Hierarchies of Base Pair Preferences in the P22 *ant* Promoter

HENRY MOYLE, CAREY WALDBURGER, AND MIRIAM M. SUSSKIND*

Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-1340

Received 31 August 1990/Accepted 12 January 1991

Oligonucleotide-directed mutagenesis was used to complete a collection of mutations in the -35 and -10 hexamers of the *ant* promoter of *Salmonella* phage P22. The effects of all 36 single-base-pair substitutions on promoter strength *in vivo* were measured in strains carrying the mutant promoters fused to an *ant-lacZ* gene on a single-copy prophage. The results of these assays show that certain consensus base pairs are more important than others; in general, the least-critical positions are among the most poorly conserved. Some mutations within the hexamers have smaller effects on promoter strength than certain mutations outside the hexamers in this and other promoters. Several different patterns of base pair preferences are observed. These hierarchies of base pair preferences correlate well (but not perfectly) with the hierarchies defined by the frequency distribution of base pairs at each position among wild-type promoters. The hierarchies observed in the *ant* promoter also agree well with most of the available information on base pair preferences in other promoters.

The first step of transcription initiation is the association of RNA polymerase with a promoter sequence on the DNA. In eubacteria, promoter specificity is determined by a family of protein subunits called sigma (σ) factors, which associate with the core enzyme ($\alpha_2\beta\beta'$) to form the holoenzyme. The major sigma factor of *Escherichia coli* and *Salmonella typhimurium* is σ^{70} , a 70-kDa protein required for recognition of most promoters. Other classes of promoters are recognized by holoenzymes containing minor sigma factors (16).

Comparisons of numerous σ^{70} -dependent promoters controlling chromosomal, plasmid, transposon, and phage genes in *E. coli* and *S. typhimurium* revealed two regions of conserved DNA sequence located approximately 10 and 35 bp upstream of the start site of transcription. Each region has a 6-bp consensus sequence, the -10 and -35 hexamers 5'-TATAAT-3' and 5'-TTGACA-3', respectively. Weakly conserved positions outside the hexamers have varied considerably in different compilations (Fig. 1) (14, 15, 32). The hexamers are separated by a spacer region with a preferred length of 17 bp. Recent genetic studies strongly suggest that both regions of the promoter are directly contacted by sigma factors (8, 11, 23, 34, 41, 47).

Evidence that the hexamers are important determinants of promoter strength has been obtained by the analysis of promoter mutations, most of which affect base pairs within the hexamers (15). In most cases, mutations that decrease promoter activity are changes away from consensus, while mutations that increase promoter activity are changes toward consensus. Various exceptions to this rule are discussed below. Kinetic studies of several promoters show that mutations in the hexamers affect the rate parameters for early steps in the interaction between the promoter and RNA polymerase, the formation of an initial, "closed" complex and/or isomerization to a transcriptionally active, "open" complex (18, 39; for a review see reference 29).

The hexamers are not the sole determinants of promoter activity, since down or up mutations outside the hexamers have been reported. Most notably, T at -15 and G at -14 (numbering positions as in Fig. 1) each improve the activity

of at least three different promoters (5, 9, 13, 22, 43; for reviews, see references 15 and 30). Other single or multisite substitutions in the spacer (1, 10, 15, 30, 41a) or in the regions upstream of the -35 hexamer (10, 27, 41a, 44) or downstream of the -10 hexamer (10, 20, 21, 24, 25) also affect activity, at least in certain promoters. However, single-base-pair substitutions in these regions rarely have effects as dramatic as those at certain positions of the hexamers. The length of the spacer is also important, presumably for proper juxtaposition of the hexamers (2, 10, 13; for a review, see reference 15).

To elucidate the determinants of promoter activity, we have been characterizing mutations in P_{ant} , a strong near-consensus promoter of phage P22 (Fig. 1; for a review, see reference 38). The P_{ant} operon includes two genes: *arc* (antirepressor control) and *ant* (antirepressor) (see Fig. 2, top). P_{ant} is negatively regulated by the products of the flanking genes *mnt* and *arc*: Mnt protein represses P_{ant} during lysogeny, and Arc protein turns down P_{ant} early after infection. The P_{ant} region also contains P_{mnt} , an overlapping, divergent promoter for transcription of the *mnt* repressor gene (Fig. 1). Because the genetics of the *ant* operon provided selections for promoter down and up mutations, we previously obtained a large number of single-base-pair changes in P_{ant} (13, 44), most of which cluster in the hexamers. To define the base pair preferences at each of these positions, we have synthesized the remaining changes and assayed the effects of all 36 mutations on promoter strength *in vivo*.

MATERIALS AND METHODS

Media. Green and LB agar (37) and M9CAA (35) media have been described previously. Supplements were used at the following concentrations: ampicillin (Amp), 50 μ g/ml; spectinomycin (Spc), 50 μ g/ml; kanamycin (Kan), 10 μ g/ml; and 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal), 30 μ g/ml.

Plasmids. pMS420 (Amp^r; 11), pMS421 (Spc^r; 12), and pMS580 (Amp^r; 12) have been described. The Arc-producing plasmid pMS414 (Spc^r Kan^r) was constructed by inserting the *EcoRI* E fragment of P22 *Kn9* (45) in the *EcoRI* site of

* Corresponding author.

FIG. 1. Three versions of the consensus promoter sequence (14, 15, 32) are aligned with the sequence of wild-type *P_{ant}*. Weakly conserved positions are indicated by lowercase letters. The -10 and -35 hexamers are bracketed. Arrows indicate the start point and direction of transcription.

pGB2, a pSC101-based *Spc^r* plasmid (7). The insert carries the wild-type *P_{ant}-arc* region and a substitution (*Kn9*) that replaces *mnt* with a *Kan^r* gene. To construct pMS813 (Amp^r *Kan^r*), which carries the -35T→A mutation, the *Ava*II fragment carrying *P_{ant}* from P22 *Kn6 P_{ant}-35T→A arc-am* phage (see below) was filled in and inserted in the filled-in *Hind*III-*Eco*RV backbone of pZ152 (46). The insert extends from an *Ava*II site within the *Kn6* substitution to an *Ava*II site at +2 with respect to the *P_{ant}* start point.

DNA sequencing. Sequencing of single-stranded plasmid and double-stranded plasmid or phage DNA was done by the dideoxy method by using Sequenase enzyme (U.S. Biochemical Corp.) and a *P_{ant}*-specific primer.

Construction of new *P_{ant}* mutations. Oligomutagenesis was done by using either (i) the method of Kunkel (26) as described previously (11) or (ii) the Amersham Corp. Mutagenesis System. The target of mutagenesis was single-stranded pMS420, which carries an *ant-lacZ* gene fusion under control of wild-type *P_{ant}*. After transformation of *E.*

coli, strains carrying mutant plasmids were usually identified by screening for lighter-blue colony color on LB-ampicillin-X-Gal plates. The *E. coli* recipient was usually TG1 (Amersham Corp.) but in some cases was TG1 carrying the compatible plasmid pMS414. This plasmid produces Arc protein, a repressor of *P_{ant}*, which improved the stability of pMS420 derivatives carrying very mild down mutations. The *P_{ant}* region of each mutant plasmid was sequenced to confirm that it carried the correct base pair change.

Construction of P22 *Kn6 P_{ant} lacZ* phages. The effects of *P_{ant}* mutations on promoter strength in vivo were assayed by using a set of P22 derivatives carrying an *ant-lacZ* gene fusion under control of *P_{ant}*. Figure 2 illustrates the structure of these phages, which has been described in detail previously (12). These “*lac* phages” carry (i) *Kn6*, a substitution that replaces *mnt* with a *Kan^r* gene, (ii) *amH1605*, an amber mutation in *arc*, (iii) *lac2*, a substitution that replaces a segment of the *ant* gene with *E. coli lac* sequences, creating an *ant-lacZ* gene, and (iv) *9-delTB*, a deletion that removes the genes for tail protein (gene 9) and antigen conversion (*a1*). These phages were constructed by two sequential plasmid-phage crosses as previously described (12). Briefly, the *P_{ant}* mutations were crossed from the mutant plasmids (see above) onto P22, generating a set of P22 *Kn6 P_{ant} arc-am* phages. The resulting phages were crossed with pMS580, which donates the *lac2* substitution and *9-delTB* deletion. The *P_{ant}* region of each of the resulting *lac* phages was sequenced to confirm that it carried the correct base pair change.

Construction of *P_{ant} ant-lacZ* lysogens. Strains that are certain to carry one copy of the *P_{ant} ant-lacZ* fusions were constructed by crossing the *Kn6 P_{ant} arc-am lac2 9-delTB* constellation onto the defective prophage in *S. typhimurium* MS2909 (Fig. 2). This strain (11) has a deletion that removes the host *proAB* locus and the leftmost third of the P22 prophage, including the *immC* region (for a review, see reference 38). Since this prophage lacks the *c2* repressor gene, it does not confer immunity to superinfecting P22. The

FIG. 2. Scheme for crossing *P_{ant}-ant-lacZ* fusions onto a partially deleted prophage. Open boxes represent genes of interest; shaded boxes represent deletions. Only the *mnt-9* region of the recipient prophage is drawn roughly to scale. The indicated crossovers replace the wild-type *mnt-a1* region of the recipient prophage with the corresponding region of the superinfecting *lac* phage. The resulting nonimmune (*c2* mutant) lysogens must carry only one copy of the *lac* fusion.

FIG. 3. Construction of P_{ant} double down mutations. P22 $Kn6$ $arc-am$ phage carrying a mutation at -8 were crossed with plasmid pMS813, which carries a mutation at -35 . The indicated crossovers generate phage recombinants that form large plaques under the plating conditions used to counterselect the parental phage (see text). Numbers denote distances available for crossovers (number of phosphates between nonhomologous bases).

crosses were done by infecting MS2909 with a high multiplicity of $Kn6$ $P_{ant} \downarrow$ $arc-am$ $lac2$ $9-delTB$ phage and plating the infected cultures on green agar-kanamycin plates. Kan^r colonies were screened for immunity to P22 ant mutant phage. Immune Kan^r lysogens must carry a $c2^+$ prophage; they presumably carry the entire genome of the $Kn6$ lac phage, integrated into the original $immC\Delta$ prophage. Kan^r lysogens that are fully sensitive to P22 ant mutant phage must carry a single $immC\Delta$ prophage and must have the $Kn6$ (Kan^r) and $9-delTB$ markers from the lac phage. (The $9-delTB$ deletion removes the P22 antigen conversion gene al ; $A1^+$ lysogens are somewhat resistant to superinfecting phage because a modification of the host receptor inhibits adsorption of virions.) Figure 2 illustrates how these nonimmune lysogens arise: two crossovers, one to the left of $Kn6$ and one to the right of $9-delTB$, replace the relevant segment of the resident prophage with that of the superinfecting phage. As expected, all of these nonimmune lysogens carry the $lac2$ substitution. The presence of the $arc-am$ mutation was confirmed by crossing the arc region of the defective $arc-am$ prophage onto a P22 derivative carrying $RU411$ (44), a deletion of P_{ant} and most of the arc gene. Recombinant phages from these crosses were then tested for presence of the $arc-am$ mutation by genetic tests (44) or by restriction analysis of phage DNA (the $arc-am$ allele destroys a $Fnu4HI$ site; 33). MS2909 and its derivatives carry the $lacI^q$ plasmid pMS421 for reasons unrelated to this study.

Construction of double mutants. The strategy used previously to combine the $-35T \rightarrow A$ and $-8T \rightarrow A$ mutations (12) was used to combine $-35T \rightarrow A$ with $-8T \rightarrow G$ or $-8T \rightarrow C$. P22 $Kn6$ $P_{ant} -8T \rightarrow G$ $arc-am$ and $Kn6$ $P_{ant} -8T \rightarrow C$ $arc-am$ phages were crossed with pMS813, a plasmid carrying the $-35T \rightarrow A$ mutation (Fig. 3). Recombinant phages that acquired the $-35T \rightarrow A$ mutation from the plasmid were selected by plating on an Su^- host at $37^\circ C$. Neither the parental phage nor P_{ant}^+ recombinants form plaques under these conditions because they have the Arc^- lethal phenotype; this phenotype is suppressed by the $P_{ant} \downarrow -35T \rightarrow A$ mutation (44). Sequence analysis of the P_{ant} region of two candidates from each cross identified recombinants carrying both $-35T \rightarrow A$ and either $-8T \rightarrow G$ or $-8T \rightarrow C$. These phages were used to construct lac phages and lysogens as described above.

β -Galactosidase assays and statistical analysis. Cells were grown to mid-log phase in M9CAA medium at $37^\circ C$. Cultures were harvested, lysates were prepared, and β -galactosidase activities were assayed as described previously (11). One lysate of each culture was assayed in duplicate or quadruplicate. The two or four measurements of specific activity

TABLE 1. Relative activities of mutant promoters

P_{ant} genotype of lac fusion	Relative activity ^a	Hierarchy based on activity ^b	Base distribution ^c
Wild type	(100)		
$-35T \rightarrow C$	2.1 ± 0.2	$T > G > A \approx C$	$T_{78} G_{10} C_9 A_3$
$-35T \rightarrow A$	2.3 ± 0.8		
$-35T \rightarrow G$	3.6 ± 0.7		
$-34T \rightarrow C$	4.4 ± 0.7	$T > G \approx A > C$	$T_{82} A_{10} G_5 C_3$
$-34T \rightarrow A$	14.7 ± 2.1		
$-34T \rightarrow G$	16.9 ± 1.0		
$-33G \rightarrow A$	1.3 ± 0.3	$G > T > C > A$	$G_{68} T_{15} C_{14} A_3$
$-33G \rightarrow C$	2.0 ± 0.1		
$-33G \rightarrow T$	12.5 ± 0.8		
$-32A \rightarrow G$	2.9 ± 0.1	$A > C > T > G$	$A_{58} T_{20} C_{13} G_{10}$
$-32A \rightarrow T$	7.1 ± 1.0		
$-32A \rightarrow C$	11.6 ± 1.7		
$-31C \rightarrow G$	3.5 ± 0.5	$C > A > T > G$	$C_{52} A_{32} T_{10} G_7$
$-31C \rightarrow T$	10.8 ± 1.9		
$-31C \rightarrow A$	17.9 ± 2.6		
$-30A \rightarrow C$	40.0 ± 5.5	$A > T > G \approx C$	$A_{54} T_{24} G_{17} C_5$
$-30A \rightarrow G$	52.2 ± 8.2		
$-30A \rightarrow T$	78.1 ± 7.0		
$-12T \rightarrow A$	1.6 ± 0.1	$T > C > G > A$	$T_{82} C_8 G_7 A_3$
$-12T \rightarrow G$	2.8 ± 0.3		
$-12T \rightarrow C$	5.0 ± 0.5		
$-11A \rightarrow C$	0.7 ± 0.1	$A > G > C; A > T$	$A_{89} T_7 C_3 G_1$
$-11A \rightarrow T$	1.1 ± 0.5		
$-11A \rightarrow G$	1.3 ± 0.1		
$-10T \rightarrow G$	29.7 ± 6.7	$T > C \approx A > G$	$T_{52} A_{26} G_{12} C_{10}$
$-10T \rightarrow A$	44.7 ± 4.2		
$-10T \rightarrow C$	52.3 ± 6.4		
$-9A \rightarrow G$	36.5 ± 4.2	$A > C > T \approx G$	$A_{59} G_{15} T_{14} C_{12}$
$-9A \rightarrow T$	37.8 ± 1.2		
$-9A \rightarrow C$	57.9 ± 4.1		
$-7T \rightarrow G$	4.2 ± 0.1	$T > A > C \approx G$	$T_{89} C_5 A_3 G_2$
$-7T \rightarrow C$	5.2 ± 0.7		
$-7T \rightarrow A$	26.2 ± 2.5		

^a P22 P_{ant} - lac monolysogens were assayed for β -galactosidase activity as described in Materials and Methods. Percent activity is given relative to that of the wild-type promoter, which is 100% and equal to 852 Barrick units.

^b Statistical tests used to compare activities are described in Materials and Methods. Consensus bases are in boldface type.

^c Subscripts are base frequencies, expressed as a percentage, among 263 wild-type promoters; data are from Harley and Reynolds (14).

were averaged and converted to relative specific activity by dividing by the average specific activity of the P_{ant}^+ controls on the same microtiter plate. The relative specific activity of each mutant promoter was measured in this way on three to nine occasions. The mean and adjusted-sample standard deviations of the three to nine values of relative specific activity are given in Tables 1 and 2.

Pairwise Student t tests were used to assess the statistical significance of differences in activity between promoters carrying mutations at the same position. In Tables 1 and 2, the confidence interval for differences considered significant was 98%. Even if the confidence limit were reduced to 95%, no additional differences would be deemed significant.

TABLE 2. Effects of mutations at position -8

Base at -8 (top strand)	Relative activity		Hierarchy ^a	Base distribution ^b
	With -35T→A ^c	Alone ^d		
A (consensus)	(100)	83.4 ± 9.1	A > C > T ≈ G	A ₄₉ C ₂₁ T ₁₉ G ₁₁
G	18.8 ± 2.7	58.4 ± 9.9		
T (wild type)	22.8 ± 2.7	(100)		
C	72.4 ± 5.5	93.1 ± 9.5		

^a Hierarchy in the presence of -35T→A (see text for rationale). Consensus base is in boldface type.

^b See Table 1, footnote c.

^c Promoters carrying -35T→A and the indicated genotype at -8 were assayed. Percent activity is relative to that of the promoter carrying -8T→A, which is 100% and equal to 90 Barrick units.

^d Promoters with the indicated genotype at -8 and no other change in *P_{ant}* were assayed. Percent activity is relative to that of the wild-type promoter, which is 100% and equal to 852 Barrick units.

RESULTS AND DISCUSSION

Twenty-one single-base-pair changes in the hexamers of *P_{ant}* were previously isolated, 20 by genetic selection (13, 44) and one by oligomutagenesis (11). We have now completed this set by constructing the remaining 15 mutations by oligomutagenesis. The relative strengths of these mutant promoters were assayed in vivo by fusing them to an *ant-lacZ* gene and measuring β -galactosidase levels produced by strains carrying the fusions on single-copy prophages (see Materials and Methods). The results of these assays are given in Tables 1 and 2 and are illustrated in Fig. 4.

Position -8 is considered separately in Table 2, since the analysis of changes at this position is conceptually complicated. Previous work showed that -8T→A, a change toward consensus, increases *P_{ant}* activity in vivo when combined with any 1 of 10 different down mutations (12). However, in the otherwise wild-type promoter, the -8T→A up mutation causes a slight decrease in promoter activity. Similarly, several other strong promoters become slightly less active in vivo when changed to make both hexamers identical with consensus, even though the final change toward consensus is different for each promoter (19, 21, 28). Such consensus promoters are defective in a late step of transcription initiation (21, 24), suggesting that the same strong contacts with

the hexamers that expedite formation of open complexes can also inhibit promoter clearance (6, 21, 36).

To determine the hierarchy of base pair preferences at position -8, the effects of mutations at this position were examined in the presence of the down mutation -35T→A, as well as in the context of the otherwise wild-type promoter (Table 2). As shown previously, the consensus A · T base pair is preferable to the nonconsensus, wild-type T · A in the presence, but not in the absence, of the down mutation at -35. Presumably, the results obtained in the presence of the -35T→A down mutation reflect the true base pair preferences at -8, whereas the pattern seen in the otherwise wild-type promoter is skewed by the promoter clearance defect caused by a perfect match with consensus.

Previously, the combined results of studies of various promoters established that each consensus base pair in the hexamers is preferable to at least two (four positions) or all three (eight positions) nonconsensus base pairs (9, 10, 27, 34, 43; for a review, see reference 15). Our data confirm that, within a single promoter, any change away from consensus in the hexamers significantly decreases promoter strength (with the caveat about perfect consensus discussed above). Some base pairs are more critical than others: any change at -11, -12, or -31 through -35 is more detrimental than any change at -8, -9, -10, or -30. These less critical positions

FIG. 4. Activities of mutant promoters plotted above the consensus sequence. Data are from Tables 1 and 2. Wild-type *P_{ant}* differs from consensus only at -8 (marked with an asterisk). The effects of mutations at -8 were analyzed in the presence of a mutation at -35.

TABLE 3. Patterns of base-pair preferences

Preferred base pair ^a	Second-best base pair	Positions
A · T	C · G	-32, -9, -8
T · A	G · C	-35
A · T	T · A	-30
T · A	A · T	-7
A · T	G · C	-12
G · C	T · A	-33
C · G	A · T	-31

^a The nine hierarchies in which the second-best base pair is unambiguous are summarized. Pairs of related patterns are grouped together (e.g., the pattern for -32, -9, and -8 is the inverse of that for -35).

are also four of the five (15) or six (14) most poorly conserved positions in the hexamers. Some of these mutations have smaller effects than some single-base-pair substitutions outside the hexamers in P_{ant} or other promoters (see introduction).

At 11 positions, the relative activity of all three mutants is either $\leq 18\%$ or $\geq 19\%$. The exception is -7, where two changes are severe and one is mild, even though this "invariant T" is one of the two most highly conserved positions in the promoter (the other is -11). In the *tyrT* promoter, T→A at the corresponding position is a severe down mutation, reducing promoter activity to 1.6% of wild type (3, 4). In P_{ant} , the mild phenotype of the -7T→A mutation might reflect its effect on an overlapping -10 hexamer (44) or on the -35 hexamer of P_{mnt} (see below).

Taking into account the statistical significance of our data (see Materials and Methods), we derived the hierarchies of base pair preferences shown in Tables 1 and 2. These hierarchies are generally in good agreement with the frequency distributions of bases at each position among 263

wild-type promoters compiled by Harley and Reynolds (14; and see last column of Tables 1 and 2). Three of four complete hierarchies (positions -33, -31, and -12, but not -32) are congruent with the rank orders defined by frequency distributions. In addition, half of the remaining partial hierarchies (positions -35, -34, -30, and -8, but not -11, -10, -9, and -7) are consistent with the hierarchies based on frequency. (For example, at -35 the hierarchy based on promoter strength, $T > G > A \approx C$, is consistent with the hierarchy based on frequency, $T > G > C > A$.) Thus, sequence conservation among wild-type promoters, which accurately predicted the best base pair at each position, does not always accurately rank the three alternatives. Most of the discrepancies occur at positions where the frequencies of two or more nonconsensus bases differ by only a few percentage points (positions -11, -10, -9, and -7). The most striking discrepancy is at position -32, where C is better than T for promoter strength, but T is substantially more frequent than C among wild-type promoters.

The hierarchies at certain positions are similar to each other, but there are clearly several different patterns of base pair preferences (Table 3). Although the consensus sequence is remarkably AT rich, our results show that in many cases T · A and A · T are not the best substitutes for each other. At seven positions where the consensus base pair is A · T or T · A and there is a clear-cut second-best alternative, it is the inverse transversion in only two cases (Table 3). Even at positions -9 and -8, which are within the region that is unwound in the open complex (-10 to +3), the strong base pair C · G is preferable to the weak base pair T · A.

In other promoters, some comparisons among nonconsensus base pairs are available, usually because two different changes away from consensus at the same position were obtained. In addition, a few down mutations change one

TABLE 4. Hierarchies in the hexamers of other promoters

Position in P_{ant} ^a	Hierarchy in P_{ant}	Hierarchy in other promoter ^b	Basis for hierarchy ^c	Reference
-35	T > G > A ≈ C	*T > C > G (<i>E. coli lac</i> ± cAMP) T > C ≈ G (<i>Tn10 tetA</i>)	2 c→nc downs ^d 2 c→nc downs ^e	34 9
-34	T > G ≈ A > C	T > A > C (<i>E. coli lac</i> ± cAMP) T > A ≈ C (<i>E. coli rrnB P1</i>)	2 c→nc downs ^d 2 c→nc downs ^e	34 10
-33	G > T > C > A	T > C ≈ A (<i>E. coli lac</i>) G > T > C (<i>E. coli recA</i>)	2 nc→nc downs ^d 2 c→nc downs ^d	11 42
-32	A > C > T > G	*A > G > C (<i>E. coli lac</i> + cAMP) *A = G > C (<i>E. coli lac</i> - cAMP) *G > A (<i>E. coli araB</i>) A > T ≈ G (<i>Tn10 tetA</i>)	2 c→nc downs ^d 2 c→nc downs ^d 1 c→nc up ^f 2 c→nc downs ^e	34 34 17 9
-31	C > A > T > G	C > T > G (<i>E. coli lac</i> ± cAMP)	2 c→nc downs ^d	34
-12	T > C > G > A	T > C > G (<i>E. coli lac</i> ± cAMP)	2 c→nc downs ^d	34
-11	A > G > C	A > G ≈ C (<i>E. coli rrnB P1</i>) A > T > G (<i>E. coli lac</i> ± cAMP)	2 c→nc downs ^e 2 c→nc downs ^d	10 34
-10	T > C ≈ A > G	*G > A (<i>E. coli argC</i>) *T > C ≈ G > A (λP_R)	1 nc→nc down ^e 3 c→nc downs ^g	31 34
-7	T > A > C ≈ G	T > G > C (<i>E. coli lac</i> ± cAMP) T > C ≈ A (<i>E. coli tyrT</i>)	2 c→nc downs ^d 2 c→nc downs ^e	34 3, 4

^a Corresponding positions in other promoters may be numbered differently.

^b Hierarchies marked by an asterisk differ from those in P_{ant} . Other promoter is shown in parentheses.

^c c, Consensus; nc, nonconsensus.

^d Promoter activities in vivo were measured by β -galactosidase assays, and statistical significance was tested as for P_{ant} . Primary data from Siegle et al. (34) on *lac* promoter mutations in a Δcya strain in the presence and absence of cAMP were provided by D. Siegle. The P_{lac} constructs from Gardella et al. (11) are missing the catabolite gene activator protein-cAMP binding site. Additional data for *recA* were from J. M. Weisemann (41b).

^e Promoter activities in vivo were determined by assaying β -galactosidase, other enzyme activities, or mRNA levels. Differences less than 1.2-fold were considered insignificant. The *argCp* mutation reduces gene expression two- to fourfold.

^f Mutant promoter is more active than the wild-type promoter by genetic test in vivo in the absence of AraC.

^g Hierarchy based on $K_B \times k_p$, the kinetic parameters for initial binding and isomerization, measured in vitro (13a).

nonconsensus base pair to another nonconsensus base pair. Most of these hierarchies in other promoters agree with those seen in P_{ant} (Table 4). Siegle et al. (33a, 34) examined the effects of numerous mutations in the *E. coli lac* promoter in a Δcya strain in the presence and absence of cAMP, which is required for activation of P_{lac} by the catabolite gene activator protein-cAMP complex. At two of eight positions, the hierarchies in P_{lac} differ from those in P_{ant} . Whereas C is better than G at -32 of P_{ant} , G is better than C at the corresponding position of P_{lac} in both the presence and absence of cAMP. Interestingly, the consensus A is better than G at this position of P_{lac} in the presence of cAMP, but A and G are equivalent in the absence of cAMP. Another remarkable finding is that A→G at this position of the *araBAD* promoter is an up mutation (17), even though it is a change away from consensus and the same change is a down mutation in several other promoters (15). At two other positions, the ranking of nonconsensus bases varies in different promoters. At -35, G is better than C in P_{ant} , whereas C is better than G in P_{lac} . At -10, A is better than C in P_{ant} , whereas C is better than A in the *E. coli argC* (31) and λP_R (13a) promoters.

These variations are interesting, since it is conceivable that the preferences at one position might depend on the identity of base pairs at other positions. However, another obvious kind of context effect that applies to many bacterial and phage promoters is the arrangement of overlapping promoters (29). The -35 and -10 hexamers of P_{ant} overlap those of the divergent promoter P_{mnt} (Fig. 1): -30 of P_{ant} is -7 of P_{mnt} and vice versa. Normally, P_{ant} is by far the dominant promoter; P_{mnt} activity is detectable only if P_{ant} is repressed by Mnt protein, which binds to an operator located from +1 to +17 with respect to P_{ant} (40). Nevertheless, the hierarchies in P_{ant} , particularly those at positions -30 and -7, may reflect the effects of mutations on P_{mnt} as well as P_{ant} . Many of the other promoters listed in Table 4 have overlapping promoters, promoterlike sequences, or activator-binding sites.

Even when the hierarchies in P_{ant} are the same as those in other promoters, the magnitudes of the effects presented here are not expected to agree precisely with the quantitative effects of the same changes in other promoters (see above). Even within P_{ant} , the effect of one mutation can vary considerably when its context is changed by the presence of another mutation (Table 2) (12). The magnitudes of the effects of P_{ant} mutations, but not the hierarchies, also appear to vary somewhat in different host strains for reasons that are currently obscure.

ACKNOWLEDGMENTS

We thank Myron Goodman for help with statistics, Gary Gussin for critical reading of the manuscript, and Debby Siegle, Carol Gross, and George Weinstock for help in compiling Table 4.

This work was supported by Public Health Services grant GM36811 from the National Institutes of Health to M.M.S. C.W. was the recipient of a Graduate Merit Fellowship from the University of Southern California.

REFERENCES

1. Auble, D. T., T. L. Allen, and P. L. DeHaseth. 1986. Promoter recognition by *Escherichia coli* RNA polymerase. Effects of substitutions in the spacer DNA separating the -10 and -35 regions. *J. Biol. Chem.* **261**:11202-11206.
2. Ayers, D. G., D. T. Auble, and P. L. deHaseth. 1989. Promoter recognition by *Escherichia coli* RNA polymerase: role of the spacer DNA in functional complex formation. *J. Mol. Biol.* **207**:749-756.
3. Berman, M. L., and J. Beckwith. 1979. Use of gene fusions to isolate promoter mutants in the transfer RNA gene *tyrT* of *Escherichia coli*. *J. Mol. Biol.* **130**:303-315.
4. Berman, M. L., and A. Landy. 1979. Promoter mutations in the transfer RNA gene *tyrT* of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **76**:4303-4307.
5. Bingham, A. H. A., S. Ponnambalam, B. Chan, and S. Busby. 1986. Mutations that reduce expression from the P2 promoter of the *Escherichia coli* galactose operon. *Gene* **41**:67-74.
6. Carpousis, A. J., J. E. Stephano, and J. Gralla. 1982. 5' nucleotide heterogeneity and altered initiation of transcription at mutant *lac* promoters. *J. Mol. Biol.* **157**:619-633.
7. Churchward, G., D. Belin, and Y. Nagamine. 1984. A pSC101-derived plasmid which shows no sequence homology to other commonly used cloning vectors. *Gene* **31**:165-171.
8. Daniels, D., P. Zuber, and R. Losick. 1990. Two amino acids in an RNA polymerase σ factor involved in the recognition of adjacent base pairs in the -10 region of a cognate promoter. *Proc. Natl. Acad. Sci. USA* **87**:8075-8079.
9. Daniels, D. W., and K. P. Bertrand. 1985. Promoter mutations affecting divergent transcription in the Tn10 tetracycline resistance determinant. *J. Mol. Biol.* **184**:599-610.
10. Gaal, T., J. Barkel, R. R. Dickson, H. A. DeBoer, P. L. DeHaseth, H. Alavi, and R. L. Gourse. 1989. Saturation mutagenesis of an *Escherichia coli* rRNA promoter and initial characterization of promoter variants. *J. Bacteriol.* **171**:4852-4861.
11. Gardella, T., H. Moyle, and M. M. Susskind. 1989. A mutant *E. coli* σ^{70} subunit of RNA polymerase with altered promoter specificity. *J. Mol. Biol.* **206**:579-590.
12. Graña, D., T. Gardella, and M. M. Susskind. 1988. The effects of mutations in the *ant* promoter of phage P22 depend on context. *Genetics* **120**:319-327.
13. Graña, D., P. Youderian, and M. M. Susskind. 1985. Mutations that improve the *ant* promoter of Salmonella phage P22. *Genetics* **110**:1-16.
- 13a. Gussin, G. (University of Iowa). Personal communication.
14. Harley, C. B., and R. P. Reynolds. 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15**:2343-2361.
15. Hawley, D. K., and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **8**:2237-2255.
16. Helmann, J. D., and M. J. Chamberlin. 1988. Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* **57**:839-872.
17. Horwitz, A. H., C. Morandi, and G. Wilcox. 1980. Deoxyribonucleic acid sequence of *araBAD* promoter mutants of *Escherichia coli*. *J. Bacteriol.* **142**:659-667.
18. Hwang, J.-J., S. Brown, and G. N. Gussin. 1988. Characterization of a doubly mutant derivative of the λP_{RM} promoter: effects of mutations on activation of P_{RM} . *J. Mol. Biol.* **200**:695-708.
19. Inouye, S., and M. Inouye. 1985. Up promoter mutations in the *lpp* gene of *Escherichia coli*. *Nucleic Acids Res.* **13**:3101-3110.
20. Jacques, J.-P., and M. M. Susskind. 1990. Pseudo-templated transcription by *Escherichia coli* RNA polymerase at a mutant promoter. *Genes Dev.* **4**:1801-1810.
21. Kammerer, W., U. Deuschle, R. Gentz, and H. Bujard. 1986. Functional dissection of *Escherichia coli* promoters: information in the transcribed region is involved in late steps of the overall process. *EMBO J.* **5**:2995-3000.
22. Keilty, S., and M. Rosenberg. 1987. Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *J. Biol. Chem.* **262**:6389-6395.
23. Kenney, T. J., K. York, P. Youngman, and C. P. Moran, Jr. 1989. Genetic evidence that RNA polymerase associated with σ^A uses a sporulation specific promoter in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* **86**:9109-9113.
24. Knaus, R., and H. Bujard. 1988. P_L of coliphage lambda: an alternative solution for an efficient promoter. *EMBO J.* **7**:2919-2923.
25. Krummel, B., and M. J. Chamberlin. 1989. RNA chain initiation by *E. coli* RNA polymerase: structural transitions of the enzyme

- in early ternary complexes. *Biochemistry* **28**:7829–7842.
26. Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* **82**:488–492.
 27. Kunkel, T. A., and P. S. Alexander. 1986. The base substitution fidelity of eucaryotic DNA polymerases: mispairing frequencies, site preferences, insertion preferences, and base substitution by dislocation. *J. Biol. Chem.* **261**:160–166.
 28. Mandecki, W., R. A. Goldman, B. S. Powell, and M. H. Caruthers. 1985. *lac* up-promoter mutants with increased homology to the consensus promoter sequence. *J. Bacteriol.* **164**:1353–1355.
 29. McClure, W. R. 1985. Mechanisms and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.* **54**:171–204.
 30. Munson, L. M., W. Mandecki, M. H. Caruthers, and W. S. Reznikoff. 1984. Oligonucleotide mutagenesis of the *lacP_{UV5}* promoter. *Nucleic Acids Res.* **12**:4011–4017.
 31. Piette, J., R. Cunin, A. Boyen, D. Charlier, M. Crabeel, F. Van Vliet, N. Glansdorff, C. Squires, and C. L. Squires. 1982. The regulatory region of the divergent *argECBH* operon in *Escherichia coli* K-12. *Nucleic Acids Res.* **10**:8031–8048.
 32. Rosenberg, M., and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* **13**:319–353.
 33. Sauer, R. T., W. Krovatin, J. DeAnda, P. Youderian, and M. M. Susskind. 1983. Primary structure of the *imm1* immunity region of bacteriophage P22. *J. Mol. Biol.* **168**:699–713.
 - 33a. Siegele, D. (Harvard Medical School). Personal communication.
 34. Siegele, D., J. C. Hu, W. A. Walter, and C. A. Gross. 1989. Altered promoter recognition by mutant forms of the σ^{70} subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **206**:591–604.
 35. Smith, H. O., and M. Levine. 1964. Two sequential repressions of DNA synthesis in the establishment of lysogeny by phage P22 and its mutants. *Proc. Natl. Acad. Sci. USA* **52**:356–363.
 36. Straney, D. C., and D. M. Crothers. 1987. A stressed intermediate in the formation of stably initiated RNA chains at the *E. coli lacUV5* promoter. *J. Mol. Biol.* **193**:267–278.
 37. Susskind, M. M., A. Wright, and D. Botstein. 1971. Superinfection exclusion by P22 prophage in lysogens of *Salmonella typhimurium*. II. Genetic evidence for two exclusion systems. *Virology* **45**:638–652.
 38. Susskind, M. M., and P. Youderian. 1983. Bacteriophage P22 antirepressor and its control, p. 347–363. In R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), *Lambda II*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 39. Szoke, P. A., T. L. Allen, and P. L. deHaseth. 1987. Promoter recognition by *Escherichia coli* RNA polymerase: effects of base substitutions in the –10 and –35 regions. *Biochemistry* **26**:6188–6194.
 40. Vershon, A. K., S. Liao, W. R. McClure, and R. T. Sauer. 1987. Bacteriophage P22 Mnt repressor DNA binding and effects on transcription *in vitro*. *J. Mol. Biol.* **195**:311–322.
 41. Waldburger, C., T. Gardella, R. Wong, and M. M. Susskind. 1990. Changes in conserved region 2 of *Escherichia coli* σ^{70} affecting promoter recognition. *J. Mol. Biol.* **215**:267–276.
 - 41a. Weinstock, G. (University of Texas Medical School). Personal communication.
 - 41b. Weisemann, J. M. 1988. Ph.D. thesis. University of Texas Medical School, Houston.
 42. Weisemann, J. M., and G. M. Weinstock. 1985. Direct selection of mutations reducing transcription or translation of the *recA* gene of *Escherichia coli* with a *recA-lacZ* protein fusion. *J. Bacteriol.* **163**:748–755.
 43. Wulff, D. L., M. Mahoney, A. Shatzman, and M. Rosenberg. 1984. Mutational analysis of a regulatory region in bacteriophage λ that has overlapping signals for the initiation of transcription and translation. *Proc. Natl. Acad. Sci. USA* **81**:555–559.
 44. Youderian, P., S. Bouvier, and M. M. Susskind. 1982. Sequence determinants of promoter activity. *Cell* **30**:843–853.
 45. Youderian, P., A. Vershon, S. Bouvier, R. T. Sauer, and M. M. Susskind. 1983. Changing the DNA-binding specificity of a repressor. *Cell* **35**:777–783.
 46. Zagursky, R. J., and M. L. Berman. 1984. Cloning vectors that yield high levels of single-stranded DNA for rapid DNA sequencing. *Gene* **27**:183–191.
 47. Zuber, P., J. Healy, H. L. Carter, III, S. Cutting, C. P. Moran, Jr., and R. Losick. 1989. Mutation changing the specificity of an RNA polymerase sigma factor. *J. Mol. Biol.* **206**:605–614.