An Unsubstituted C2 Hydrogen of Adenine Is Critical and Sufficient at the −11 Position of a Promoter to Signal Base Pair Deformation*

Received for publication, February 2, 2004, and in revised form, February 26, 2004

Hee Jung Lee, Heon Man Lim,†‡§ and Sankar Adhya†§

From the Department of Biology, College of Natural Sciences, Chungnam National University, Taejon 305-764, Republic of Korea and the Laboratory of Molecular Biology, NCI, National Institutes of Health, Bethesda, Maryland 20892-4264

The conserved AT base pair at the −11 position of the promoters in Escherichia coli is very sensitive to substitutions. In vitro transcription with the galP1 promoter having a natural or unnatural base in either strand at position −11 showed that only a purine base with no side group at C2 in the nontemplate strand is transcriptionally potent; neither a purine with an amino group at C2 nor a pyrimidine support transcription. The amino group at C6 in the omnipresent adenine at −11 does not play any role in promoting transcription. The nature of the base, complementary or noncomplementary, at −11 in the template strand also does not influence transcription. We propose that the adenine, by becoming extra-helical, interacts with an amino acid(s) of the 2.3–2.4 region of σ for which an unsubstituted C2 hydrogen is critical.

Transcription initiation executed by a bacterial RNA polymerase holoenzyme (Eo) starts with its binding to a promoter DNA sequence (formation of a closed complex) and is followed by a multistep process leading to the formation of an open complex. The open complex is competent to initiate polymerization of ribonucleotides according to the DNA sequences of the template strand (1–4). Both DNA (−10 region of the promoter) and the multisubunit enzyme go through major conformational changes in becoming an open complex (4–9). Although the exact structural changes that they undergo during open complex formation are unknown, recent crystallographic and fluorescence resonance energy transfer studies of various bacterial Eo and their parts provided some clue to the structure of the open complex (10–15). It is believed that in a fully mature open complex, about 15 base pairs in DNA (from −12 to +3 relative to the transcription start site) become single-stranded by the so-called “melting” process, which we will refer to as DNA deformation (16–18). Several amino acid residues in the 2.3–2.4 region of the σ subunit were inferred to make consequential contacts with bases in the −10 region of the promoter: Phe427, Tyr430, and Trp433 in the formation of the open complex, Lys414, Lys415, Tyr430, Phe427, and Tyr434 in binding to single-stranded DNA following DNA deformation, and Tyr430 in the stabilization of the open complex (4, 16, 18–27). Although these amino acid residues were shown to strategically lie on one face of the σ subunit to be able to perform the two functions in open complex formation, RNA polymerase binding and DNA deformation (10, 11), the details of the interactions from structural viewpoints remain unknown. Studies of transcription initiation with specifically designed DNA templates suggested that DNA deformation nucleates at around position −10–11 (9, 25–28). The omnipresent AT base pair at the position −11 was specifically suggested to play a “master” role in signaling DNA deformation (29). We investigated the properties of the −11 base pair, which are critical in interacting with σ and in signaling base pair deformation.

EXPERIMENTAL PROCEDURES

In Vitro Transcription Assay—In vitro transcription reactions were performed in a 25-μl volume following the procedure of Choy and Adhya (30). The σ20–saturated RNA polymerase from Escherichia coli was purchased from Epicenter Technologies (Madison, WI). The reaction mixture (22.5 μl) contained 4 nm PCR-generated DNA template (2 nm for negatively supercoiled plasmid DNA), 20 nm RNA polymerase, 20 nm Tris acetate, pH 7.5, 10 mM magnesium acetate, 200 mM potassium glutamate, and 20 units of rRNasin (Promega, Madison, WI). The reactions were preincubated at 37 °C for 5 min and started by the addition of 2.5 μl of NTP (2 mM of ATP, GTP, and CTP, 0.2 mM UTP, and 2.5 μCi of [α-32P]UTP). They were incubated at 37 °C for 10 min and stopped by the addition of 25 μl of RNA loading buffer (95% deionized formamide, 5 mM EDTA, 0.025% bromphenol blue, and 0.025% xylene cyanol). Samples were boiled for 2 min and chilled on ice, and 2.5 μl of samples were electrophoresed on 8% polyacrylamide-urea sequencing gels. The transcripts were quantified by using a PhosphorImager (Amersham Biosciences).

DNA Template Preparation—Synthetic oligonucleotides (105 nucleotides long) containing various nucleotides at −11 position of the gal P1 promoter were obtained from Genosys (The Woodlands, TX) and purified by PAGE. 358-bp-long DNA templates for transcription containing adenine, 2-aminopurine, 2,6-diaminopurine, or purine (experiments 1, 8, 12, and 16 in Table I) were prepared by PCR using the pSA852 plasmid DNA as template as described previously (29). The position and the size of the primers are indicated in Fig. 2. The PCR products were treated with the PCR purification kit (Qiagen, Valencia, CA) and electroeluted in agarose gel. The desired product embedded in the gel was cut out and electroeluted. The resulting DNA fragments were further cleaned by G-50 spin column (Amersham Biosciences). DNA templates with mismatched base pairs (experiments 5, 7, 9, 11, 13, and 15) were prepared as follows. The complementary strands of DNA templates containing mismatched bases at −11 position were synthesized individually using an asymmetric PCR method that employs only a single primer, and the relevant pairs of amplified single-stranded DNA were duplexed by mixing the corresponding PCR reaction products. After hybridization, the products were visualized by agarose gel electrophoresis. The hybridized DNA band was cut out and purified as described above for double-stranded PCR products. All PCR and hybridization-generated DNA templates contained two mutations (−12G to T and −22T to G), which largely inactivate the P2 promoter (31). A site-directed mutagenesis method (32) was used to change the pSA852 plasmid into one carrying guanine, thymine, or cytosine instead of adenine at the −11 position (experiments 2–4 in Table I). The latter constructs were directly used for in vitro transcription as supercoiled DNA templates.

* This work was supported in part by a grant of the Korea Health 21 R & D project, Ministry of Health and Welfare, Republic of Korea (01-PJ10-PG6-01GM02-0002). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence may be addressed. Tel.: 84-42-821-6276; E-mail: hmlim@cu.ac.kr (for H. M. Lim) or Tel.: 301-496-2495; Fax: 301-480-7687; E-mail: sadhys@helix.nih.gov (for S. Adhya).
RESULTS

The fact that the presence of 2-aminopurine at −11 of the nontemplate strand in the gal promoters made the promoter defective for transcription suggested that the position of the substitution group in the purine ring at −11 of the nontemplate strand is the source of the inhibitory effect.

Substitution of the Nontemplate Strand Adenine C2H at −11 Is Deleterious—The structural difference between 2-aminopurine and adenine (6-aminopurine) is in the position of the amino group (Fig. 1). We addressed the question whether the presence of an amino group at the C2 position or the absence of an amino group at the C6 position or both in 2-aminopurine prevented transcription. We performed in vitro transcription assays with E. coli RNA polymerase on DNA templates that contained various unnatural bases at −11 position of the P1 promoter of gal. The DNA templates were supercoiled plasmid or fragments generated by PCR or hybridization of single-stranded PCR products (Fig. 2). The bases used were purine, adenine, 2-aminopurine, 2,6-diaminopurine, guanine, as well as the two pyrimidines, thymine, and cytosine. The P2 promoter of gal in the templates used in these studies was inactivated unless mentioned (31). The sample results of the transcription assays are shown in Fig. 3, and the entire data set are summarized in Table I. As shown in Fig. 3, the wild type template (with adenine at −11) produced two transcripts, a 135-nt-long P1 RNA and a 121-nt-long P3 RNA. The P3 RNA was not made when a supercoiled plasmid DNA was the template (29, 33). The DNA template having 2-aminopurine at −11 of the P1 promoter did not yield transcripts from P1 (lane 3).

2,6-diaminopurine with amino groups at both C2 and C6 positions of the purine ring was present at −11 of P1, the DNA template did not yield P1 transcripts either (lane 4). These results showed that the presence of an amino group at C6 does not ensure the ability of a purine base at −11 to allow transcription from P1. However, we observed that a DNA template having a purine base (no amino group in either position) at −11 of P1 made RNA from the promoter as much as the wild type template (lane 5). DNA sequencing confirmed that the template strand base complementary to the unnatural bases (2-aminopurine, 2,6-diaminopurine, and purine) in the template strand was thymine (Table I), showing that the substitutions at the −11 position of the nontemplate strand were solely respons-

1 The abbreviation used is: nt, nucleotide.
sible for the observed behavior of the templates.

The above results demonstrated that if an amino group is placed at C2 of the purine ring at –11 in the nontemplate strand of the P1 promoter, the substitution becomes detrimental to transcription initiation. The presence of an amino group at C6 of the purine base is not important. This conclusion was further supported by the use of guanine at the –11 position of the nontemplate strand (lane 6). Guanine has an amino group at C2 and oxygen at C6 of the purine ring (Fig. 1). The template with guanine in the nontemplate and cytosine in template strand was inactive for P1 transcription (Fig. 3, lane 6). The –11 position did not accept a pyrimidine base (cytosine or thymine) either, suggesting a strict requirement of an adenine or a purine at –11 (Table I). Our observation of the essential nature of a purine base with an unmodified C2 hydrogen at –11 for promoter activity is consistent with a previous report that depurination of adenine, while stimulating open complex formation when at position –12, –10, or +1, inhibited the process when at position –11 (28).

The Nature of the Template Strand Base at –11 Is Inconsequential—To test whether the base in the template strand at the –11 position plays any role, several templates were made in which the base at –11 nontemplate strand, adenine, 2-aminopurine, purine, guanine, cytosine, or thymine, was paired against a non-Watson-Crick complement. The templates were tested for P1 transcription. The results showed that the transcriptional ability of the templates does not depend upon the nature of the base in the nontemplate strand (Table I). A mismatch did not change the behavior of the base in the nontemplate strand; adenine and purine remained transcriptionally active, and 2-aminopurine, guanine, and cytosine showed a mutational effect irrespective of the nature of the base in the other strand. Thus, having a heteroduplex pairing at the most critical position of the –10 region of the promoter is not a problem in the promoter function. (Thymine paired with a noncomplementary base was not tested.) The PCR-generated DNA templates containing different bases at –11 of P1 did not significantly change the P3 RNA synthesis. The innocuous effect of base substitutions at some positions of the template strand as opposed to that in the nontemplate strand was reported before (16).

**DISCUSSION**

We established that an unsubstituted C2 hydrogen of a purine base is essential and sufficient at the nontemplate strand of the critical –11 position for the promoter to be transcriptionally competent. If an amino acid side chain(s) in the σ 2.3–2.4 region makes a direct contact with the –11 adenine, then either the interaction involves the C2 hydrogen (a likely hydrophobic linkage) or the substitution of the C2 hydrogen causes a steric problem for the interaction. In B-DNA structure, the C2 hydrogen of an adenine is available in the minor groove (34). However, the line-up of base-edge interaction potential (with a protein) through the minor groove in an A:T base pair is precisely the same as that in a T:A base pair: H bond acceptor-H atom-H bond acceptor (Fig. 4). Since the consensus sequence (TATAAT) of the –10 element has only A:T or T:A base pair, it is difficult to imagine how an amino acid side chain in the σ protein can discriminate between the –11 and other positions in the neighborhood while establishing a contact using the minor groove. One likelihood is that the –11 adenine becomes extrahelical thus providing specificity to σ. Indeed it was suggested previously that –11 adenine interacts with σ by becoming stacked between two aromatic amino acid residues, Tyr430 and Trp433, in the 2.3 region of the subunit (9). However, we believe that a base intercalation between two adjacent planar aromatic side chains, e.g. Trp433 and Trp434 (a separation of about 3.5 Å) or between Tyr430 and Trp433 (a separation of 5.1 Å) in the protein α-helix may not be energetically feasible. An alternative idea is that the extrahelical base contacts amino acid side chain(s) through base edge. As opposed to the sandwich structure, an extrahelical base interaction with a single aromatic residue, prevalent in many reported structures of DNA-protein complexes with flipped out bases, is favored by deHaseth and colleagues (35–41). Whether the –11 adenine interaction with σ involves base intercalation between aromatic rings or base edge and a single amino acid side chain, our results suggest that an unsubstituted C2 hydrogen in the slipped out –11 adenine is critical for the contact. Although the adenine residue at –11 is the first base to be distorted to initiate the process of isomerization (29), the subsequent deformation of the base pairs in the –10 region is not a sequential process beginning at –11 (42). We postulate that the –11 adenine interaction (intercalation or base edge-amino acid side chain interaction) induces major changes in the structure of RNA polymerase instructing other amino acid residues in σ to make individual contacts with the remaining base pairs of the –10 region and deform them with a pattern that is detected by the DNA sequence of the promoter. Incidentally, about 5% of E. coli promoters do not have an adenine at position –11 (43). We presume that these promoters use another position in the –10 region to initiate open complex formation.

**REFERENCES**

1. Chamberlin, M. J. (1974) *Annu. Rev. Biochem.* 43, 721–775
2. McClure, W. R. (1985) *Annu. Rev. Biochem.* 54, 171–204
3. Record, M. T. J., Reznikoff, W. S., Craig, M. L., McQuade, K. L., and Schlax,

---

**TABLE I**

| Source | P | P | P | P | P | H | H | H | H | P | P | H | H | H | P |
|-------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Non-template | A | G | C | T | A | G | C | 2AP | 2AP | 2AP | 2AP | Pu | Pu | Pu | Pu |
| Template | T | C | G | A | A | A | T | A | G | C | T | A | G | C | T |
| Transcription | Y | N | Y | N | N | N | N | Y | Y | Y | N | Y | Y | Y | N |
Role of C2 Hydrogen of Adenine at −11 in Promoter Opening

P. J. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., ed) pp. 792–820, American Society for Microbiology, Washington, D. C.
4. deHaseth, P. L., and Helmann, J. D. (1995) J. Mol. Biol. 16, 817–824
5. deHaseth, P. L., Zupancic, M. L., and Record, M. T. J. (1998) J. Bacteriol. 180, 3019–3025
6. Craig, M. L., Tsodikov, O. V., McQuade, K. L., Schlax, P. E. J., Capp, M. W., Saecker, R. M., and Record, M. T. J. (1998) J. Mol. Biol. 283, 741–756
7. Borowiec, J. A., and Gralla, J. D. (1986) Biochemistry 25, 5051–5057
8. Callaci, S., Heyduk, E., and Heyduk, T. (1999) Mol. Cell 3, 229–238
9. Fenton, M. S., Lee, S. J., and Gralla, J. D. (2000) EMBO J. 19, 1130–1137
10. Malhtra, A., Severinoa, E., and Darst, S. A. (1996) Cell 87, 127–136
11. Campbell, E. A., Masuda, S., Sun, J. L., Muzzin, O., Olson, C. A., Wang, S., and Darst, S. A. (2002) Cell 108, 795–807
12. Vassylev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassyleva, M. N., Borukhev, S., and Yokoyama, S. (2002) Nature 417, 712–719
13. Murakami, K. S., Masuda, S., and Darst, S. A. (2002) Science 296, 1280–1284
14. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002) Science 296, 1285–1290
15. Meckler, V., Korkhjonova, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidza, A. N., Niu, W., Ehrlich, Y. W., Levy, R., and Ehrlich, R. H. (2002) Cell 108, 599–614
16. Roberts, C. W., and Roberts, J. W. (1996) Cell 86, 495–501
17. Helmann, J. D., and Chamberlin, M. J. (1988) Annu. Rev. Biochem. 57, 839–872
18. Juang, Y. L., and Helmann, J. D. (1996) J. Mol. Biol. 235, 1470–1488
19. Gardella, T., Moyle, H., and Stassiek, M. M. (1988) J. Mol. Biol. 206, 579–590
20. Kenney, T. J., York, K., Youngman, P., and Moran, C. P. J. (1988) Proc. Natl. Acad. Sci. U. S. A. 86, 139–1413
21. Siegele, D. A., Hu, J. C., Walter, W. A., and Gross, C. A. (1989) J. Mol. Biol. 206, 591–603
22. Tatti, K. M., Jones, C. H., and Moran, C. P. J. (1991) J. Bacteriol. 173, 7828–7833
23. Waldburger, C., Gardella, T., Wong, R., and Susskind, M. M. (1990) J. Mol. Biol. 215, 267–276
24. Zuber, P., Healy, J., Carter, H. L., 3rd, Cutting, S., Moran, C. P. J., and Losick, R. (1989) J. Mol. Biol. 206, 605–614
25. Jones, C. H., and Moran, C. P. J. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 1958–1962
26. Panaghiou, G., Aiyar, S. E., Bobb, K. L., Hayward, R. S., and de Haseth, P. L. (2000) J. Mol. Biol. 299, 1217–1230
27. Helmann, J. D., and deHaseth, P. L. (1999) Biochemistry 38, 5959–5967
28. Li, X. Y., and McClure, W. R. (1998) J. Biol. Chem. 273, 23558–23566
29. Lim, H. M., Lee, H. J., Roy, S., and Adhya, S. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 14849–14852
30. Choy, H. E., and Adhya, S. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 472–476
31. Bushby, S., Alba, H., and de Crombrugghe, B. (1992) J. Mol. Biol. 154, 211–227
32. Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V., and Kopka, M. L. (1982) Science 216, 475–485
33. Tomsic, M., Tsujikawa, L., Panaghiou, G., Wang, Y., Azok, J., and deHaseth, P. L. (2001) J. Biol. Chem. 276, 31891–31896
34. Helmann, J. D., and Chamberlin, M. J. (1988) Annu. Rev. Biochem. 57, 839–872
35. Helmann, J. D., and Chamberlin, M. J. (1988) Annu. Rev. Biochem. 57, 839–872
36. Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000) EMBO J. 19, 758–766
37. Davies, D. R., Goryshin, I. Y., Reznikoff, W. S., and Raymert, I. (2002) Science 299, 77–85
38. Keifer, J. R., Mao, C., Braman, J. C., and Beese, L. S. (1998) Nature 391, 384–387
39. Pues, H., Bleimling, N., Holz, B., Wolcke, J., and Weinhold, E. (1999) Biochemistry 38, 1426–1434
40. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) Nature 384, 87–92
41. Xiao, G., Tordova, M., Jagadeesh, J., Drohat, A. C., Stivers, J. T., and Gilliland, G. L. (1999) Proteins 35, 15–24
42. Roy, S., Lim, H. M., Lui, M., and Adhya, S. (2004) EMBO J. 23, 869–875
43. Hawley, D. K., and McClure, W. R. (1983) Nucleic Acids Res. 11, 2237–2255
