

A "master" in base unpairing during isomerization of a promoter upon RNA polymerase binding

Heon Man Lim*[†], Hee Jung Lee*[†], Siddhartha Roy*[‡], and Sankar Adhya*[§]

*Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892-4255; [†]Department of Biology, College of Natural Sciences, Chungnam National University, Taejeon 305-764, Republic of Korea; and [‡]Department of Biophysics, Bose Institutes, Calcutta 700054, India

Contributed by Sankar Adhya, September 28, 2001

Isomerization of a closed to open complex of a promoter upon RNA polymerase binding involves base unpairing at the -10 region. After potassium permanganate sensitivity of unpaired thymine residues, we studied base unpairing at the -10 region during isomerization upon RNA polymerase binding at the $P1$ and $P3$ promoters of the *gal* operon. Substitution of adenine by 2-amino purine (2-AP) at the invariable A-T base pair at the -11 position of $P1$ and $P3$ prevented unpairing not only at that position but also at the other downstream positions, suggesting a "master" role of the adenine base at -11 of the template strand in overall base unpairing. 2-AP at -11 did not inhibit the formation of RNA polymerase-promoter complex and subsequent isomerization of the polymerase. Substitution of adenine by 2-AP at several other positions did not affect thymine unpairing. Changing the position of the amino group from C6 in adenine to C2 in 2-AP is mutational only at the master switch position, -11 .

Under ordinary physiological conditions, transcription in *Escherichia coli* is executed by the RNA polymerase holoenzyme comprising $\alpha_2\beta\beta'\sigma^{70}$ subunits. Transcription initiation follows specific binding of RNA polymerase to the promoter. The initial binary complex (closed complex) undergoes structural changes (isomerization) in both proteins and DNA to make an open complex that is competent for subsequent template-dependent polymerization of ribonucleotides (1, 2). Whether isomerization of RNA polymerase and DNA occurs simultaneously or one after the other is unknown. The σ^{70} subunit plays a key role in recognizing and making specific contacts with base pairs at the -35 and -10 regions of the promoter (3). In the open complex, DNA at the -10 region presumably becomes single-stranded for RNA polymerase to start reading the exposed template strand and polymerizing ribonucleotides (4–10). The bases in the nontemplate strand make specific interactions with amino acid residues of σ^{70} (11). The consequence of these interactions is believed to stabilize the unstable nature of the unpaired DNA region (11–13). The consensus DNA sequence of the -10 region is 5'-TATAAT-3' in the nontemplate strand, with the first T being at the -12 position (1). To elucidate the molecular procedures of base unpairing during open complex formation, altered forms of DNA (depurinated, nicked, forked, or with single-stranded gap) in the -10 region were used as templates (10, 14–19). These studies suggested that base unpairing involves an initial distortion or localized melting event around -11 (16, 20). In this paper, we provide results by using intact duplex DNA to demonstrate that the adenine residue itself at the position -11 of the nontemplate strand plays a master role in the initiation of base unpairing.

Materials and Methods

Materials. *E. coli* RNA polymerase holoenzyme was purchased from Epicentre Technologies (Madison, WI). cAMP receptor protein (CRP), purified to 98% homogeneity by FPLC (Amersham Pharmacia) from an *E. coli* strain carrying the *crp* gene in a multicopy plasmid pHA5, was a gift from S. Garges (National Cancer Institute, Bethesda). Synthetic oligodeoxynucleotides containing 2-amino purine (2-AP) substitutions (106-bases long)

were obtained from Genosys (The Woodlands, TX) or Trilink BioTechnologies (San Diego). The final product of synthetic oligonucleotides was purified by PAGE.

Preparations of DNA Templates Containing 2-AP. The 2-AP-containing DNA templates were prepared by PCR in which one of the primers was 2-AP containing synthetic oligo. The second primer corresponded to the segment downstream to a transcription terminator (Fig. 1). The DNA template for PCRs was plasmid pSA852. We prepared 358-bp-long PCR DNA fragments that generate a 125-nt-long RNA transcript from the *gal P1* promoter in *in vitro* transcription reactions. After the PCR, the 358-bp DNA fragments were purified by electroelution (Bio-Rad) followed by ethanol precipitation.

In Vitro Transcription. *In vitro* transcription reactions were performed at 25°C in a 50- μ l volume following the procedure of Choy and Adhya (21). The initial reaction mixture (45 μ l) contained 10 nM DNA template (2 nM for supercoiled pSA852), 20 nM RNA polymerase, 20 mM Tris acetate, pH 7.5, 10 mM magnesium acetate, 200 mM potassium glutamate, and 40 units of rRNasin (Promega). CRP, when present, was added to the initial reaction mixture at 50 nM. The cAMP contraction was 100 μ M. The mixture was incubated at 25°C for 20 min, and 5 μ l of NTP (2 mM of ATP, GTP, and CTP, 0.2 mM UTP, and 5 μ Ci [α -³²P]UTP) was added. The mixture was incubated further for 10 min at 25°C, and the reaction was terminated by the addition of 50 μ l of STOP solution (BRL). Samples were boiled for 2 min, and 3 μ l of each sample was loaded onto 8% polyacrylamide-urea sequencing gels to analyze the RNA. Quantitation of the transcripts was done by using a PhosphorImager (Molecular Dynamics).

DNase I Footprinting Assay. The complex formation reaction was performed in a 50- μ l transcription buffer at 25°C for 20 min. The final concentrations of RNA polymerase and DNA were 50 nM and 4 nM, respectively. This reaction was treated with DNase I (1.5 units, BRL) for 4 min at 25°C. DNase I was inactivated by a phenol/chloroform (1:1) extraction followed by a G-50 spin column (Amersham Pharmacia). The pattern was generated by PCR using a ³²P-labeled single primer followed by a DNA-sequencing gel electrophoresis.

Electrophoretic Mobility-Shift Assay. The complex formation was identical to that of the DNase I footprinting assay. One microliter of properly diluted RNA polymerase was mixed with 10 μ l of ³²P-labeled 358-bp DNA template. The final concentrations of RNA polymerase for each reaction are shown in Fig. 4. The

Abbreviations: CRP, cAMP receptor protein; 2-AP, 2-amino purine.

[§]To whom reprint requests should be addressed at: Laboratory of Molecular Biology, National Cancer Institute, Building 37, Room 2E16, 37 Convent Drive, Bethesda, MD 20892-4255. E-mail: sadhya@helix.nih.gov.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

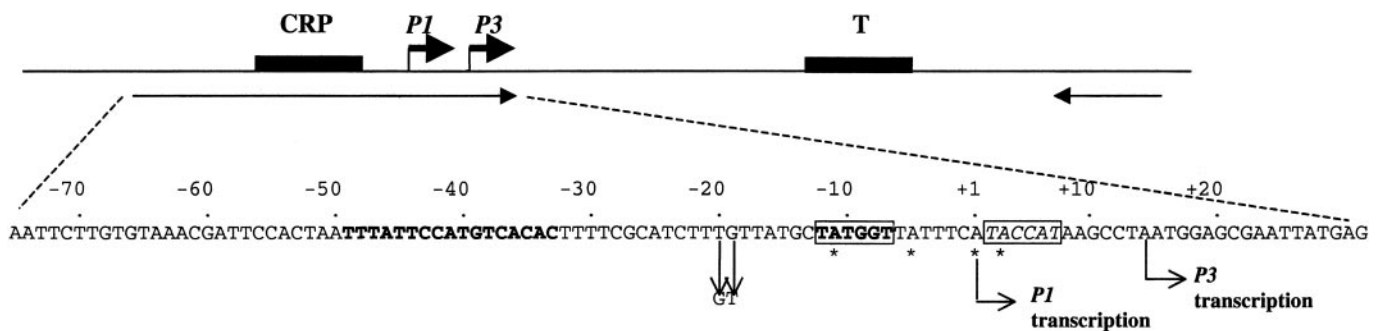


Fig. 1. The *gal* promoter region of *E. coli* showing the *P1* and *P3* promoters in the plasmid pSA852 (D. Lewis, unpublished data). pSA852 contains a 288-bp segment of the *gal* control region (−197 to +91) followed by a transcription terminator (21). The CRP binding site is in bold. The transcription start site of *P1* is denoted as +1. The −10 region, the *P1* promoter, is in bold and boxed, and that of *P3* is boxed. The adenine residues with the asterisks were replaced individually by 2-AP in the 106-base-long rightward primer. DNA templates (385 bp) for *in vitro* transcription containing 2-AP were generated by PCR using the rightward and leftward primers (shown by arrows) from pSA852. The 106-base-long primer also contained −20T to G and −19T to G changes to inactivate the *P2* promoter. T, transcription termination signal.

DNA concentration was 200 pM. The reaction was incubated at 25°C for 20 min followed by the addition of heparin to a final concentration of 50 μg/ml. Five microliters of the reaction was analyzed by 4% PAGE.

Potassium Permanganate Assay. The complex formation was identical to that of the DNase I footprinting assay. The reaction was treated with 5 μl of 80 mM KMnO₄ and quenched by the addition of 5 μl of 14.4 M β-mercaptoethanol followed by a phenol/chloroform (1:1) extraction. The DNA was cleaned further by G-50 spin column and precipitated by ethanol. The modified residues were identified by PCR using ³²P-labeled primers followed by electrophoresis in DNA-sequencing gel.

Results and Discussion

2-AP is an adenine analog in which the amino group is present at the C2 of the purine ring instead of at C6. 2-AP can substitute for adenine and pair with thymine without changing much of the physicochemical properties of double-stranded polynucleotide (22–26). 2-AP acts as an adenine and pairs with thymine when present in the template strand during polynucleotide synthesis by DNA polymerase (see the experiments mentioned in the Fig. 5 legend). Yet as shown below, 2-AP has an effect on transcription initiation only when present in a specific position in the −10 region of the *P1* and *P3* promoters of the *gal* operon of *E. coli*. 2-AP-containing DNA (Fig. 1) was generated by PCR for use as templates of transcription by RNA polymerase *in vitro*. The *gal* operon has been shown to contain four promoters: *P1*, *P2*, *P3*, and *P4*[†] (27–29). Of these, *P2* was inactivated by substitution mutations (Fig. 1), and *P4*, which is located further upstream, was left out in these templates made by PCR. Thus, these templates showed synthesis of 135- and 121-nt-long RNA molecules corresponding to the *P1* and *P3* promoters as described below. We replaced the four adenine residues at positions −11, −5, +1, and +3 in the −10 region of the nontemplate strand of *P1* by 2-AP individually. Wild-type DNA and DNA containing 2-AP at position −5 or +3 produced as much *P1* RNA as from wild type (Fig. 2). The template with 2-AP at position +1 made *P1* RNA at 10% of the wild-type level. However, the RNA molecules from *P1* were conspicuously absent from the template with 2-AP at position −11. The results of defective *P1* RNA synthesis with 2-AP at the −11 position were interesting considering the minor structural difference between adenine

and 2-AP and similarities in the physical and biochemical properties of adenine and 2-AP-containing DNA as mentioned. When the adenine at position −11 of the nontemplate strand of the template was replaced by guanine, cytosine, or thymine, the resulting *P1* promoter also became totally defective in the corresponding RNA synthesis (data not shown). An adenine → thymine transversion at the −11 position was shown previously to inactivate the *P1* promoter both *in vivo* and *in vitro* (30, 31). Together, these results showed that the adenine at the −11 position of *P1* is critical for transcription initiation and are consistent with adenine being the most invariant at this position in the consensus sequence (32).

The 121-nt-long RNA from the *P3* promoter was not observed in wild-type supercoiled DNA template but was found only with PCR-generated linear templates (Fig. 2). In the latter examples, *P3* was made generally in lower amounts relative to *P1* RNA both from wild type and 2-AP-containing DNA templates, as was observed previously (33), except when the 2-AP was present at position +3. With 2-AP at +3, no *P3* RNA was observed. The +3 position of *P1* corresponds to the −11 position of *P3* (33). Thus similar to *P1*, transcription initiation from *P3* was abolished when 2-AP replaced adenine at the corresponding −11 position.

CRP in conjunction with cAMP enhances the *P1* promoter and represses *P3* (21, 31, 33). Consistently, the presence of cAMP and CRP stimulated transcription 3–4-fold from *P1* and inhibited the same from *P3* with DNA templates without 2-AP (Fig. 2). The cAMP- and CRP-mediated stimulation of *P1* RNA synthesis was observed also with templates having 2-AP present at position −5, +1, or +3 by the same amount as was observed with DNA without 2-AP. Although the intrinsic amount of *P1* RNA synthesis was significantly lower in the template with 2-AP at the +1 position, CRP resurrected this *P1* defect to the wild-type level. Unlike 2-AP at +1, cAMP and CRP barely helped the defect of *P1* transcription in DNA containing 2-AP

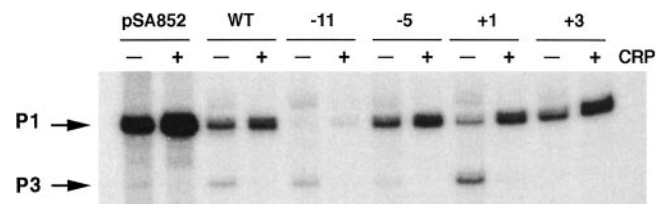


Fig. 2. *In vitro* transcription of the *P1* promoter containing 2-AP at various positions are indicated in the figure in the absence (−) or presence (+) of CRP. pSA852, supercoiled DNA template containing no 2-AP; WT, DNA containing no 2-AP; arrows, *P1* and *P3* transcripts.

[†]The newly discovered *P4* promoter also was named as *P3* in the original publication (27). We have changed that to *P4* in this report, with the permission of the authors, to avoid confusion.

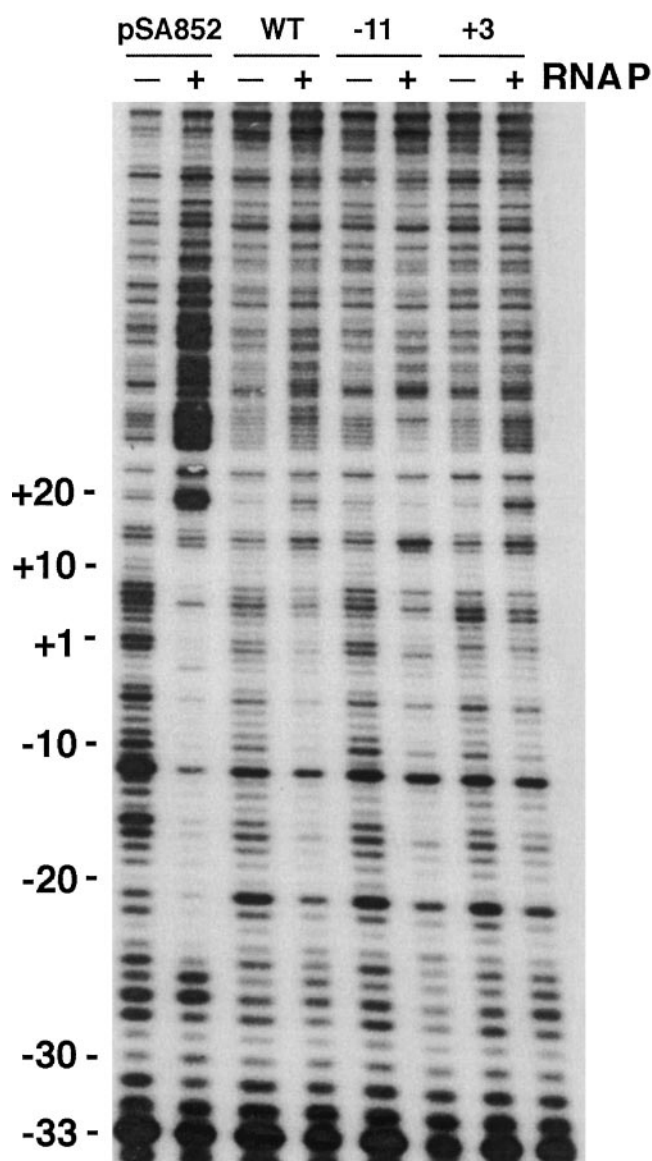


Fig. 3. DNase I footprinting of the DNA templates in the absence (–) or presence (+) of RNA polymerase (RNAP). pSA852 is supercoiled plasmid DNA. WT, DNA containing no 2-AP; –11 and +3, the positions of 2-AP in the DNA, when present.

at position –11. cAMP and CRP inhibited transcription from *P3* not only with the wild-type DNA, as shown previously (33), but also in DNA containing 2-AP at position –11, –5, or +1.

The binding of RNA polymerase to wild type and 2-AP-containing DNA templates was followed by DNase I protection experiments and electrophoretic mobility-shift assays. The binding of RNA polymerase to the *P3* promoter was not detectable under our experimental conditions, perhaps because of a very low affinity of *P3*. The results of DNase I protection showed that the presence of RNA polymerase affected the digestion pattern of the DNA segment from the –28 to +14 position of *P1* not only in DNA without 2-AP (both supercoiled and PCR-generated), as expected, but also in DNA containing 2-AP at position –11 (Fig. 3). The effect was less pronounced for DNA with 2-AP at position +3. The results suggested RNA polymerase binds to the *P1* promoter with 2-AP at position –11. The binding was confirmed by gel electrophoresis experiments in the presence of heparin (Fig. 4). The formation of a heparin-resistant complex

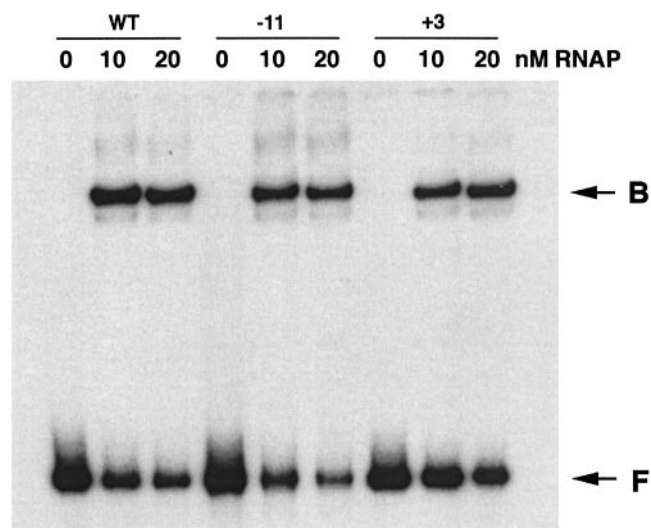


Fig. 4. Electrophoretic mobility-shift assays for RNA polymerase (RNAP) binding to 2-AP-containing DNA templates in the presence of heparin. WT, DNA containing no 2-AP; –11 and +3, the positions of 2-AP in DNA, when present; B, the RNA polymerase-DNA binary complex; F, unbound DNA.

demonstrates structural changes, i.e., isomerization of RNA polymerase. DNA without 2-AP and DNA with 2-AP at position –11 or +3 complexed with RNA polymerase, thus retarding the mobility of the bound DNA in the same amount and to the same extent relative to free DNA during electrophoresis. Thus, 2-AP substitution at position –11, although it did not allow transcription, did not prevent binding and isomerization of RNA polymerase.

We studied base unpairing at the –10 region of DNA containing 2-AP at different positions. Base unpairing was monitored by the treatment of the DNA-protein complexes with KMnO_4 (34). KMnO_4 generates hydroxy radicals, which detect unpaired thymines in DNA duplexes. The location of the unpaired thymines in both template and nontemplate strands was detected by primer extension (35). The unpairing of thymine residues in open complexes at the *P1* promoter has been identified previously by KMnO_4 reaction (36, 37). These authors showed that the major opening of thymines occurred at the –11, –5, and +3 positions in the nontemplate strand. Fig. 5 shows the results of thymine unpairing with or without 2-AP-containing DNA. Thymine unpairing was observed under our assay conditions at positions –11 and +3 in the nontemplate strand and at positions –2, +2, +13, and +16 in the template strands of DNA without 2-AP. Of these, thymines at position –11 in the nontemplate strand and at –2 and +2 in the template strand were unpaired during open complex formation at the *P1* promoter. The thymine unpairing at +3 in the nontemplate strand and at +13 and +16 in the template strand were caused by open complex formation at *P3*; unpairing at these positions was not observed in supercoiled DNA as expected from the transcription results described above. Note that the +3, +13, and +16 positions correspond to the –11, –2, and +2 positions of *P3*. We investigated the effect of 2-AP substitutions of adenine opposite to thymines at positions –11, –5, +1, and +3 on thymine unpairing by KMnO_4 treatment (Fig. 5). It is clear that 2-AP at positions –5, +1, and +3 did not affect the thymine unpairing pattern at the *P1* promoter; thymine residues at positions –11, –2, and +2 were sensitive to KMnO_4 as with DNA without 2-AP. Interestingly, the presence of 2-AP at position –11 of *P1* has a pleiotropic effect; it not only abolished unpairing of thymine opposite to 2-AP at –11 but also failed to show unpairing at positions –2 and +2 of *P1*. The presence of 2-AP

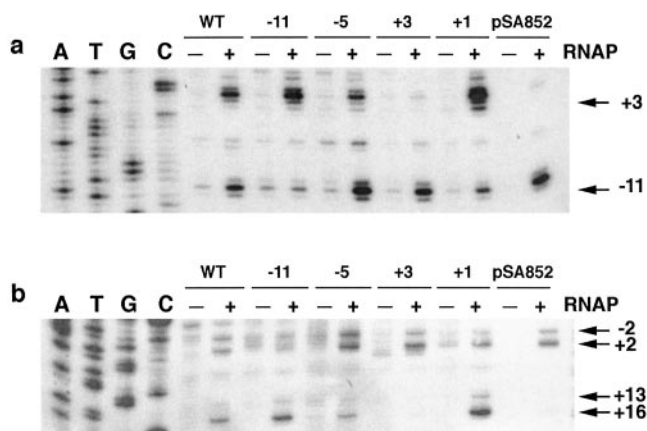


Fig. 5. Probing of thymine unpairing in DNA by KMnO_4 . (a) Nontemplate strand. (b) Template strand. WT, DNA containing no 2-AP. Numbers on the top indicate the positions of 2-AP in DNA. RNAP, RNA polymerase. pSA852 is supercoiled plasmid DNA. The positions of each KMnO_4 -reacted unpaired thymine are indicated by the numbers on the right with arrows. DNA sequencing ladders for A, T, G, and C were generated by using the DNA sequencing kit (BRL). The DNA template used in sequencing was the 358-bp PCR DNA fragment containing 2-AP at -11 position; the results show that the complementary base for 2-AP after amplification is T.

at $+3$ also showed the same pleiotropic effect in *P3*; it not only abolished unpairing of thymine opposite to 2-AP at $+3$ but also failed to show unpairing at positions $+13$ and $+16$. These results clearly showed that when 2-AP resides at the corresponding -11 position of *P1* and *P3*, none of the thymines unpaired.

Despite the fact that adenine and 2-AP are isomers that are indistinguishable in several physical and biochemical properties

as discussed, 2-AP behaved differently from adenine when it replaced the latter at the invariant -11 position in the *P1* and *P3* promoters of gal. Although DNA containing 2-AP at -11 did bind RNA polymerase, it prevented unpairing and subsequent transcription, perhaps because the amino group at C6 of the purine makes a critical contact with the σ subunit, which is essential for unpairing. It has been suggested that the adenine at -11 interacts with the σ^{70} subunit by becoming stacked between two aromatic amino acid residues, Trp-433 and Tyr-430 (12, 38, 39). It is noteworthy that 2-AP at several other positions of the two *gal* promoters did not show any mutational effect. Unlike at other positions, 2-AP at position -11 at *P1* not only prevented unpairing at that position but also prevented base unpairing at other positions located downstream, thus showing a polar effect. These results suggest that unpairing of the invariant A·T base pair -11 is a forerunner event and acts as the “master” of unpairing of bases. It has been suggested that DNA “melting” by RNA polymerase starts with a nucleation in the -12 to -4 region and extends up to $+2$ position downstream (18). Our finding of -11 A·T being the master position further pinpoints the actual start of the process of the nucleation. There may be a hidden rule in the pattern of base unpairing in a promoter guided by overall DNA sequence of the region; nevertheless, base unpairing shows polarity starting at a master position and moving downstream. Although no information is available about how RNA polymerase opens base pairs at a promoter, our demonstration of formation of heparin-resistant complex without base unpairing indicates that isomerization of RNA polymerase precedes that of DNA.

We thank Dale Lewis for the gift of several plasmids, Angela Fox for technical assistance, and Susan Garges and Jay Gralla for critical suggestions.

- McClure, W. R. (1985) *Annu. Rev. Biochem.* **54**, 171–204.
- Record, M. T. J., Reznikoff, W. S., Craig, M. L., McQuade, K. L. & Schlaw, P. J. (1996) in *Escherichia Coli and Salmonella: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), pp. 792–820.
- Gross, C. A., Lonetto, M. & Losick, R. (1992) in *Transcriptional Regulation*, eds. Yamamoto, K. & McKnight, S. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 129–176.
- Buc, H. & McClure, W. R. (1985) *Biochemistry* **24**, 2712–2722.
- Craig, M. L., Suh, W.-C. & Record, M. T. J. (1995) *Biochemistry* **34**, 15624–15632.
- deHaseth, P. L. & Helmann, J. D. (1995) *Mol. Microbiol.* **16**, 817–824.
- Roe, J. H., Burgess, R. R. & Record, M. T. (1984) *J. Mol. Biol.* **176**, 495–521.
- Roe, J. H., Burgess, R. R. & Record, M. T. (1985) *J. Mol. Biol.* **184**, 441–453.
- Severinov, K. & Darst, S. A. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 13481–13486.
- Zaychikov, E., Denissova, L., Meier, T., Gotte, M. & Heumann, H. (1997) *J. Biol. Chem.* **272**, 2259–2267.
- Roberts, C. W. & Roberts, J. W. (1996) *Cell* **86**, 495–501.
- Helmann, J. D. & Chamberlin, M. J. (1988) *Annu. Rev. Biochem.* **57**, 839–872.
- Juang, Y.-L. & Helmann, J. D. (1994) *J. Mol. Biol.* **235**, 1470–1488.
- Burns, H. & Minchin, S. (1994) *Nucleic Acids Res.* **22**, 3840–3845.
- Chen, Y.-F. & Helmann, J. D. (1997) *J. Mol. Biol.* **267**, 47–59.
- Helmann, J. D. & deHaseth, P. L. (1999) *Biochemistry* **38**, 5959–5967.
- Juang, Y.-L. & Helmann, J. D. (1995) *Biochemistry* **34**, 8465–8473.
- Li, X.-Y. & McClure, W. R. (1998) *J. Biol. Chem.* **273**, 23558–23566.
- Meier, T., Schickor, P., Wedel, A., Cellai, L. & Heumann, H. (1995) *Nucleic Acids Res.* **23**, 988–994.
- Tomsic, M., Tsujikawa, L., Panaghie, G., Wang, Y., Azok, J. & deHaseth, P. L. (2001) *J. Biol. Chem.* **274**, 31891–31896.
- Choy, H. E. & Adhya, S. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 472–476.
- McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N. & Scholtissek, S. (1987) *Biochemistry* **26**, 7238–7245.
- Nordlund, T. M., Andersson, S., Nilsson, L., Rigler, R., Graslund, A. & McLaughlin, L. W. (1989) *Biochemistry* **28**, 9095–9103.
- Sowers, L. C., Fazakerley, G. V., Eritja, R., Kaplan, B. E. & Goodman, M. F. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5434–5438.
- Wu, P., Nordlund, T. M., Gildia, B. & McLaughlin, L. W. (1990) *Biochemistry* **29**, 6508–6514.
- Xu, D., Evans, K. O. & Nordlund, T. M. (1994) *Biochemistry* **16**, 9592–9595.
- Sur, R., Debnath, D., Mukhopadhyaya, J. & Parrack, P. (2001) *Eur. J. Biochem.* **268**, 2344–2350.
- Adhya, S. & Miller, W. (1979) *Nature (London)* **279**, 492–494.
- Ponnambalam, S., Chan, B. & Busby, S. (1988) *Mol. Microbiol.* **2**, 165–172.
- Busby, S. (1982) *J. Mol. Biol.* **154**, 211–227.
- Musso, R. E., Di Lauro, R., Adhya, S. & de Crombrughe, B. (1977) *Cell* **12**, 847–854.
- Hawley, D. K. & McClure, W. R. (1983) *Nucleic Acids Res.* **11**, 2237–2255.
- Ponnambalam, S., Spassky, A. & Busby, S. (1987) *FEBS Lett.* **219**, 189–196.
- Sasse-Dwight, S. & Gralla, J. D. (1989) *J. Biol. Chem.* **264**, 8074–8081.
- Sasse-Dwight, S. & Gralla, J. D. (1991) *Methods Enzymol.* **208**, 146–168.
- Burns, H. D., Ishihama, A. & Minchin, S. D. (1999) *Nucleic Acids Res.* **27**, 2051–2056.
- Chan, B., Minchin, S. & Busby, S. (1990) *FEBS Lett.* **267**, 46–50.
- Fenton, M. S., Lee, S. J. & Gralla, J. D. (2000) *EMBO J.* **19**, 1130–1137.
- Panaghie, G., Aiyar, S. E., Bobb, K. L., Hayward, R. S. & de Haseth, P. L. (2000) *J. Mol. Biol.* **299**, 1217–1230.