Fine structure of the promoter–σ region 1.2 interaction

Shanil P. Haugen, Wilma Ross, Michele Manrique, and Richard L. Gourse*

Department of Bacteriology, University of Wisconsin, 1550 Linden Drive, Madison, WI 53706

We recently proposed that a nontemplate strand base in the discriminator region of bacterial promoters, the region between the −10 element and the transcription start site, makes sequence-specific contacts to region 1.2 of the σ subunit of *Escherichia coli* RNA polymerase (RNAP). Because rRNA promoters contain sequences within the discriminator region that are suboptimal for interaction with σ1.2, these promoters have the kinetic properties required for regulation by the RNAP-binding factors DksA and ppGpp. Here, we use zero-length cross-linking and mutational, kinetic, and footprinting studies to map RNAP interactions with the nontemplate strand bases at the junction of the −10 element and the discriminator region in an unregulated rRNA promoter variant and in the λP$_R$ promoter. Our studies indicate that nontemplate strand bases adjacent to the −10 element bind within a 9-aa interval in σ1.2 (residues 99–107). We also demonstrate that the downstream-most base on the nontemplate strand of the −10 hexamer cross-links to σ region 2, and not to σ1.2. Our results refine models of RNAP–DNA interactions in the promoter complex that are crucial for regulation of transcription initiation.

promoter element | RNA polymerase | transcription initiation | discriminator region | −10 element

Interactions between bacterial RNA polymerase holoenzyme (RNAP; $\alpha_2\beta\beta'\omega\sigma$) and the promoter can determine not only its basal strength but also its regulation. (In this report, σ always refers to σ^{70}, the major σ factor.) Four promoter elements are generally recognized as making sequence-specific contacts with RNAP (1, 2) (Fig. 1): the UP element (bound by the C-terminal domains of the two α subunits); the −35 hexamer (bound by σ region 4.2); the extended −10 element (bound by σ region 3.0); and the −10 hexamer (bound by σ region 2.3–2.4). Recently, an additional element immediately downstream of the −10 hexamer, within the discriminator region, was proposed to bind to σ region 1.2 (3, 4).

The term "discriminator" was coined by Travers (5) more than 25 years ago to describe a G+C-rich region downstream from the −10 hexamer in stable RNA (rRNA and tRNA) promoters, and it was proposed that the G+C content of this region was important for maintaining proper regulation of stable RNA promoters (6, 7). High G+C content was proposed to impede strand separation, leading to promoter regulation. However, it was found that a C to G substitution 2 nt downstream from the −10 hexamer in the rRNA promoter *rrnB* P1 (*rrnB* P1 C-7G) eliminated its regulation, suggesting that the actual sequence of the discriminator region, in addition to its high G+C content, is crucial for control of transcription (3).

Footprinting, photocross-linking, and genetic approaches led to the conclusion that the nontemplate strand base two positions downstream from the −10 element in the *rrnB* P1 C-7G promoter contacts σ1.2. When the base at the analogous position in all other promoters investigated was a C (either naturally or by mutation), competitor-resistant complexes formed with RNAP were much shorter-lived than the same promoters with G at this position, suggesting that the σ1.2 interaction with this element in the discriminator region can occur in the context of most/all σ^{70}-dependent promoter sequences and that this contact can contribute to the longevity typical of open complexes. Surveys of *Escherichia coli* promoter sequences show no preference for a specific base at this position, in stark contrast to the strong preference for C at this

position in rRNA and tRNA promoter sequences (ref. 8 and data not shown). These observations led to the proposal that rRNA promoters have evolved to make weak σ1.2 contacts with the discriminator element, resulting in short-lived competitor-resistant complexes that are therefore susceptible to the effects of regulatory factors such as ppGpp and DksA (3).

σ1.2 (residues ≈96–127) is evolutionarily well conserved among group 1 and group 2 σ factors (9), suggesting that it has a crucial function. Furthermore, even before its precise role in sequence-specific promoter recognition was identified, several substitutions in σ1.2 were shown to render RNAP defective for transcription initiation (10). Although structural information is not available for the proposed σ1.2–RNAP interaction because the available RNAP–DNA cocrystal does not contain DNA downstream of the −10 element (11), crystal structures of RNAP holoenzymes from *Thermus aquaticus* and *Thermus thermophilus* in the absence of promoter DNA revealed that σ1.2 consists of two α helices oriented at ≈90° with respect to one another (12, 13). Some models of the promoter complex predict that residues 93–108 would be in close proximity to the nontemplate strand of the discriminator (3, 14), but other models place the discriminator region further away from σ1.2 (15). It has also been suggested that σ1.2 controls −10 element recognition (16).

To refine our understanding of transcription initiation and its regulation, in this work we use a range of approaches to define promoter–RNAP interactions at the junction of the −10 hexamer and the discriminator region. We localize two discriminator base contacts to a 9-aa segment in σ1.2, confirming the role of σ1.2 in sequence-specific promoter recognition, and we demonstrate that the most downstream base in the nontemplate strand of the −10 hexamer interacts with σ2, not σ1.2. We suggest that similar methods can be used to construct higher resolution models of other parts of the promoter complex.

Results

rrnB P1 C-7G Cross-Links to σ1.2 Between Residues 99 and 107. Previously, we localized the σ1.2–discriminator interaction to residues 99–132 (3), too long a segment to position the interaction precisely in models of the open complex. To define the interaction more precisely, we created an *rrnB* P1 template containing a 6-thiodeoxyguanine at position −7 in the promoter (2 nt downstream from the −10 hexamer) and performed cross-linking experiments with RNAPs reconstituted with wild-type (WT) σ, σ that had been engineered to contain a single cysteine residue at Cys-107, or σ lacking the first 98 aa (Δ1.1 RNAP) (3). Thio-substituted bases minimally disrupt DNA structure, are highly photoreactive, will cross-link to a variety of amino acids, allow detection of essentially zero-length interactions, and form cross-links at higher wave-

Fig. 1. Schematic diagram (not to scale) of sequence-specific promoter–RNAP interactions. Lines connect promoter elements (*Lower*) with the RNAP modules in α or σ that bind to them (*Upper*) (1, 2). UP Element, upstream element. −35 and −10 hexamers are indicated. Ext, extended −10 element; Dis, discriminator element; +1, transcription start site; CTD, α C-terminal domain; NCR, nonconserved region in σ.

lengths of UV where cross-links to non-thio-substituted bases are minimized (17).

After cross-linking, σ was purified by using an N-terminal hexahistidine affinity tag and digested with the cysteine-specific cleavage reagent 2-nitro-5-thiocyanobenzoic acid (NTCBA). As is typical with this reagent, multiple bands were observed because the digestions do not go to completion and because of low levels of nonspecific cleavage (Fig. 2). Nevertheless, unambiguous conclusions could be drawn. The patterns for RNAP with WT σ (cysteines at 132, 291, and 295; Fig. 2B, lane 2) and with σΔ1.1 (lane 1) were the same as observed (3). With σΔ1.1, every cleavage product was smaller than the corresponding cleavage product from WT σ, indicating that the cross-link is to the N-terminal fragment of σΔ1.1, residues 99–132 (for a schematic diagram, see Fig. 2C). Because the smallest fragment from digestion of the WT σ (1–132; lane 2) migrated slightly slower than the smallest band from digestion of Cys-107 RNAP (lane 3), we conclude that the smallest fragment in lane 3 contains residues 1–107. Because fragments containing residues 99–132 and residues 1–107 both cross-linked to the promoter with the photoactivated base at −7, we conclude that the cross-link maps to the interval between residues 99 and 107.

Fig. 2. Identification of amino acid interval in σ region 1.2 that contacts promoter position −7 in the discriminator region of *rrnB* P1 C-7G. (*A*) −10 element and discriminator region sequence in the *rrnB* P1 C-7G promoter used for cross-linking. X at −7 is 6-thiodeoxyguanine. (*B*) Mapping of the *rrnB* P1 C-7G cross-link to amino acids 99–107 in σ. RNAPs were reconstituted with Δ1.1 σ (lane 1), with WT σ (lane 2), or with σ containing a single cysteine at position 107 (lane 3). WT σ contains cysteines at positions 132, 291, and 295. After cross-linking, complexes were treated with NTCBA (cleaves at cysteines) and examined by SDS/PAGE and PhosphorImaging. Fragments from cleavages at positions 291 vs. 295 are not distinguishable in these profiles. Bands that migrate slightly faster than those labeled 99–291/5 and 1–291/5 likely derive from a nonspecific cleavage and usually were formed at much lower efficiency than those labeled 99–291/295 or 1–291/5 (see Fig. 5). (*C*) Position of cross-link (star), inferred from *B*.

Fig. 3. Lifetimes of competitor-resistant complexes formed by the WT *rrnB* P1 or *rrnB* P1 C-7G promoters with σ1.2 alanine-substituted RNAP mutants. Half-lives of complexes were determined with a transcription assay (*SI Materials and Methods*). (*A*) Absolute half-lives. White bars, WT *rrnB* P1. Black bars, *rrnB* P1 C-7G. (*B*) Ratio of half-lives, *rrnB* P1 WT to *rrnB* P1 C-7G complex, by using RNAPs with single-alanine substitutions in σ1.2.

Alanine Substitutions in σ1.2 Alter the Lifetime of the Promoter Complex. If the *rrnB* P1 C-7G promoter formed long-lived competitor-resistant complexes because of a specific interaction between σ1.2 and the G residue at position −7, we reasoned that substitution(s) in σ1.2 that preferentially decreased the lifetime of the C-7G complex (relative to their effects on the WT *rrnB* P1 complex) would be good candidates for participating in interactions with the discriminator region. Therefore, we purified σ subunits with single-alanine substitutions at every position between residues 99 and 107, assembled them with core RNAP to form the holoenzyme, and determined the half-lives of competitor-resistant complexes formed on the *rrnB* P1 C-7G and the WT *rrnB* P1 promoters (Fig. 3).

In contrast to the short-lived complex formed by WT RNAP with WT *rrnB* P1, WT RNAP formed a long-lived complex with the *rrnB* P1 C-7G promoter (≈80 min; Fig. 3A), as reported (3). None of the alanine substitutions had large effects on the lifetime of the complex formed by the WT promoter, as expected if the interaction of σ1.2 with the discriminator region in this promoter was already weak. In contrast, two of the alanine substitutions in σ1.2, Y101A and M102A, preferentially decreased the lifetime of the *rrnB* P1 C-7G complex (by 20- to 30-fold), to approximately the same lifetime as that of the WT promoter complex. The substitutions flanking Y101A and M102A also reduced the lifetime of the complex formed by *rrnB* P1 C-7G, but not as much as Y101A and M102A (Fig. 3A).

To illustrate the specific effects of the alanine-substituted σ subunits on promoter complex half-life, i.e., their abilities to distinguish between a nontemplate C or G at promoter position −7, Fig. 3B shows the ratio of the lifetime of the RNAP complex containing WT *rrnB* P1 relative to that containing *rrnB* P1 C-7G. Elimination of either the Tyr-101 or the Met-102 amino acid side chain in σ resulted in WT and C-7G *rrnB* P1 promoter complexes with similar absolute half-lives (ratio only slightly <1.0; Fig. 3B). In contrast, the other alanine-substituted RNAPs had ratios ≪1.0, more similar to that of the WT RNAP. The inability of the σY101A and σM102A RNAPs to distinguish between the WT and C-7G *rrnB* P1 promoters made these amino acids likely candidates for interaction with the discriminator region.
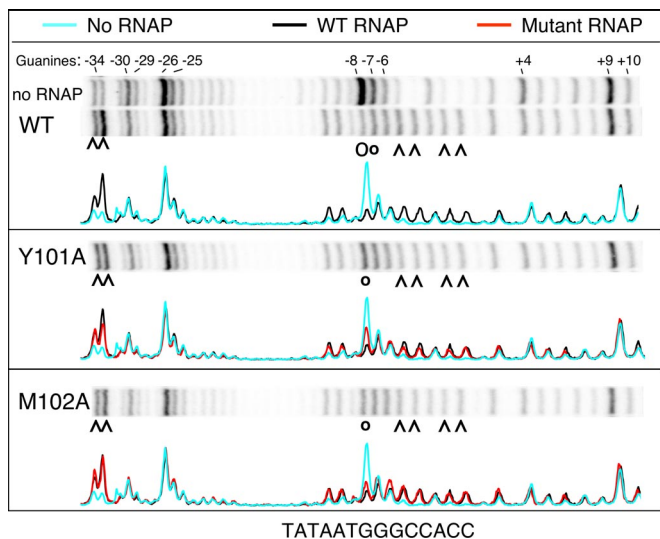
BIOCHEMISTRY

Fig. 4. DMS protection footprints of the *rrnB* P1 C-7G promoter (nontemplate strand) with WT RNAP and RNAPs containing single-alanine substitutions in σ. Methylation of guanines, adenines, and unpaired cytosines by DMS arrests *Taq* polymerase, generating DNA fragments whose end points correspond to the position of the modification. Bands were not detected in controls with untreated DNA (data not shown). The identity of the RNAP is indicated at the left in each footprint. Carets indicate enhancements; circles show protections. Scans of the no-RNAP lane (blue line), the WT RNAP lane (black line), and the σ mutant RNAP lane (red line) are superimposed as indicated. (*Middle*) Y101A. (*Bottom*) M102A. Enhanced DMS reactivities of G residues in the −35 hexamer and C residues in the strand opening region are discussed in *Results*.

Alanine Substitutions in σ1.2 Alter Protection of Discriminator Region Bases by RNAP. Previous dimethyl sulfate (DMS) protection footprints with WT RNAP and *rrnB* P1 C-7G promoter fragments showed that the guanine bases on the nontemplate strand at −8 and −7 were protected by RNAP (3, 18), indicating that RNAP made a very close approach to the discriminator region. We performed DMS protection footprints with a subset of the alanine-substituted σ RNAPs to assess their effects on discriminator region contacts. Because several of the mutant RNAPs formed short-lived complexes on *rrnB* P1 C-7G and because complexes decay faster on linear DNA than supercoiled templates (19), we performed the complexes on supercoiled templates, modified the DNA with DMS, and then detected the methylated bases by arrest of primer extension (see ref. 20 and Fig. 4 legend).

As observed, WT RNAP strongly protected position G-8 in the *rrnB* P1 C-7G promoter and also protected G-7, but to a lesser extent (3) [compare blue line (no RNAP) with black line (WT RNAP) in the scans in Fig. 4]. Enhanced DMS reactivity was observed at positions in the −35 hexamer with all of the RNAPs, as observed (3), and at C residues −5, −4, −2, and −1. The signals at −5 to −1 have also been observed (18) and likely result from methylation of distorted or melted cytosines in the open complex. These enhancements served as an internal control, indicating that the promoter was bound by the mutant RNAPs to extents similar to those of WT RNAP under these conditions.

The footprints formed by the WT and mutant RNAPs differed only in the discriminator, supporting the model that σ1.2 interacts with this region of the promoter. The σY101A RNAP (*Middle*, red line) and σM102A RNAP (*Bottom*, red line) reproducibly displayed less protection of G-8 than the WT RNAP, and the protection at G-7 was completely lost. Therefore, in conjunction with the cross-linking data and effects of the alanine substitutions on complex lifetime, shown above, the footprinting data suggest that either

Tyr-101 or Met-102 in σ (or both) contacts the discriminator region directly.

The Discriminator Region in λP_R Cross-Links to σ1.2. The cross-linking, footprinting, and kinetic results described above and in ref. 3 indicate that the *rrnB* P1 C-7G discriminator region interacts with σ1.2. Effects of a C vs. G at the analogous position in several other promoters (i.e., 2 nt downstream from the −10 hexamer) on the lifetimes of promoter complexes suggested that an interaction with σ1.2 might occur at those promoters as well (3). To address the generality of the discriminator region–σ1.2 interaction and its presence in a naturally occurring promoter, we performed cross-linking experiments with λP_R, in which the WT sequence contains Gs at each of the two positions just downstream from the −10 hexamer, −5 and −6. [In contrast to *rrnB* P1, where transcription starts 9 bp downstream from the −10 hexamer, transcription starts 7 bp downstream from the −10 hexamer in λP_R (Fig. 2*A*, Fig. 5*A*). Therefore, position −5 in λP_R corresponds to −7 in *rrnB* P1, −6 in λP_R corresponds to −8 in *rrnB* P1, and −7 in λP_R corresponds to −9 in *rrnB* P1.] Double-stranded λP_R templates were created containing a single photoreactive nt analog (6-thiodeoxyguanine) on the nontemplate strand either at −5 (λP_R −5) or at −6 (λP_R −6). A template was also created containing the photoreactive nt analog 4-thiothymidine at the highly conserved −7 position [the "invariant T" (8)], the most downstream position in the −10 hexamer (λP_R −7).

With WT RNAP, all three templates cross-linked primarily to σ, although weaker bands corresponding in size to either β or β' were also observed [supporting information (SI) Fig. 7*A*]. The identities of these weaker cross-links have not yet been pursued.

The cross-links made by λP_R −5 (Fig. 5*B*) and λP_R −6 (Fig. 5*C*) were mapped as described above for the cross-links to the *rrnB* P1 C-7G template (Fig. 2 and ref. 3) by comparing the cleavage patterns of complexes formed by WT RNAP, σΔ1.1 RNAP, and RNAPs containing σ variants with single cysteines at residues 95 (Cys-95 RNAP) or 107 (Cys-107 RNAP; see also schematic diagrams in Fig. 5*D*). The cleavage patterns of WT and σΔ1.1 RNAP on λP_R −5 (Fig. 5*B*, lanes 1 and 2) and λP_R −6 (Fig. 5*C*, lanes 1 and 2) closely resembled the patterns observed with *rrnB* P1 C-7G (Fig. 2*B* and ref. 3), indicating that the cross-links were between σ residues 99 and 132. A large fragment was generated after cross-linking and digestion of Cys-95 (Fig. 5*B*, lane 4, and 5*C*, lane 3), indicating that the cross-link was to the C-terminal fragment, σ residues 95–613. In contrast, Cys-107 produced a fragment corresponding in size to residues 1–107 (Fig. 5*B*, lane 5, and 5*C*, lane 4). Taken together, the data indicate that the bases on the nontemplate strand in λP_R one and two positions downstream from the −10 hexamer cross-link within the interval 99–107 in σ.

Cross-linking of RNAPs with the promoter fragment containing the photoreactive base on the nontemplate strand at the most downstream position in the −10 hexamer, λP_R −7 (Fig. 5*E*), resulted in patterns different from those formed by the λP_R templates with the photoreactive base at −6 or −5. NTCBA digestion of WT σ (lanes 1 and 3) or σ containing a single cysteine at residue 132 (Cys-132 RNAP; lane 2) produced a large fragment of identical mobility. Because WT and Cys-132 σ each contain cysteine-132, but the fragment cross-linking to position −7 is quite large, it likely consists of residues 132–613 and not the 1–132 fragment that contains σ1.2. However, the identity of the smallest fragment generated from WT σ (identified below as 291/295–613) could not be determined from these data alone because two potential fragments (291/295–613 or the partial digestion product 1–291/295) would be of similar size.

RNAPs containing σ with a single cysteine at Cys-376 (in region 2.1) or at Cys-442 (in region 2.4) were used to map the λP_R −7 cross-link more precisely (Fig. 5*E*, lanes 3–5). Cross-linking and digestion of Cys-376 σ (lane 4) resulted in a band that migrated faster than the smallest product from digestion of WT σ (1–291/295
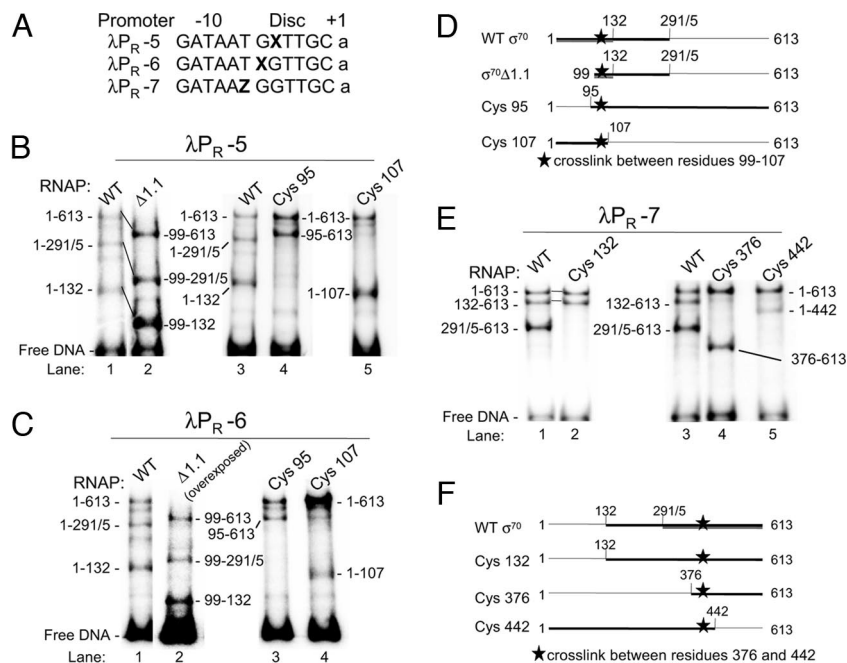
Fig. 5. Mapping of cross-links to λP$_R$ nontemplate promoter bases G-5, G-6, and T-7. (A) −10 element and discriminator region sequences in the λP$_R$ cross-linking templates. X is 6-thiodeoxyguanine, and Z is 4-thiothymidine. (B) Mapping of λP$_R$ −5 cross-link to residues 99–107. Lanes 1 and 2 and lanes 3–5 are from separate gels. (C) Mapping of λP$_R$ −6 cross-link to residues 99–107. A longer exposure is provided for the σΔ1.1 RNAP (lane 2) because of poor recovery in the experiment shown. (D) Schematic diagrams of cross-links with λP$_R$ −5 and λP$_R$ −6. (E) Mapping of λP$_R$ −7 cross-link to residues 376–442. (F) Schematic diagrams of cross-link with λP$_R$ −7.

or 291/295–613) (lane 3). The size of this fragment suggests that it contains the smaller of the two potential digestion products, the 376–613 fragment (237 aa), because a fragment extending from residue 1 to 376 would likely have migrated slower than either the 1–291/295 or 291/295–613 products. Because the cross-link is C-terminal to residue 376, the smallest cross-linked fragment in the WT digest must contain residues 291/295–613.

Because cross-linking and digestion of Cys-442 σ (lane 5) resulted in a band that migrated between the WT products (132–613 and 291/295–613) and slower than the 376–613 fragment, this band must be fragment 1–442. In conjunction with the results reported above, we conclude that the cross-link between λP$_R$ −7 and σ maps between residues 376 and 442, within σ region 2 (amino acids ≈375–452).

Effects of Substitutions in σ1.2 on RNAP–Promoter Interactions. We next used cross-linking efficiency as a semiquantitative means of identifying residues within σ1.2 likely to interact with the discriminator region DNA bases. Promoter complexes were formed from λP$_R$ −5, λP$_R$ −6, or λP$_R$ −7 and RNAPs containing WT σ or σ subunits with single-alanine substitutions for residues 99–107. Cross-links were induced with UV, and the efficiency of cross-link formation to σ was assessed by SDS/PAGE (Fig. 6A). Promoter binding by the mutant RNAPs was assessed in parallel by filter binding. Under these conditions, σY101A RNAP was only ≈70% as efficient as WT RNAP in forming competitor-resistant λP$_R$ complexes, but the other RNAPs bound promoter DNA as efficiently as WT RNAP (data not shown).

With the λP$_R$ −5 template, the M102A substitution in σ almost completely eliminated cross-linking (≈16% as efficient as WT RNAP), whereas the other σ mutants had much smaller effects. (The cross-linked band was a doublet; the weak upper band was unaffected by any of the alanine substitutions in this region of σ and was not included in the quantitation.) The R99A, M100A, and R103A RNAPs cross-linked with slightly higher efficiency than WT RNAP. The basis for these increases is unclear, although in theory each of these substitutions could subtly alter the local environment in the complex, creating conditions more favorable for cross-linking. We conclude that Met-102 is the most likely residue in σ1.2 to interact with the base at −5 in λP$_R$, consistent with the results with *rrnB* P1, where the M102A substitution caused the greatest

decrease in complex half-life and discrimination between the WT and the C-7G promoters (Fig. 3).

Several of the σ1.2 alanine substitutions increased the efficiency of cross-linking of RNAP to λP$_R$ −6, but none reduced cross-linking as dramatically as M102Aσ on the λP$_R$ −5 template. The cross-linking efficiency of RNAP containing Y101Aσ was ≈60% of that for WT RNAP, but this number may overestimate the reduction because this RNAP was only 70% as efficient as WT RNAP in forming competitor-resistant λP$_R$ complexes under these conditions (see above). Although these data do not identify the residue(s) in σ1.2 that cross-link to λP$_R$ −6, they reinforce the conclusion that the interaction of Met-102 is specific to the base at −5 (see *Discussion*).

RNAP cross-links to λP$_R$ −7 were nearly abolished by the Y101A substitution in σ1.2, and the M102A substitution reduced cross-linking efficiency to 30% of that with WT RNAP. Because the λP$_R$ −7 cross-link mapped to σ region 2 and not to σ1.2, we conclude the Y101A and M102A substitutions (and the σ1.2–discriminator region interaction) must affect the −10 hexamer–σ region 2 interaction indirectly (see ref. 16 and *Discussion*).

Discussion

Interactions with σ at the −10 Hexamer–Discriminator Junction. The results reported here refine our understanding of the promoter–RNAP complex by demonstrating that a 9-aa segment in σ1.2 interacts directly with the region of the promoter just downstream from the −10 hexamer. Our data also demonstrate that the invariant T at the downstream edge of the −10 hexamer, an extremely conserved base in bacterial promoters (8), contacts σ region 2. Our data thus define the boundaries in the interacting surfaces of σ at the junction of the −10 element and the discriminator region.

Multiple lines of evidence suggest that Met-102 in σ contacts the base two positions downstream from the −10 hexamer: (*i*) M102A strongly reduced the lifetime of the promoter complex, and RNAP containing this mutant σ was unable to distinguish between a nontemplate strand G and C at position −7 in *rrnB* P1 (Fig. 3). (*ii*) M102A RNAP lost the ability to protect G-7 in *rrnB* P1 C-7G in DMS footprints (Fig. 4). (*iii*) M102A dramatically reduced RNAP cross-linking efficiency to the λP$_R$ −5 template (Fig. 6A). (*iv*) Met-102 is surface-exposed in holoen-

BIOCHEMISTRY

Fig. 6. Identification of residues in $\sigma 1.2$ interacting with the nontemplate strand of the discriminator. (*A*) Cross-linking efficiencies of templates with alanine-substituted RNAPs. (*Upper*) Representative gels for each template. (*Lower*) Histograms show averages and standard deviations for each RNAP (at least three experiments for each promoter). Standard deviations for RNAPs with largest effects (M102A RNAP with λP_R -5, and Y101A RNAP with λP_R -7) were <8%. Cross-linking efficiency is normalized to efficiency with WT RNAP (dashed line). (*B*) Model of path of discriminator region DNA in *Thermus* open complex, modified slightly from ref. 14 to illustrate discriminator interaction with *E. coli* $\sigma 1.2$ identified in this work. RNAP is oriented so that the viewer is looking into the secondary channel. Template strand is purple, nontemplate strand is red, β and β' are gray, and most of σ is in blue spacefill. Residues in σ corresponding to *E. coli* Met-102 and Arg-99 are in light blue and pink, respectively. Other residues in helix 1 of $\sigma 1.2$ are in vivid yellow, and the remainder of $\sigma 1.2$ is in pale yellow. Tyr-101 is buried (adjacent to Met-102) on the opposite face of the helix. In the model, $\sigma 2.3$ (light orange) most closely approaches the most downstream base in the -10 hexamer.

zyme structures and is positioned appropriately for interaction with the discriminator region in models of the open complex (Fig. 6*B*). Contacts between methionine side chains and DNA have been reported (21–23), but they are far from common.

An alignment of this segment of region 1.2 in σ homologs is provided in SI Table 1. The alignment indicates that this region is highly conserved in evolution, supporting its functional importance, that Tyr-101 is virtually universal, and that only conservative substitutions (leucine or isoleucine) are found in place of Met-102.

Although the Y101A substitution dramatically affected interactions with the discriminator region, Tyr-101 is a less likely candidate

for the contacting residue, primarily because it is not surface-exposed in the available x-ray structures of RNAP (Fig. 6*B* and legend). In theory, helix movement upon DNA binding could expose Tyr-101 to solvent (and DNA); if so, an attractive model would be that the tyrosine side chain stacks on the base at position -5 in a flipped out conformation, such as has been proposed for interactions between aromatic amino acid side chains in σ region 2 and base(s) in the -10 hexamer (11, 24).

DMS footprints of *rrnB* P1 C-7G with R99A RNAP resulted in dramatic enhancements at positions G-6 and G-7, without altering protection of G-8 (SI Fig. 7). R99A also resulted in a relatively strong cross-link between the λP_R -5 template and a higher molecular weight band (most likely β and/or β'), in addition to the cross-link to σ (SI Fig. 7). Although the role of σ residue Arg-99 remains uncertain, we suggest that Arg-99 could contact the DNA backbone, constraining the path of the DNA and preventing the discriminator region from interacting with β and/or β'. The R99A substitution could disrupt this backbone interaction and/or change the conformation of σ, allowing greater discriminator region access to both DMS and β and/or β'.

Strong protection of the base immediately downstream of the -10 hexamer (-8 in *rrnB* P1, -6 in λP_R) from methylation by DMS indicates that RNAP also closely approaches this nt (18, 25). This base cross-linked to $\sigma 1.2$, and alanine substitutions within $\sigma 1.2$ reduced protection of this base from methylation by DMS (Fig. 4). However, the identity of the amino acid residue(s) interacting with this nt is uncertain because none of the $\sigma 1.2$ substitutions abolished cross-linking to G-6 in λP_R (Fig. 6*A*). Multiple explanations could account for this result. (*i*) More than one amino acid could interact with this base, so that one alanine substitution would not be sufficient to eliminate cross-linking. (*ii*) If the principal contact were eliminated by mutation, a local rearrangement in the protein might allow neighboring side chains to form a cross-link. (*iii*) Because thio-substituted nucleotides are not chemoselective (17), the substituted alanine might be able to cross-link to the reactive base. In any case, because a G to C mutation at -8 in *rrnB* P1, the analogous position in that promoter, had no effect on complex half-life or on regulation (3), the consequence (if any) of the interaction between $\sigma 1.2$ and the base at $-8(-)$ on transcription initiation remains to be determined.

Interactions of σ with the -10 Hexamer. The T:A base pair at the last position in the -10 hexamer is highly conserved in bacterial promoters (8) and crucial for open complex formation (26). Nontemplate strand bases further upstream in the -10 element interact with σ region 2 (1, 11). Our observation that nontemplate bases adjacent to the -10 hexamer interact with $\sigma 1.2$ raised the question as to which part of σ contacts the downstream-most base in the -10 element. We show here that the nontemplate base at this position (-7 in λP_R, -9 in *rrnB* P1) cross-links within $\sigma 2$.

The σ substitutions M102A and Y101A both diminished cross-linking of RNAP to λP_R -7 (Fig. 6*A*). Thus, $\sigma 1.2$ must indirectly affect interactions between $\sigma 2$ and the -10 hexamer, supporting the observations of Zenkin *et al.* (16) that $\sigma 1.2$ plays a role in -10 hexamer recognition by $\sigma 2$. Tyr-101 is buried in the structure of RNAP holoenzyme (refs. 12 and 13 and Fig. 6*B*), as discussed above, and it interacts intimately with both the β' coiled-coil and σ region 2.1, the surfaces most critical for binding of σ to core RNAP. It would therefore not be surprising if the Y101A substitution indirectly affected -10 element interactions with RNAP in addition to disrupting the discriminator region interaction with the adjacent amino acid, Met-102. Our results further emphasize the central importance of $\sigma 1.2$ in general and of Tyr-101 in particular for promoter recognition.

Role of $\sigma 1.2$–Discriminator Interactions in Regulation of Transcription Initiation. The results described here and in our previous work (3) suggest that the discriminator element in many, if not all, Eσ^{70}-

dependent promoters contacts σ1.2, but the strength of this interaction is sequence-dependent. Our results are consistent with the proposal that the optimal nontemplate sequence for this interaction is 5′-GGG-3′ for the 3 nt adjacent to the −10 hexamer (4), and the lifetime of promoter complexes increases with the strength of this interaction (3). We emphasize that the consequence of the interaction on transcriptional output depends on the intrinsic kinetics of the promoter; stronger interactions with σ1.2 appear to improve transcription only if the promoter complex has an intrinsically short lifetime (3). At rRNA promoters, stronger interactions with σ1.2 abolish transcription regulation (3).

Although the identity of the base two positions downstream of the −10 hexamer does not affect the rate of formation of closed complexes (3, 27), σ1.2 interactions with DNA could potentially affect other transcriptional events. For example, recent results from our laboratory suggest that the σ1.2–discriminator region interaction affects *rrnB* P1 transcription start site selection (P. Chandrangsu, S.P.H., W.R., and R.L.G., unpublished data). Furthermore, a pseudo-10 element downstream from the transcription start site for the late promoter of bacteriophage λ is followed by G residues that appear to serve as a pseudodiscriminator element, facilitating the promoter proximal pause required for λQ-dependent antitermination (28).

Future Directions. We propose that rRNA promoters have evolved to make a suboptimal interaction between the discriminator element and σ1.2, which contributes to the kinetic properties of the rRNA promoter complex that facilitate its regulation by factors that bind to RNAP. In theory, these or other factors could also modulate the discriminator element–σ1.2 interaction directly to regulate promoters. Finally, we also note that the methods used here could be used to identify contacts with RNAP at other positions in the promoter and/or to define promoter contacts with RNAP in the intermediates that precede open complex formation.

Materials and Methods

Plasmids and Proteins. Plasmids containing the *rrnB* P1 promoter and variants and *rpoD* constructs coding for σ and variants, are listed in SI Table 2, and their construction is described in *SI Materials and Methods*. Core RNAP and mutant and WT σ subunits were purified as described (3, 29).

Cross-Linking and Mapping. *rrnB* P1 C-7G template construction and cross-linking were performed as described in ref. 3. λP$_R$ templates containing a zero-length cross-linker (see Fig. 5A) were prepared by annealing three oligonucleotides including a thio-substituted nucleotide analog at the specified position (Trilink Biotechnology). The annealed oligonucleotides were ligated, and the fragment was then gel-purified. After UV irradiation, cross-linked complexes were either separated by PAGE or were purified on Ni-agarose, treated with NTCBA, and then analyzed on gels (3). Further details are presented in *Results*, the figure legends, *SI Materials and Methods*, and ref. 3.

Competitor-Resistant Complex Decay. Decay rates were measured by using a transcription-based assay as described in ref. 3 and *SI Materials and Methods*. Briefly, promoter complexes were formed with WT RNAP or the σ mutant RNAPs on supercoiled plasmids containing WT *rrnB* P1 or *rrnB* P1 C-7G promoters. The fraction of complexes remaining at times after addition of competitor was determined from the amount of RNA product produced after the addition of NTP (3).

DMS Protection Footprinting. The procedure was modified from ref. 20. Promoter complexes were formed on supercoiled plasmids, and DNA bases protected by RNAP from modification by DMS were detected by primer extension. Additional details are in *Results* and *SI Materials and Methods*.

1. Helmann J, deHaseth P (1999) Protein–nucleic acid interactions during open complex formation investigated by systematic alteration of the protein and DNA-binding partners. *Biochemistry* 38:5959–5967.
2. Gourse R, Ross W, Gaal T (2000) UPs and downs in bacterial transcription initiation: The role of the α subunit of RNA polymerase in promoter recognition. *Mol Microbiol* 37:687–695.
3. Haugen S, et al. (2006) rRNA promoter regulation by nonoptimal binding of σ region 1.2: An additional recognition element for RNA polymerase. *Cell* 125:1069–1082.
4. Feklistov A, et al. (2006) A basal promoter element recognized by free RNA polymerase σ subunit determines promoter recognition by RNA polymerase holoenzyme. *Mol Cell* 23:97–107.
5. Travers A (1980) Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. *J Bacteriol* 141:973–976.
6. Lamond A, Travers A (1985) Stringent control of bacterial transcription. *Cell* 41:6–8.
7. Pemberton I, Muskhelishvili G, Travers A, Buckle M (2000) The G+C-rich discriminator region of the *tyrT* promoter antagonises the formation of stable preinitiation complexes. *J Mol Biol* 299:859–864.
8. Harley C, Reynolds R (1987) Analysis of *E. coli* promoter sequences. *Nucleic Acids Res* 15:2343–2361.
9. Lonetto M, Gribskov M, Gross C (1992) The σ^{70} family: Sequence conservation and evolutionary relationships. *J Bacteriol* 174:3843–3849.
10. Baldwin N, Dombroski A (2001) Isolation and characterization of mutations in region 1.2 of *Escherichia coli* σ^{70}. *Mol Microbiol* 42:427–437.
11. Murakami K, Masuda S, Campbell E, Muzzin O, Darst S (2002) Structural basis of transcription initiation: An RNA polymerase holoenzyme–DNA complex. *Science* 296:1285–1290.
12. Murakami K, Masuda S, Darst S (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* 296:1280–1284.
13. Vassylyev D, et al. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417:712–719.
14. Lawson C, et al. (2004) Catabolite activator protein: DNA binding and transcription activation. *Curr Opin Struct Biol* 14:10–20.
15. Artsimovitch I, et al. (2004) Structural basis for transcription regulation by alarmone ppGpp. *Cell* 117:299–310.
16. Zenkin N, et al. (2007) Region 1.2 of the RNA polymerase σ subunit controls recognition of the −10 promoter element. *EMBO J* 26:955–964.
17. Meisenheimer K, Koch T (1997) Photocross-linking of nucleic acids to associated proteins. *Crit Rev Biochem Mol Biol* 32:101–140.
18. Newlands J, Ross W, Gosink K, Gourse R (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.
19. Leirmo S, Gourse R (1991) Factor-independent activation of *Escherichia coli* rRNA transcription. I. Kinetic analysis of the roles of the upstream activator region and supercoiling on transcription of the *rrnB* P1 promoter *in vitro*. *J Mol Biol* 220:555–568.
20. Artsimovitch I, Kahmeyer-Gabbe M, Howe M (1996) Distortion in the spacer region of Pm during activation of middle transcription of phage Mu. *Proc Natl Acad Sci USA* 93:9408–9413.
21. Raumann B, Rould M, Pabo C, Sauer R (1994) DNA recognition by β-sheets in the Arc repressor–operator crystal structure. *Nature* 367:754–757.
22. Gao Y, et al. (1998) The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA. *Nat Struct Biol* 5:782–786.
23. Murphy F, Sweet R, Churchill M (1999) The structure of a chromosomal high mobility group protein–DNA complex reveals sequence-neutral mechanisms important for non-sequence-specific DNA recognition. *EMBO J* 18:6610–6618.
24. Panaghie G, Aiyar S, Bobb K, Hayward R, de Haseth P (2000) Aromatic amino acids in region 2.3 of *Escherichia coli* σ^{70} participate collectively in the formation of an RNA polymerase–promoter open complex. *J Mol Biol* 299:1217–1230.
25. Siebenlist U, Simpson R, Gilbert W (1980) *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell* 20:269–281.
26. Roberts C, Roberts J (1996) Base-specific recognition of the nontemplate strand of promoter DNA by *E. coli* RNA polymerase. *Cell* 86:495–501.
27. Cook V, deHaseth P (2007) Strand opening-deficient *Escherichia coli* RNA polymerase facilitates investigation of closed complexes with promoter DNA: Effects of DNA sequence and temperature. *J Biol Chem* 282:21319–21326.
28. Ring B, Yarnell W, Roberts J (1996) Function of *E. coli* RNA polymerase σ factor σ^{70} in promoter-proximal pausing. *Cell* 86:485–493.
29. Ross W, Schneider D, Paul B, Mertens A, Gourse R (2003) An intersubunit contact stimulating transcription initiation by *E coli* RNA polymerase: Interaction of the α C-terminal domain and σ region 4. *Genes Dev* 17:1293–1307.

BIOCHEMISTRY