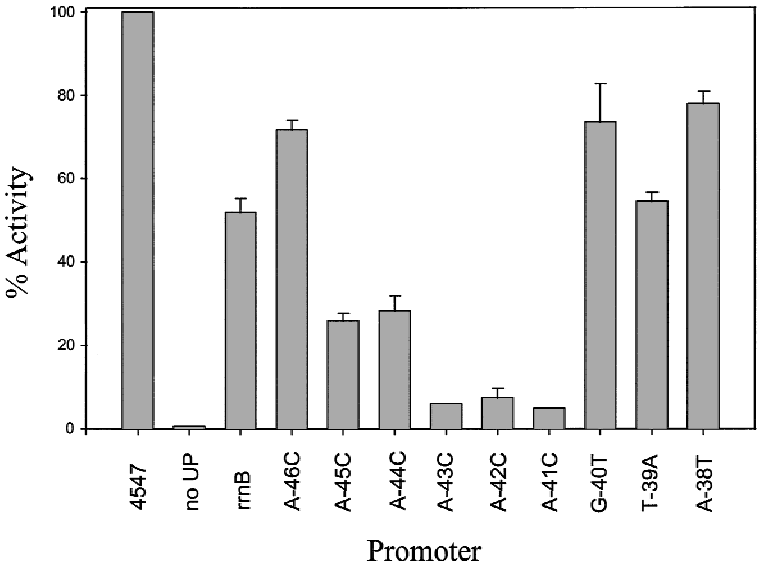
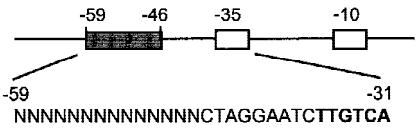


Figure 2. Effects of substitution mutations on proximal subsite function. Transcription activities were determined from measurements of β-galactosidase activities from monolysogens containing promoter–*lacZ* fusions and are expressed relative to the activity of the parent, the *rrnB* P1 promoter containing proximal subsite 4547 (see Fig. 1). The mutant strains are RLG4523 (A-46C), RLG4522 (A-45C), RLG4524 (A-44C), RLG4525 (A-43C), RLG4526 (A-42C), RLG4527 (A-41C), RLG4528 (G-40T), RLG4529 (T-39A), and RLG4530 (A-38T). For comparison, activities of *rrnB* P1 promoters containing the intact *rrnB* P1 UP element (rrnB; RLG3074) or lacking an UP element (no UP; RLG3097) are also shown.

A



B

Name and No. of isolates		Relative Activity	Distal subsite Sequence
4518	(2)	16	agaaaaatattttg
4517	(1)	15	gcatttttttttca
4512*	(1)	14	gtaaaaatttttta
4520	(1)	13	acgtatttttttta
4521	(1)	12	gaaaaatattttg
4511	(2)	12	taaaaaatatttta
4515	(1)	12	taagttttttttta
4503	(1)	11	aaaaatttattttg
4507*	(1)	10	cgaaaaaaaattta
4509	(1)	10	ccggtttttttttta
4500	(1)	9	taaatttttttttt
4516	(1)	9	gaaaaaaatagttg
4514	(1)	7	atatgttttttta
4505	(1)	6	agatttatttttct
4504	(1)	5	gataaaaatagttg
4510	(1)	5	gtatgatttttta
4508	(1)	5	ataaaatattttat
4519	(1)	4	gcaaatatattttt
4501	(1)	4	tgtaataatttta
rrnB P1 Distal		9	agaaaattatttta
No UP		1	gactgcagtggtac
4513		16	ggaaaatttttttt

C

	-58	-56	-54	-52	-50	-48	-46
A	31	37	78	63	63	47	47	31	26	16	0	0	5	53
T	21	21	11	26	26	53	53	69	74	84	89	100	84	21
G	37	21	11	11	11	0	0	0	0	0	11	0	0	26
C	11	21	0	0	0	0	0	0	0	0	0	0	11	0

Figure 3. Distal subsites. Sequences and relative activities of 19 promoters selected for binding of RNAP in vitro and screened for high transcription in vivo. Details are described in Fig. 1, except that the promoters contained different distal regions (filled rectangle; –59 to –46) and the SUB sequence in the proximal region (–45 to –38 CTAGGAAT). The randomized residues are indicated as N and displayed in context (nontemplate strand) below the schematic. The –35 hexamer is in boldface type. Distal region sequences are shown for the 19 promoters, for an *rrnB* P1 construct containing only the distal region of the *rrnB* P1 UP element (*rrnB* P1 distal; RLG3099), and for a construct lacking an UP element (No UP; RLG3097). Distal subsite 4513 is described in the text.

ing-site-selected distal subsites. We had previously constructed overlapping triple substitutions in the *rrnB* P1 distal subsite and measured their effects as promoter–*lacZ* fusions to obtain information about individual residues important for function (Estrem 1998). All triple substitutions in the distal subsite decreased transcription at least threefold, and the substitution centered at position –52 decreased transcription the most (approximately sixfold).

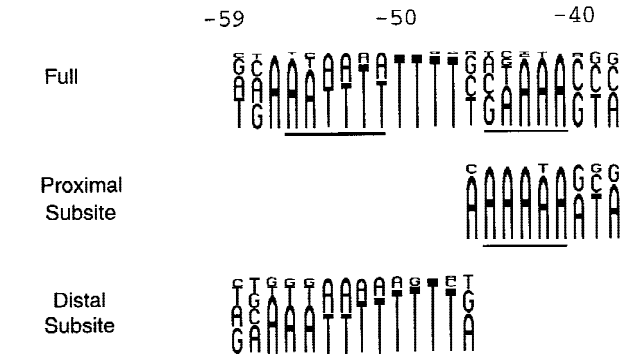
Relationship between the consensus full UP element and the consensus subsite sequences

The distributions of nucleotides at each position in the selected proximal and distal subsite sequences are pictured in diagram form in Figure 4A and compared with the distribution obtained in the previously - described full UP element selection (Estrem et al. 1998). Fig. 4B presents the derived consensus sequences.

The consensus proximal subsite sequence is related to the corresponding proximal region in the consensus full UP element, but differs in substantive ways. The consensus proximal subsite includes the three specified positions from the corresponding segment of the consensus full UP element, –41, –42, and –43, but it also contains five additional specified positions, with strong preference for A at –44, –45, and –46 and for purine at –38 and –40 (Fig. 4B).

In contrast, the consensus distal subsite sequence is almost identical to that of the corresponding sequence

A



B

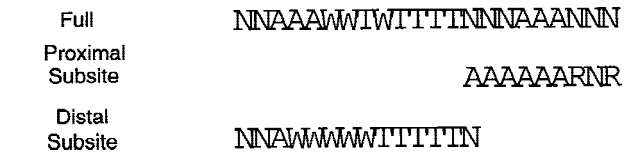


Figure 4. Consensus sequences. (*A*) Frequency diagrams of residues in the binding-selected full UP elements (from Estrem et al. 1998) and in the binding-selected proximal and distal subsites (from Figs. 1C and 3C). Each nucleotide is represented as a letter proportional in size to its frequency at that position in the selected population. The nontemplate strand positions protected by RNAP in hydroxyl radical footprints (Estrem et al. 1998; Fig. 6) are indicated by lines. (*B*) Consensus subsite sequences based on the nucleotide frequencies. One nucleotide is indicated when it is present in >70% of the population, or two when together they represent 95% or more of the population. W = A or T; R = A or G; N = no single base pair present in 70% of the population and no 2 bp make up 95% of the population.

from the *rrnB* P1 and consensus full UP elements (Estrem et al. 1998). We constructed promoters containing the distal subsite from *rrnB* P1 or full UP element 4192 (Estrem et al. 1998) and containing a nonfunctional proximal region. The resulting UP elements stimulated transcription in vivo 9- and 16-fold, respectively, consistent with their sequence similarity to the binding-selected distal subsites (*rrnB* P1 Distal and 4513; Fig. 3B).

The most active proximal and distal subsite sequences (4549 and 4513; Figs. 1 and 3), were combined to create a composite UP element (4541; –59 5′-GGAAAATTTT-TTTAAAAAAAGA-3′–38). The stimulatory effect of the resulting composite UP element was 340-fold (data not shown), which is very similar to the effect of the consensus full UP element (330-fold; Estrem et al. 1998). Nevertheless, the stimulatory effect is far below that expected for the product of the effects of the two individual subsites (16-fold × 170-fold = 2720-fold), suggesting that the observed 330- to 340-fold increase represents the limit for activation of the *rrnB* P1 core promoter in vivo and/or that in consensus full UP elements the two subsites do not function independently (see Discussion).

The consensus proximal and distal subsites stimulate transcription through interactions with αCTD

In vitro transcription experiments were performed to establish that individual consensus proximal and distal subsites, by themselves, stimulate transcription through interactions with αCTD. The consensus proximal and distal subsites increased transcription 10- and 9-fold, respectively (Fig. 5). (Under the same conditions, a consensus full UP element and the *rrnB* P1 UP element stimulated transcription by 47- and 21-fold, respectively.) The single base pair substitutions in the proximal subsite that decreased transcription in vivo (Fig. 2) also decreased transcription in vitro (data not shown). We conclude that the individual consensus subsites stimulate transcription and that this stimulation requires no components other than promoter DNA and RNAP.

We note that the consensus proximal subsite stimulated transcription less well in vitro than in vivo (10-fold vs. 130-fold; Figs. 5 and 1), whereas the consensus distal subsite stimulated transcription similarly in vitro and in vivo (~9-fold vs. 16-fold, respectively; Figs. 5 and 3). The quantitative difference in vitro versus in vivo for the effect of the proximal subsite may reflect differences in limiting steps to which the assays are sensitive, differences in solution conditions, differences in supercoiling, or the absence/presence of potential accessory factors.

To assess the dependence of transcription stimulation on αCTD–DNA interaction, we performed parallel in vitro transcription experiments with two mutant RNAP derivatives: αΔ235 RNAP, which completely lacks the αCTD; and αR265A RNAP, which has a single amino acid substitution that disrupts αCTD–DNA interaction (Gaal et al. 1996; Murakami et al. 1996). The individual consensus proximal and distal subsites, like the *rrnB* P1 and consensus full UP element, failed to stimulate transcription with αΔ235 RNAP and αR265A RNAP (Fig. 5). We conclude that transcription stimulation by individual consensus subsites absolutely requires αCTD–DNA interaction.

The consensus proximal and distal subsites are binding sites for αCTD

We performed hydroxyl radical DNA footprinting experiments using RNAP and promoters containing only a consensus proximal subsite or only a consensus distal subsite (Fig. 6A–D). In each case, strong protection (i.e., protection comparable to that in the –35 element region) was observed in the consensus subsite, and only weak protection was observed in the nonconsensus subsite.

We also performed hydroxyl radical DNA footprinting experiments using purified α and promoters containing only a consensus proximal subsite or only a consensus distal subsite (Fig. 6A,B,E,F). In each case, preferential protection was observed in the consensus subsite. (Weaker protection was observed also in the nonconsensus subsite and ~10 bp downstream from the consensus proximal subsite. α is a dimer, therefore, the weak protection may be attributable to nonspecific interactions with the second αCTD.) We conclude that a single consensus subsite is sufficient for binding αCTD, both with RNAP and with purified α.

Transcription stimulation by the consensus proximal subsite, but not by the consensus distal subsite, requires only one copy of αCTD

RNAP contains two α subunits: α^I and α^{II} (where α^I is

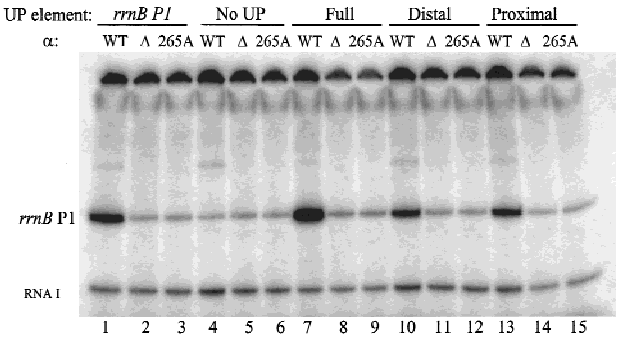


Figure 5. In vitro transcription. Plasmid templates contained *rrnB* P1 core promoters with either the *rrnB* P1 UP element (*rrnB* P1; lanes *1–3*; pRLG4238), no UP element (No UP; lanes *4–6*; pRLG4210), a consensus full UP element [4192 (Estrem et al. 1998); lanes *7–9*; pRLG3278], a consensus distal subsite [4513; lanes *10–12*; pRLG4214], or a consensus proximal subsite [4547; lanes *13–15*; pRLG4213]. Plasmids were transcribed with reconstituted RNAPs at concentrations that resulted in equivalent transcription from the *lacUV5* promoter [2.7 nM for RNAP with wild-type α (WT); 17 nM for RNAP with αΔ235 (Δ); or 9 nM with αR265A (265A)]. The 220-nucleotide *rrnB* P1 transcript terminates at an *rrnB* T1 terminator in the vector. The vector-encoded RNA I transcript (~110 nucleotides) is also indicated. The transcription buffer was as described (Ross et al. 1993), except the reactions contained 170 mM NaCl and no KCl.

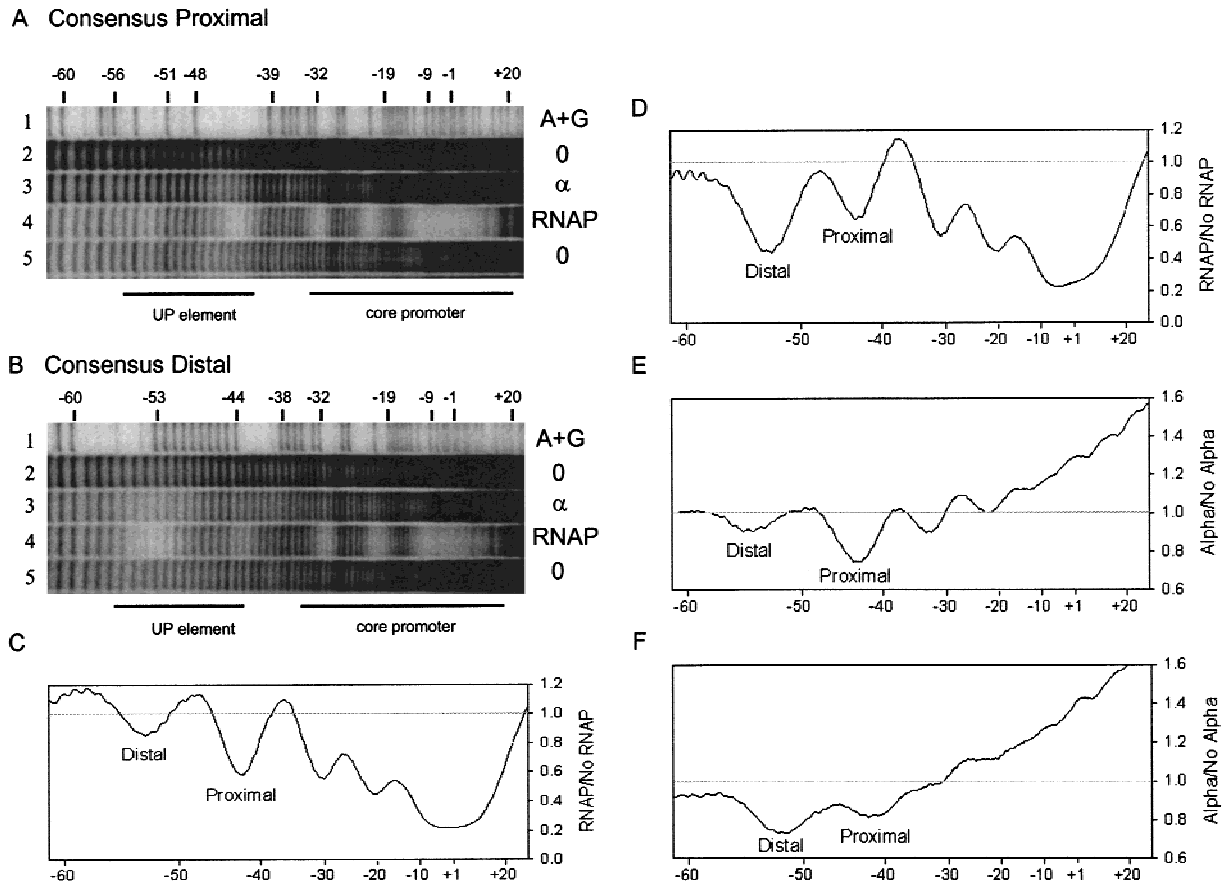


Figure 6. Hydroxyl radical footprints of RNAP and purified α on promoters with consensus subsites. (*A*) *rrnB* P1 core promoter with a consensus proximal subsite (4547). The top of the gel is at *right*. DNA fragments were labeled in the template strand at position –66. (Lane *1*) A+G sequence markers; (lanes *2,5*) no protein (different amounts of sample); (lane *3*)6μM purified α; (lane *4*) 16 nM wild-type RNAP. Lines indicate the positions of the core promoter and –40 to –60 region (UP element). (*B*) *rrnB* P1 core promoter with a consensus distal subsite (4513). Lanes are the same as in *A*. Scans of lanes *3* and *4* from each gel are shown in *C–F* and represent a ratio of radioactivity at each position in the promoter with protein/without protein (see Materials and Methods for details). (*C*) Promoter containing consensus proximal subsite with RNAP holoenzyme. (*D*) Promoter containing consensus distal subsite with RNAP holoenzyme. (*E*) Promoter-containing consensus proximal subsite with purified α. (*F*) Promoter containing consensus distal subsite with purified α.

defined as the subunit that interacts with the β subunit; see Heyduk et al. 1996). To determine whether transcription stimulation by UP element subsites requires αCTD of α^I, αCTD of α^{II}, or both αCTDs, we prepared and analyzed two oriented-α RNAP derivatives: $\alpha^I/\alpha\Delta^{II}$, in which only α^I contains αCTD; and $\alpha\Delta^I/\alpha^{II}$, in which only α^{II} contains αCTD. To prepare oriented-α RNAP, we took advantage of the R45A substitution in α, which results in an α that is unable to interact with β, and thus is unable to serve as α^I (Kimura and Ishihama 1995; Murakami et al. 1997a). We coexpressed genes encoding one α derivative with the R45A substitution and a hexahistidine affinity tag and a second α derivative without the R45A substitution and hexahistidine tag, lysed the cells, and isolated RNAP using metal-ion-affinity chromatography (see Materials and Methods; W. Niu and R.H. Ebright, in prep.).

We performed in vitro transcription experiments with the oriented-α RNAP derivatives and a promoter containing a consensus full UP element. Both $\alpha^I/\alpha\Delta^{II}$ and $\alpha\Delta^I/\alpha^{II}$ transcribed the promoter about one-third as well as wild-type RNAP (Fig. 7, left). The reduction in promoter activity in vitro on elimination of one αCTD was almost as much as the reduction in activity on elimination of one consensus subsite of the consensus full UP element (Fig. 5). We performed parallel experiments with the *rrnB* P1 UP element, which contains a moderately effective proximal subsite but a fully effective distal subsite (Figs. 1 and 3). The oriented-α RNAP derivatives transcribed the *rrnB* P1 promoter only about one-fourth as well as the wild-type RNAP (data not shown). We conclude that both αCTD^I and αCTD^{II} are required for maximal transcription of promoters containing two consensus or near-consensus UP element subsites.

Next, we performed in vitro transcription experiments with the oriented-α RNAP derivatives on promoters containing only a single consensus proximal subsite or a single consensus distal subsite. Both $\alpha^I/\alpha\Delta^{II}$ and $\alpha\Delta^I/\alpha^{II}$

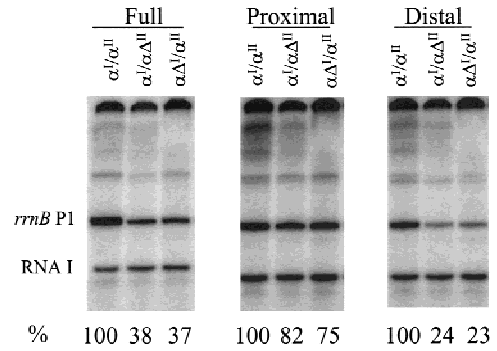


Figure 7. In vitro transcription with oriented-α RNAP derivatives. Plasmids containing promoters with the indicated UP elements were transcribed with reconstituted wild-type RNAP (α^I/α^{II}) or with oriented-α RNAP derivatives ($\alpha^I/\alpha\Delta^{II}$ and $\alpha\Delta^I/\alpha^{II}$). The transcription buffer was as described previously (Ross et al. 1993) except the reactions contained 160 mM NaCl instead of KCl. The templates were supercoiled plasmids containing the *rrnB* P1 core promoter with (*left*), consensus full UP element (4192; pRLG3278); (*middle*) consensus proximal subsite (4547; pRLG4213); or (*right*), consensus distal subsite (4513; pRLG4214). The *rrnB* P1 transcript and the vector-derived RNA I transcript are indicated. Transcriptional activities were quantified by PhosophorImager analysis and are expressed under each lane as a percentage (%) of transcription with the wild-type RNAP on the same template. (See Fig. 5 for relative activities of the three promoters with wild-type RNAP.)

transcribed the promoter with only a consensus proximal subsite nearly as well as wild-type RNAP (~80% as well as wild-type RNAP; Fig. 7, middle). In contrast, the oriented-α RNAP derivatives transcribed the promoter with only a distal subsite much less well than did wild-type RNAP (<25% as well as wild-type RNAP; Fig. 7, right). We conclude that only a single αCTD is required for efficient transcription stimulation by a consensus proximal subsite and that αCTD^I and αCTD^{II} can function interchangeably for this purpose. We also conclude that, in contrast, αCTD^I and αCTD^{II} are both required for efficient transcription stimulation by a consensus distal subsite (see Discussion).

Occupancy of the consensus proximal subsite, but not the consensus distal subsite, requires only one copy of αCTD

To analyze interactions between oriented-α RNAP derivatives and UP element subsites directly, we performed hydroxyl radical DNA footprinting experiments (Fig. 8). Both oriented-α RNAP derivatives, $\alpha^I/\alpha\Delta^{II}$ and $\alpha\Delta^I/\alpha^{II}$, protected the proximal subsite regions of the *rrnB* P1 full UP element and consensus full UP element to the same extent as wild-type RNAP, but protected the distal subsites in the two promoters much less well than did wild-type RNAP (Fig. 8A,B and corresponding PhosphorImager scans Fig. 8D,E). Strikingly, preferential protection of the proximal subsite region was observed even with a promoter having a nonconsensus proximal subsite and a consensus distal subsite (Fig. 8C,F).

We conclude that only a single αCTD is required for interaction with the proximal subsite, and that both αCTD^I and αCTD^{II} can function interchangeably for this purpose. We conclude that, in contrast, both αCTD^I and αCTD^{II} are required for efficient interaction of RNAP with the consensus distal subsite. These conclusions are consistent with the conclusions of the previous section that only a single αCTD is required for transcription stimulation by the consensus proximal subsite, but that both αCTDs are required for transcription stimulation by the consensus distal subsite.

Discussion

UP elements consist of subsites, each of which constitutes a binding site for αCTD

We demonstrate here that UP elements consist of proximal and distal subsites, and we define the consensus sequences for these subsites. The sequences of the consensus proximal and distal subsites are both A+T -rich but are significantly different (–46 5′-AAAAAARNR-3′ –38 vs. –57 5′-AWWWWWTTTTT-3′ –47). The relative tolerance for either A or T at some positions in both the proximal and distal subsites (see Results; Figs. 1 and 3) most likely reflects the binding of αCTD to DNA primarily in the minor groove (W. Ross and R.L. Gourse, unpubl.), where there is usually little discrimination between A and T residues (Seeman et al. 1976; see also Kielkopf et al. 1998). Because each subsite binds an identical peptide (αCTD), the differences in the subsite consensus sequences must reflect the different locations of the two subsites within the RNAP–promoter complex, and thus the different potential molecular interactions for αCTD bound at the two locations. Factors that might differentially influence sequence preferences in the proximal subsite include requirements for possible interactions between αCTD and αNTD or between αCTD and σ region 4 bound at the –35 element (see below).

The sequence of the consensus proximal subsite differs not only from that of the consensus distal subsite, but also from the sequence of the corresponding segment of the consensus full UP element. The fact that the corresponding sequences within the consensus proximal subsite and the consensus full UP element differ indicates that binding of an αCTD at the proximal subsite is altered by binding of the other αCTD at the distal subsite (see also W. Ross and R.L. Gourse, unpubl.). Factors that might differentially constrain the proximal subsite sequence in the context of a full UP element include sequence requirements for potential αCTD–αCTD interactions and/or for DNA bending in or adjacent to the proximal subsite (see below).

Both consensus subsites include A or T tracts that are likely to deviate somewhat in structure from standard B-form DNA (Koo et al. 1986; Young et al. 1995), and we suggest that some aspect of A-tract structure may contribute to α recognition. The stimulatory effect of A tracts on transcription when fused upstream of core promoters often has been attributed to effects of DNA struc-

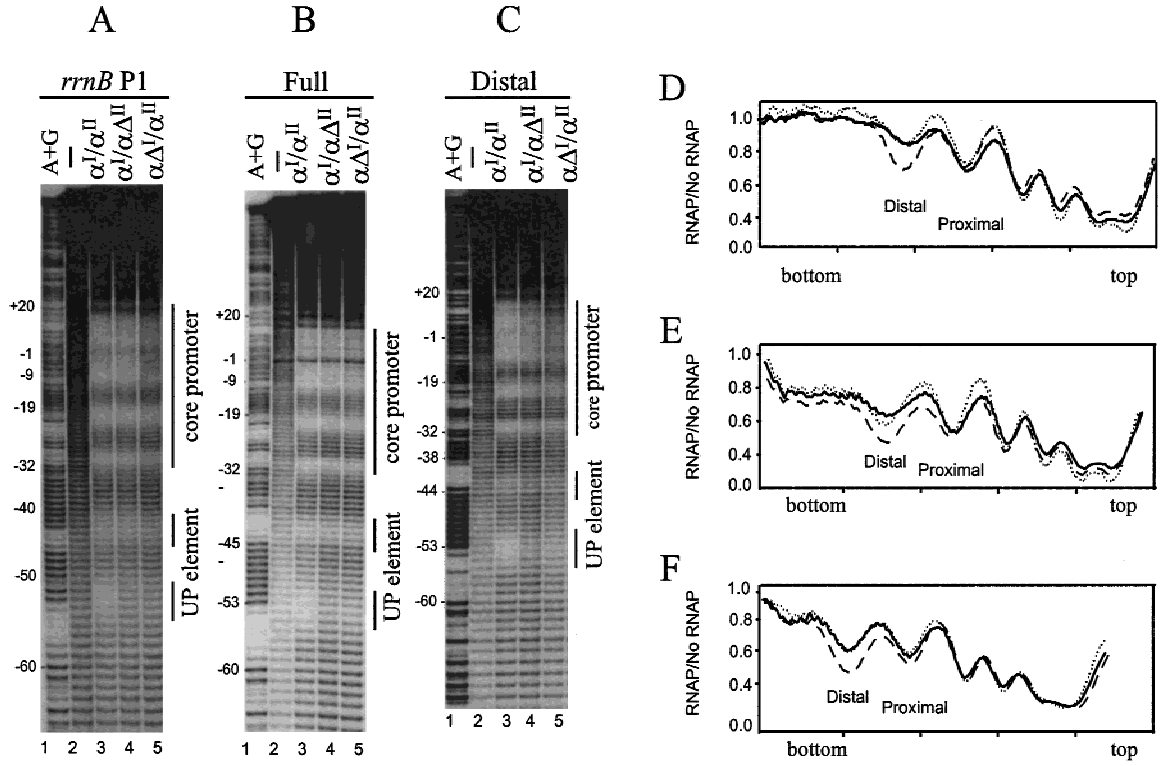


Figure 8. Hydroxyl radical footprints with oriented-α RNAP derivatives. DNA fragments containing the *rrnB* P1 core promoter and different UP elements were labeled in the template strand at position –66. (*A*) *rrnB* P1 UP element. (*B*) consensus full UP element (4192). (*C*) consensus distal subsite (4513). Vertical lines to the *right* of each panel indicate the positions of the core promoter and UP element subsites. (Lane *1*) A+G sequence markers; (lane *2*) no RNAP; (lane *3*) wild-type RNAP (α^I/α^{II}, 22 nM); (lane *4*) oriented-α RNAP ($\alpha^I/\alpha\Delta^{II}$, 8 nM); (lane *5*) oriented-α RNAP ($\alpha\Delta^I/\alpha^{II}$, 4 nM). PhosphorImager scans of the footprints with the three RNAPs are superimposed in *D* (*rrnB* P1; lanes from *A*), *E* (consensus full; lanes from *B*), and *F* (consensus distal; lanes from *C*). Each line is the ratio of radioactivity with RNAP/without RNAP. (Dashed line) Wild-type RNAP; (solid line) oriented-α RNAP ($\alpha^I/\alpha\Delta^{II}$); (dotted line) oriented-α RNAP ($\alpha\Delta^I/\alpha^{II}$). The scans of the footprints with the wild-type and two oriented-α RNAP derivatives are superimposed (normalized) in the core promoter region. The top and bottom of the gel are indicated in *D–F*.

ture (bending) per se. However, we recently demonstrated that A-tract–αCTD interactions account for the observed stimulation (Aiyar et al. 1998).

Our results establish that transcription stimulation by, and protection of, the consensus proximal subsite requires only a single αCTD (Figs. 7–9). We infer that the consensus proximal subsite constitutes a binding site for a single copy of αCTD. Our results further establish that two copies of αCTD are required for maximal transcription stimulation by, and protection of, a consensus full UP element (Figs. 7–9). We infer that the consensus distal subsite also constitutes a binding site for a single copy of αCTD. We note that the observation that function of a consensus proximal subsite requires only one copy of αCTD rules out the possibility that αCTD dimerization (Blatter et al. 1994; Jeon et al. 1997) is required for sequence-specific αCTD–DNA interaction.

The proximal subsite is preferentially occupied by αCTD

Several observations suggest that the proximal subsite region, by virtue of its location within the RNAP–promoter complex, is the preferred binding site for αCTD. First, the consensus proximal subsite is more effective than the consensus distal subsite in transcription stimulation in vivo. Second, the consensus proximal subsite, but not the consensus distal subsite, can stimulate transcription with RNAP derivatives containing only a single copy of αCTD. Third, αCTD preferentially occupies the proximal subsite region in RNAP–promoter complexes containing only one copy of αCTD, even in a promoter with a nonconsensus proximal subsite and a consensus distal subsite (Figs. 7–9).

We suggest four (not mutually exclusive) possible explanations for preferential occupancy of the proximal subsite region by αCTD. All four derive from the fact that the proximal subsite is located closer to the core promoter than the distal subsite (rather than from a difference in intrinsic affinity of the two subsite DNA sequences for αCTD). First, binding of αCTD to the proximal subsite may place less constraint on the linker connecting αCTD to the remainder of RNAP. Second, binding of αCTD to the proximal subsite may demand less DNA bending to bring the subsite close to the core promoter. Third, binding of αCTD to the proximal sub-

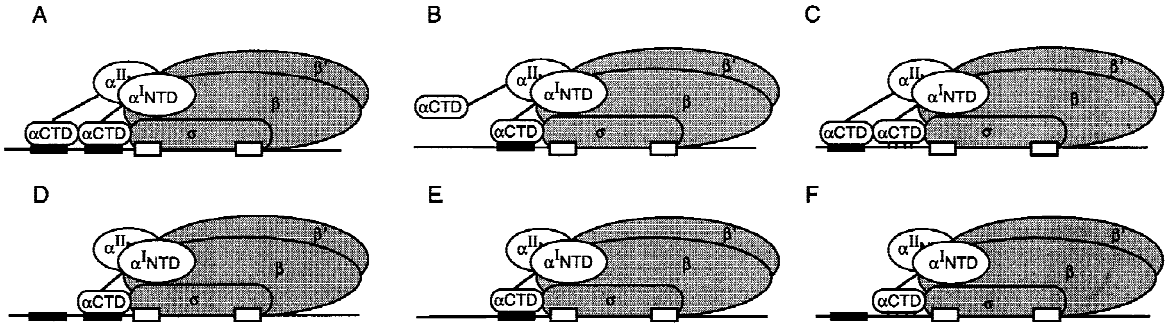


Figure 9. Summary of results and models for αCTD interactions with UP element subsites. The results are shown with wild-type RNAP in *A–C* and with an oriented-α RNAP derivative in *D–F*. Both oriented-α RNAP derivatives (i.e., $\alpha^{I}/\alpha\Delta^{II}$ and $\alpha\Delta^{I}/\alpha^{II}$) interacted similarly with the three types of UP elements [consensus full (*A*,*D*); consensus proximal (*B*,*E*) consensus distal (*C*,*F*)]; for simplicity, only one orientation is shown. Wild-type RNAP is pictured as binding to both subsites at a promoter containing a consensus full UP element (*A*), to the proximal subsite at a promoter containing only a consensus proximal subsite (*B*), or to the distal subsite at a promoter containing only a consensus distal subsite (*C*). In the latter case, nonspecific interactions occur with the nonconsensus proximal region (represented by vertical lines). Oriented-α RNAP derivatives interacted with only the proximal subsite on a promoter containing a consensus full UP element (*D*), only a consensus proximal subsite (*E*), or only a consensus distal sequence (*F*). Protection of the nonconsensus proximal region (indicated by vertical lines) is present in the latter case, but this has no functional consequence.

site may position αCTD to make favorable protein–protein interactions with αNTD. Fourth, binding of αCTD to the proximal subsite may position αCTD to make favorable protein-protein interactions with σ, specifically with σ region 4 bound at the –35 element.

The proposal that αCTD in the proximal subsite interacts with σ region 4, analogously to transcriptional activators that bind in the –40 region and interact with σ region 4 (Li et al. 1994; Lonetto et al. 1998), is especially attractive. We have recently identified mutants of αCTD outside of the DNA-binding determinant and mutants of σ region 4 that result in specific defects in transcription stimulation by the consensus proximal subsite (W. Ross, A. Mertens, D. Schneider, and R.L. Gourse; H. Chen, A. Kapanidis, H. Tang, and R.H. Ebright, both unpubl.). In addition, the proposed αCTD–σ interaction could provide an explanation for the observation that UP elements can affect not only the initial binding of RNAP to promoter DNA to form the closed complex, but also the isomerization of the closed complex to the open complex (Rao et al. 1994; Strainic et al. 1998), because σ is involved in both of these processes (Hochschild and Dove 1998; Helmann and deHaseth 1999).

αCTD at the proximal subsite assists binding of αCTD to the distal subsite

Our results establish that two copies of αCTD are required for function of a consensus distal subsite (Figs. 7 and 8). We propose that binding of a first copy of αCTD in the proximal subsite region cooperatively assists a second copy of αCTD in binding to a consensus distal subsite (Fig. 9A,C). This proposed cooperativity does not require a sequence-specific interaction of the first copy of αCTD with proximal subsite DNA; thus, the phenomenon is observed even with a promoter having a nonconsensus proximal subsite (Fig. 6). We suggest two (non-mutually exclusive) models to explain the proposed cooperativity. First, αCTD in the proximal subsite region may make favorable protein–protein interactions with αCTD at the distal subsite. Second, the presence of both copies of αCTD may result in the formation of a DNA bend in, or adjacent to, the proximal subsite, facilitating binding of αCTD to the distal subsite. We note that position –44 in the consensus full UP element–RNAP complex (Estrem et al. 1998) and positions –38 and –39 in the *rrnB* P1–RNAP complex (Gourse 1988; Ross et al. 1993) are hypersensitive to DNase I cleavage, consistent with DNA bending within or at the downstream boundary of the proximal subsite.

In complexes containing wild-type RNAP and a promoter with either a consensus distal subsite or a full UP element, the proximal subsite is less completely protected from hydroxyl radical attack than the distal subsite. In contrast, the proximal subsite is well protected in a promoter complex with only a consensus proximal subsite (Fig. 6; Newlands et al. 1991; Ross et al. 1993; Estrem et al. 1998). Although we do not fully understand this phenomenon, we suggest that the incomplete protection of the proximal subsite does not reflect poor occupancy of this region of the complex by αCTD, but rather reflects differences in the details of the αCTD–DNA interaction when αCTD is specifically versus nonspecifically bound to DNA, that is, DNA binding of the distally located αCTD alters the sequence-specific proximal subsite interaction (W. Ross and R.L. Gourse, unpubl.).

αCTD^{I} and αCTD^{II} can function interchangeably

Our results with oriented-α RNAP derivatives indicate that αCTD^{I} and αCTD^{II} are interchangeable for UP element subsite recognition. Furthermore, αCTD^{I} and αCTD^{II} are also interchangeable for CAP-dependent

transcription of the *lac* promoter (W. Niu and R.H. Ebright, unpubl.). These results support and extend previous indications (Newlands et al. 1992; Zhou et al. 1994; Murakami et al. 1997b; Aiyar et al. 1998; Belyaeva et al. 1998) that there is a remarkable degree of flexibility in the positioning of αCTD^{I} and αCTD^{II} with respect to the rest of the RNAP–promoter complex, a phenomenon that likely results from the long unstructured linker between the two domains of α (Blatter et al. 1994; Jeon et al. 1997).

Our findings contradict the proposal of Murakami et al. (1997a) that there is a fixed relationship of αCTD^{I} and αCTD^{II} relative to the proximal and distal subsites. These investigators based their proposal on the results of DNA affinity cleaving experiments with an RNAP-derivative containing acetimido-benzyl-EDTA:Fe incorporated at residue 269 of αCTD^{II}. Because cysteine 269 is within the DNA-binding helix of αCTD (Gaal et al. 1996), and because even conservative amino acid substitutions (e.g., C269A, C269S) severely reduce αCTD–DNA binding and UP element-dependent transcription (Gaal et al. 1996; T. Gaal, H. Tang, R.H. Ebright, and R.L. Gourse, unpubl.), we suspect that incorporation of the DNA cleaving agent interferes with sequence-specific DNA interaction by αCTD^{II}. Therefore, we suggest that Murakami et al. (1997a) inadvertently created the functional equivalent of the oriented-α RNAP $\alpha^{I}/\alpha\Delta^{II}$, an RNAP derivative that (unlike wild-type RNAP) binds with the underivatized αCTD (αCTD^{I}) preferentially in the proximal region. These investigators did not report DNA experiments with an RNAP derivative having the cleaving agent incorporated in αCTD^{I}. We predict that such experiments would likewise indicate preferential binding of the underivatized αCTD, in this case αCTD^{II}, in the proximal subsite region.

Implications for promoter architecture

We have analyzed the *Escherichia coli* genome sequence to estimate the frequency of promoters that contain near-consensus subsites or full UP elements. For the purposes of this discussion, we define near consensus as 0–2 differences from consensus per subsite or 0–4 differences from consensus per full UP element. Table 1 presents the statistics for *E. coli* mRNA, tRNA, or rRNA promoters having near-consensus subsites or full UP elements. Table 2 provides the identities of these promoters.

Several conclusions can be drawn from this analysis. First, numerous *E. coli* promoters contain single near-consensus subsites. Second, promoters with a single near-consensus subsite are significantly more common than promoters with a near-consensus full UP element. Third, near-consensus proximal and distal subsites and full UP elements occur significantly more frequently in stable RNA (rRNA and tRNA) promoters.

It is important to emphasize that several UP element subsites with only a moderate match to consensus have been shown to stimulate transcription by an amount that correlates generally with similarity to consensus (e.g., see Fig. 1, *rrnB* P1 proximal; Ross et al. 1998).

Table 1. *Near-consensus UP elements in* E. coli *promoters*

	mRNA (2501)[a]	tRNA (33)[a]	rRNA (14)[a]	Total (2548)[a]
Consensus full[b]	10 (0.4%)	3 (9.1%)	3 (21%)	16 (0.63%)
Proximal subsite[c]	76 (3.0%)	8 (24%)	5 (36%)	89 (3.5%)
Distal subsite[c]	28 (1.1%)	1 (3%)	3 (21%)	32 (1.3%)

[a]The numbers in parentheses refer to the numbers of promoters searched from the *E. coli* promoter database (provided by A. Huerta and J. Collado-Vides, Universidad Nacional Autónoma de México, Cuernavaca). This database contains confirmed promoters and promoters predicted from sequence analysis (for sequences, see http://www.cifn.unam.mx/Computational_Biology/E.coli-predictions/).
[b]Four or fewer mismatches to consensus.
[c]Two or fewer mismatches to consensus.

Therefore, Tables 1 and 2 (which include only those promoters with near-consensus subsites) underestimate the number of promoters with sequences that are likely to function as UP elements.

The fact that each of the two copies of αCTD in RNAP can interact with an UP element subsite, together with the fact that the two copies of αCTD are flexibly tethered to the remainder of RNAP (Blatter et al. 1994; Jeon et al. 1997), allows for the evolution of additional, more complex classes of UP element-dependent promoters. Thus, promoters exist with functional subsites further upstream than the positions described here (Newlands et al. 1992; Aiyar et al. 1998), with multiple alternative functional distal subsites (Aiyar et al. 1998), or with UP element subsites and adjacent activator protein-binding sites that function cooperatively through αCTD-activator interactions (Murakami et al. 1997b; Belyaeva et al. 1998; Noel and Reznikoff 1998; Law et al. 1999). The modular quality of promoter structure thus provides the potential for multiple input signals to be received by a single transcription initiation complex.

Implications for transcription regulation

Our results establish that consensus proximal subsites, consensus distal subsites, and full UP elements are differently affected by functional inactivation of one αCTD, with consensus proximal subsites showing almost no change in function, consensus distal subsites showing almost complete loss of function, and full UP elements showing partial loss of function (Fig. 7). These differences in effects of functional inactivation of one αCTD potentially can be exploited for differential promoter regulation. For example, bacteriophage T4 Alt catalyzes ADP ribosylation of Arg-265 of one copy of αCTD in RNAP, a post-translational modification that functionally inactivates that copy of αCTD (K. Severinov, W. Ross, H. Tang, L. Snyder, A. Goldfarb, R.L. Gourse, and R.H. Ebright, unpubl.). We expect that Alt-mediated ADP-ribosylation would differentially affect promoters with UP elements containing consensus proximal and/or distal subsites. Furthermore, we speculate that there could be other post-translational modifi-

Table 2. *Promoters in* E. coli *genome with near consensus UP elements*

	Full UP elements (4 or fewer mismatches)	Proximal subsites (2 or fewer mismatches)	Distal subsites (2 or fewer mismatches)
mRNA	add as1A, cspB, cspE, envR, hemL, hisL, ilvGMEDAp1, rpmFp1, 2118180	cspAp1, cspB, dinG, eco, fadL, gcvR, gidB, glnS, gut, hisL, hisS, hrpA, hupA, ilvGMEDAp2, lit, lpp, metG, polA, ppa, purH, recA, rob, srmB, sulA, syd, tdcR, thdF, tpiA, tpx, tsr, ugpp1, xylE, yadD, ybbB, ybeD, yehA, yfiD, yfig, ygfE, ygjE, ygjI, yhdW, yhiS, yibD, yidC, yjbA, yjgP, yjhD, yjiD, yjiT, yjjN, yohJ, 332725, 333657, 389475, 886646, 889312, 914128, 1168296, 1214698, 1431698, 1445540, 1627239, 1631646, 1732459, 1908123, 2183937, 2454832, 2680877, 2783031, 2890237, 2903664, 2983617, 3107570, 3170227, 3203897, 3578769	alpA, cirp2, envR, hdeD, hisL, ilvGMEDAp1, narU, ndk, phnA, ppsA, recN, rpmFp1, tsr, ycgB, yhaI, yhbY, yhiX, yifK, yjfZ, 240189, 535810, 675934, 851820, 1213282, 1215012, 1218824, 1906572, 2118180
tRNA	argX, metT, valU	argX, asnU, aspV, glyW, metT, metZ, serT, serV	valU
rRNA	rrnAp1, rrnBp1, rrnCp1	rrnAp2, rrnBp2, rrnCp2, rrnDp1, rrnGp1	rrnAp1, rrnBp1, rrnCp1

For predicted promoters of unnamed genes, the numerical designations refer to the first position in the open reading frame. For actual promoter sequences, refer to http://www.cifn.unam.mx/Computational_Biology/E.coli-predictions/.

cations, small-molecule effectors, or protein effectors that functionally inactivate one αCTD and thus differentially affect promoters with consensus proximal subsites, distal subsites, and full UP elements.

Materials and methods

Synthesis of promoter populations containing randomized proximal or distal upstream sequences

rrnB P1 promoter fragments used in the first round of in vitro selection were synthesized by annealing partially complementary top and bottom strand oligonucleotides and by use of T7 DNA polymerase as described (Estrem et al. 1998). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA) or the University of Wisconsin Biotechnology Center, or were donated by NSC Technologies (Mt. Prospect, IL). The top strand oligonucleotide contained random sequences in either the proximal or distal UP element subsite. The oligonucleotide with a random proximal subsite contained (from upstream to downstream) an *Eco*RI site, *rrnB* P1 sequence from –66 to –60, 5′-GACTGCAGTGGTA-3′ from –59 to –47 (SUB sequence; Rao et al. 1994), random bases from –46 to –38, and *rrnB* P1 sequence from –37 to +1 (see also Fig. 1). The oligonucleotide with a random distal subsite contained an *Eco*RI site, *rrnB* P1 sequence from –66 to -60, random bases from –59 to –46, 5′-CTAGGAAT-3′ from –45 to –38 (SUB sequence; Rao et al. 1994), and *rrnB* P1 sequence –37 to +1 (see also Fig. 3). The bottom strand oligonucleotide for synthesizing both promoter populations contained a *Hin*dIII site and *rrnB* P1 sequence from +50 to –17. Seventeen proximal and eight distal promoter fragments were sequenced without selection after cloning into phage λ to confirm that the frequencies of each of the 4 bases in the random regions were approximately equal.

UP element selection and screen

The selection was modeled after previous in vitro selections for protein-binding sites on nucleic acids (Blackwell and Weintraub 1990; Pollock and Treisman 1990; Tuerk and Gold 1990; Wright et al. 1991). In the first round of selection, radioactively labeled promoter fragments [0.5 μg; ~3×10^{12} DNA molecules, that was in excess of the 5×10^6 (4^9) or 6.4×10^9 (4^{14}) molecules needed to ensure that all sequence combinations were represented in the proximal subsite or distal subsite selections, respectively] were incubated with RNAP for 4 min, and bound fragments were separated from unbound by gel electrophoresis as described previously (Estrem et al. 1998). For the second and subsequent rounds of selection, promoter fragments were amplified by PCR from gel-isolated RNAP-promoter complexes (Estrem et al. 1998). The PCR primers contained all of the nonrandomized promoter positions to reduce the frequency of PCR-generated mutations in the core promoter region that might increase binding by RNAP (Estrem et al. 1998). RNAP-binding reactions were carried out under progressively more stringent conditions (lower RNAP concentration and shorter reaction times). The progress of the selection was monitored by sequencing representatives of the selected populations following eight (for the proximal) and six (for the distal) rounds of selection. A total of 13 cycles of RNAP binding, separation on gels, and PCR were carried out for each selection.

In vitro-selected promoters were fused to *lacZ* in phage λ and screened for high promoter activity on MacConkey lactose indicator plates (Estrem et al. 1998). The promoter regions of the selected *lacZ* fusions were sequenced after PCR of DNA obtained directly from plaques. Three promoters from the proximal subsite selection and nine from the distal subsite selection were discarded, because they contained deletions or core promoter mutations. β-Galactosidase activities were determined from monolysogens of strain NK5031 that had grown exponentially at least three generations in Luria–Bertani medium (LB; Ross et al. 1998).

Site-directed promoter mutations

rrnB P1 promoters (–66 to +50) containing only the *rrnB* P1 proximal or distal subsite sequences (RLG3098 and RLG3099, respectively) or containing proximal subsite 4547 with single base pair substitutions were synthesized by PCR with mutagenic top strand oligonucleotides and bottom strand oligonucleotides complementary to the plasmid vector as described previously (Ross et al. 1998).

In vitro transcription

Promoter fragments were cloned into pRLG770 (Ross et al.

1990). Supercoiled DNA concentrations were determined both spectrophotometrically and by quantitation of the vector encoded RNA I transcripts under conditions of RNAP excess (40 nM). Transcription was carried out as described previously (Ross et al. 1993), except that reactions contained 0.6 nM DNA and different salt concentrations (see Figs. 5 and 7). Reconstituted RNAPs (Gaal et al. 1996; Tang et al. 1996) were used at concentrations that resulted in equivalent transcription from the *lacUV5* promoter (2.7 nM for RNAP containing wild-type α, 9 nM for α R265A, 17.4 nM αΔ235), and oriented-α RNAP derivatives (see below) were used at concentrations that resulted in equivalent transcription from the *rrnB* P1 promoter lacking UP element sequences (2 nM [α^I/α^{II}], 3.5 nM [$\alpha^I/\alpha^{II}\Delta$], 24 nM [$\alpha^I\Delta/\alpha^{II}$]). Gels were analyzed by PhosophorImager (Molecular Dynamics).

Hydroxyl radical footprinting

rrnB P1 promoter templates with different UP elements were generated by PCR from plasmids pRLG4213 (UP element 4547), pRLG4214 (4513), pRLG3278 (4192), and pRLG4238 (–66 to +50 *rrnB* P1 promoter) with vector-specific primers, digested at a primer-encoded *Bam*HI site (upstream of the *Eco*RI site), and end labeled with α[^{32}P]dGTP (DuPont). Labeled fragments were purified, and hydroxyl radical footprint reactions were performed as described previously, except that in footprinting reactions with purified α, the buffer contained 50 mM KCl and 9 mM NaCl (Estrem et al. 1998), and in oriented-α RNAP footprints, the RNAP-binding reaction was for 30 min at 37°C; RNAP-bound complexes were isolated from 5% acrylamide gels following hydroxyl radical cleavage, eluted by diffusion, and purified with an Elutip (Schleicher & Schuell). Footprinting reactions were analyzed on 10% acrylamide-8 M urea gels and quantified by phosphorimaging with ImageQuant software (Molecular Dynamics), normalized with Microsoft Excel, and plotted with SigmaPlot 4.0. Normalization of scans was done using a region outside of the protein-binding site (–65 to –80) to correct for loading differences. Band intensities in different lanes were calculated as ratios of the lane with protein to that without protein. The scans are presented as a sliding average to correct for slight differences and nonlinearities in electrophoretic migration.

Construction of oriented-α RNA polymerases

Plasmids pREII-NHα, pREII-NHα(1–235), pREII-NHα45A, and pREII-NHα45A(1–235) encode amino-terminally hexahistidine-tagged α, αΔ235, [Ala-45]α, and [Ala-45]αΔ235, respectively, under control of the tandem *lpp*P–'*lac*PUV5 promoter, confer ampicillin resistance, and have pBR322-derived origins of replication (Niu et al. 1996; W. Niu, and R.H. Ebright, in prep.). Plasmid pWN-NFα(1–235) encodes amino-terminally Flag (Kodak)-tagged αΔ235 under the control of the tandem *lpp*P–'*lac*PUV5 promoter, confers kanamycin resistance, and has a pSC101-derived origin of replication (W. Niu and R.H. Ebright, in prep.).

Wild-type RNAP(α^I/α^{II}) and oriented-α heterodimeric RNAP α^I/[Ala-45]α$\Delta235^{II}$ RNAP ($\alpha^I/\alpha^{II}\Delta$) were prepared from strains XL1-Blue/pREII-NHα and XL1-Blue/pREII-NHα45AΔ235, respectively, using Ni^{2+}–NTA agarose chromatography and Mono-Q chromatography as described by Niu et al. (1996). Oriented-α heterodimeric RNAP $\alpha\Delta235^I$/[Ala-45]α^{II} RNAP ($\alpha^I\Delta/\alpha^{II}$) was prepared from strain XL1-Blue/pREII-NHα45A/pWN-NFαΔ235 cultured in 4 × LB containing 200 μg/ml ampicillin and 20 μg/ml kanamycin, with Ni^{2+}–NTA agarose chromatography and Mono-Q chromatography as described by Niu et al. (1996), followed by Anti-Flag M2 immunoaffinity chromatography as follows. Samples were dialyzed against 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, and 5% glycerol; adsorbed onto 1 ml of anti-Flag M2 affinity gel (Kodak); washed with 3 × 12 ml of the same buffer; and eluted with 2 × 0.5 ml each of the same buffer containing 50 μg/ml, 75 μg/ml, 100 μg/ml, and 200 μg/ml Flag peptide (Kodak). Peak fractions were pooled, dialyzed against 25 mM Tris-HCl (pH 7.9), 100 mM NaCl, 0.1 mM EDTA, 0.1 mM dithiothreitol, and 50% glycerol and stored in aliquots at –20°C.

E. coli *genome sequence analysis*

Sequences of 253 confirmed mRNA promoters, 2248 predicted mRNA promoters, 33 tRNA promoters, and 14 rRNA promoters from *E. coli* were obtained from Araceli Huerta and Julio Collado-Vides (http://www.cifn.unam.mx/Computational_Biology/E.coli-predictions/). Each promoter sequence was 31 nucleotides in length, including the proposed –35 hexamer and 25 nucleotides upstream of the –35 hexamer. Sequences were searched in GCG version 9.0 using the command FINDPATTERNS with the parameters –DAT = AAAAAARNR(N){7,7} –ONE–MIS = 2 for the proximal subsite, –DAT = AWWWWWTTTTT(N){16,16} –ONE–MIS = 2 for the distal subsite, and –DAT = AAAWWTWTTTTNNNAAASNN(N){7,7} –ONE–MIS = 4 for the full UP element.

Acknowledgments

This work was supported by National Institutes of Health grants GM37048 to R.L.G. and GM41376 to R.H.E. and by an Howard Hughes Medical Institute Investigatorship to R.H.E. S.T.E. was supported by a National Institutes of Health Genetics predoctoral training grant and by a Hatch grant from the United States Department of Agriculture. Z.W.S.C. was supported by a National Science Foundation program, Research Experience for Undergraduates. We thank Araceli Huerta and Julio Collado-Vides for providing the *E. coli* promoter database and we thank Katsuhiko Murakami and Akira Ishihama for providing enzymes used in early stages of these experiments.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked 'advertisement' in accordance with 18 USC section 1734 solely to indicate this fact.

References

Aiyar, S.E., R.L. Gourse, and W. Ross. 1998. Upstream A-tracts increase bacterial promoter activity through interactions with the RNA polymerase a subunit. *Proc. Natl. Acad. Sci.* **95:** 14652–14657.

Bartlett, M.S. and R.L. Gourse. 1994. Growth rate-dependent control of the *rrnB* P1 core promoter in *Escherichia coli*. *J. Bacteriol.* **176:** 5560–5564.

Belyaeva, T.A., V.A. Rhodius, C.L. Webster, and S.J. Busby. 1998. Transcription activation at promoters carrying tandem DNA sites for the *Escherichia coli* cyclic AMP receptor protein: Organisation of the RNA polymerase alpha subunits. *J. Mol. Biol.* **277:** 789–804.

Blackwell, T.K. and H. Weintraub. 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250:** 1104–1110.

Blatter, E.E., W. Ross, H. Tang, R.L. Gourse, and R.H. Ebright. 1994. Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78:** 889–896.

Busby, S. and R.H. Ebright. 1994. Promoter structure, promoter

recognition, and transcription activation in prokaryotes. *Cell* **79:** 743–746.

Czarniecki, D., R.J. Noel, Jr. and W.S. Reznikoff. 1997. The –45 region of the Escherichia coli *lac* promoter: CAP-dependent and CAP-independent transcription. *J. Bacteriol.* **179:** 423–429.

Dombroski, A.J., W.A. Walter, M.T. Record, Jr., D.A. Siegele, and C.A. Gross. 1992. Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell* **70:** 501–512.

Estrem, S.T. 1998. 'The identification of an UP element consensus and characterization of its interaction with the α subunit of RNA polymerase,' Ph.D. thesis. University of Wisconsin, Madison, WI.

Estrem, S.T., T. Gaal, W. Ross, and R.L. Gourse. 1998. Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci.* **95:** 9761–9766.

Fredrick, K., T. Caramori, Y.F. Chen, A. Galizzi, and J.D. Helmann. 1995. Promoter architecture in the flagellar regulon of *Bacillus subtilis*: High-level expression of flagellin by the sigma D RNA polymerase requires an upstream promoter element. *Proc. Natl. Acad. Sci.* **92:** 2582–2586.

Gaal, T., W. Ross, E.E. Blatter, H. Tang, X. Jia, V.V. Krishnan, N. Assa-Munt, R.H. Ebright, and R.L. Gourse. 1996. DNA-binding determinants of the α subunit of RNA polymerase: Novel DNA-binding domain architecture. *Genes & Dev.* **10:** 16–26.

Gourse R.L. 1988. Visualization and quantitative analysis of complex formation between *E. coli* RNA polymerase and an rRNA promoter *in vitro. Nucleic Acids Res.* **16:** 9789–9809.

Hawley, D.K. and W.R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11:** 2237–2255.

Helmann, J.D. and P.L. deHaseth. 1999. Protein-nucleic acid interactions during open complex formation investigated by systematic alteration of the protein and DNA binding partners. *Biochemistry* **38:** 5959–5967.

Heyduk, T., E. Heyduk, K. Severinov, H. Tang, and R.H. Ebright. 1996. Determinants of RNA polymerase α subunit for interactions with β, β′, and α subunits: Hydroxyl-radical protein footprinting. *Proc. Natl. Acad. Sci.* **93:** 10162–10166.

Hochschild, A. and S.L. Dove. 1998. Protein-protein contcats that activate and repress prokaryotic transcription. *Cell* **92:** 597–600.

Igarashi, K. and A. Ishihama. 1991. Bipartite functional map of the *E. coli* RNA polymerase alpha subunit: Involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **65:** 1015–1022.

Jeon, Y.H., T. Yamazaki, T. Otomo, A. Ishihama, and Y. Kyogoku. 1997. Flexible linker in the RNA polymerase alpha subunit facilitates the independent motion of the C-terminal activator contact domain. *J. Mol. Biol.* **267:** 953–962.

Josaitis, C.A., T. Gaal, W. Ross, and R.L. Gourse. 1990. Sequences upstream of the –35 hexamer of *rrnB* P1 affect promoter strength and upstream activation. *Biochim. Biophys. Acta* **1050:** 307–311.

Kielkopf, C.L., S. White, J.W. Szewczyk, J.M. Turner, E.E. Baird, P.B. Dervan, and D.C. Rees. 1998. A structural basis for recognition of A-T and T-A base pairs in the minor groove of B-DNA. *Science* **282:** 111–115.

Kimura, M. and A. Ishihama. 1995. Functional map of the alpha subunit of *Escherichia coli* RNA polymerase: Amino acid substitution within the amino-terminal assembly domain. *J. Mol. Biol.* **254:** 342–349.

Koo, H.S., H.M. Wu, and D.M. Crothers. 1986. DNA bending at adenine-thymine tracts. *Nature* **320:** 501–506.

Law, E., N. Savery, and S. Busby. 1999. Interactions between the *Escherichia coli* cyclic AMP receptor protein and the C-terminal domain of the α subunit of RNA polymerase at Class I promoters. *Biochem J.* **237:** 415-423.

Leirmo, S. and R.L. Gourse. 1991. Factor-independent activation of rRNA transcription. I. Kinetic analysis of the roles of the upstream activator region and supercoiling on the *rrnB* P1 promoter in vitro. *J. Mol. Biol.* **220:** 555–568.

Li, M., W.R. McClure, and M.M. Susskind. 1997. Changing the mechanism of transcriptional activation by phage κ repressor. *Proc. Natl. Acad. Sci.* **94:** 3691–3696.

Lonetto, M.A., V. Rhodius, K. Lamberg, P. Kiley, S. Busby, and C. Gross. 1998. Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase σ^{70} subunit. *J. Mol. Biol.* **284:** 1353–1365.

Murakami, K., N. Fujita, and A. Ishihama. 1996. Transcription factor recognition surface on the RNA polymerase alpha subunit is involved in contact with the DNA enhancer element. *EMBO J.* **15:** 4358–4367.

Murakami, K., M. Kimura, J.T. Owens, C.F. Meares, and A. Ishihama. 1997a. The two alpha subunits of *Escherichia coli* RNA polymerase are asymmetrically arranged and contact different halves of the DNA upstream element. *Proc. Natl. Acad. Sci.* **94:** 1709–1714.

Murakami, K., J.T. Owens, T.A. Belyaeva, C.F. Meares, S.J. Busby, and A. Ishihama. 1997b. Positioning of two alpha subunit carboxy-terminal domains of RNA polymerase at promoters by two transcription factors. *Proc. Natl. Acad. Sci.* **94:** 11274–11278.

Newlands, J.T., W. Ross, K.K. Gosink, and R.L. Gourse. 1991. Factor-independent activation of *Escherichia coli* rRNA transcription II characterization of complexes of *rrnB* P1 promoters containing or lacking the upstream activator region with *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **220:** 569–583.

Newlands, J.T., C.A. Josaitis, W. Ross, and R.L. Gourse. 1992. Both fis-dependent and factor-independent upstream activation of the *rrnB* P1 promoter are face of the helix dependent. *Nucleic Acids Res.* **20:** 719–726.

Newlands, J.T., T. Gaal, J. Mecsas, and R.L. Gourse. 1993. Transcription of the *Escherichia coli rrnB* P1 promoter by the heat shock RNA polymerase (E sigma 32) in vitro. *J. Bacteriol.* **175:** 661–668.

Niu, W., Y. Kim, G. Tau, T. Heyduk, and R.H. Ebright. 1996. Transcription activation at class II CAP-dependent promoters: Two interactions between CAP and RNA polymerase. *Cell* **87:** 1123–1134.

Noel, R. and W.S. Reznikoff. 1998. CAP, the –45 region, and RNA polymerase: Three partners in transcription initiation at *lacP1* in *Escherichia coli. J. Mol. Biol.* **282:** 495–504.

Pollock, R. and R. Treisman. 1990. A sensitive method for the determination of protein-DNA binding specificities. *Nucleic Acids Res.* **18:** 6197–6204.

Rao, L., W. Ross, J.A. Appleman, T. Gaal, S. Leirmo, P.J. Schlax, M.T. Record, Jr., and R.L. Gourse. 1994. Factor independent activation of *rrnB* P1. An "extended" promoter with an upstream element that dramatically increases promoter strength. *J. Mol. Biol.* **235:** 1421–1435.

Ross, W., J.F. Thompson, J.T. Newlands, and R.L. Gourse. 1990. *E.coli* Fis protein activates ribosomal RNA transcription in vitro and in vivo. *EMBO J.* **9:** 3733–3742.

Ross, W., K.K. Gosink, J. Salomon, K. Igarashi, C. Zou, A. Ishihama, K. Severinov, and R.L. Gourse. 1993. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262:** 1407–1413.

Ross, W., S.E. Aiyar, J. Salomon, and R.L. Gourse. 1998. *Escherichia coli* promoters with UP elements of different strength: Modular structure of bacterial promoters. *J. Bacteriol.* **180:** 5375–5383.

Savery, N.J., G.S. Lloyd, M. Kainz, T. Gaal, W. Ross, R.H. Ebright, R.L. Gourse, and S.J. Busby. 1998. Transcription activation at Class II CRP-dependent promoters: Identification of determinants in the C-terminal domain of the RNA polymerase alpha subunit. *EMBO J.* **17:** 3439–3447.

Seeman, N.C., J.M. Rosenberg, and A. Rich. 1976. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci.* **73:** 804–808.

Strainic, M.G., J.J. Sullivan, A. Velevis, and P.L. deHaseth. 1998. Promoter recognition by *Escherichia coli* RNA polymerase: Effects of the UP element on open complex formation and promoter clearance. *Biochemistry* **37:** 18074–18080.

Tang, H., Y. Kim, K. Severinov, A. Goldfarb, and R.H. Ebright. 1996. *Escherichia coli* RNA polymerase holoenzyme: Rapid reconstitution from recombinant alpha, beta, beta′, and sigma subunits. *Methods Enzymol.* **273:** 130–134.

Tuerk, C. and L. Gold. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249:** 505–510.

Wright, W.E., M. Binder, and W. Funk. 1991. Cyclic amplification and selection of targets (CASTing) for the myogenin consensus binding site. *Mol. Cell Biol.* **11:** 4104–4110.

Young, M.A., J. Srinivasan, I. Goljer, S. Kumar, D.L. Beveridge, and P.H. Bolton. 1995. Structure determination and analysis of local bending in an A-tract DNA duplex: Comparison of results from crystallography, nuclear magnetic resonance, and molecular dynamics simulation on d(CGCAAAAAT-GCG). *Methods Enzymol.* **261:** 121–144.

Zhou, Y., T.J. Merkel, and R.H. Ebright. 1994. Characterization of the activating region of *Escherichia coli* catabolite activator protein (CAP). II. Role at Class I and Class II CAP-dependent promoters. *J. Mol. Biol.* **243:** 603–610.