Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 555–559, January 1984
Genetics

Mutational analysis of a regulatory region in bacteriophage λ that has overlapping signals for the initiation of transcription and translation

(P_{RE} promoter/λ *c*II gene/positive regulation/initiation of protein synthesis)

DANIEL L. WULFF*, MICHAEL MAHONEY*, ALLAN SHATZMAN†‡, AND MARTIN ROSENBERG†‡

*Department of Biological Sciences, State University of New York, Albany, NY 12222; and †Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20205

Communicated by Rollin Hotchkiss, October 7, 1983

ABSTRACT **The positively regulated P_{RE} promoter of phage λ structurally overlaps with the ribosome-binding and NH_2-terminal coding region of the regulatory protein (*c*II) that activates P_{RE} transcription. We have isolated and characterized 27 different point mutations that occur within the 36-base-pair overlapping region. A comparison of genetic crossover data with nucleotide separations as determined by DNA sequence analysis reveals that recombination frequencies are greatly depressed at very short distances. Moreover, recombination frequency is critically dependent upon the precise nucleotide sequence of the crossover region for distances of five nucleotides or less. The mutations define precise positions and sequences that are important to (*i*) P_{RE} promoter function, (*ii*) translation of the *c*II gene, and (*iii*) *c*II gene function. Mutational changes that affect the function of one element in this region concomitantly define phenotypically silent alterations in the other two elements. Mutations deficient in promoter function (P^-_{RE} or *cy*) are clustered in two regions that lie ≈10 and ≈35 nucleotides before the initial base of P_{RE} mRNA, analogous to mutations in other promoters. P^-_{RE} mutations in the −10 region alter bases that are conserved in prokaryotic promoters, but P^-_{RE} mutations in the −35 region do not affect bases that are normally conserved in other promoters. Several mutations deficient in *c*II gene activity affect the initiation of *c*II protein synthesis, including an A→G change four bases outside the *c*II coding region, and AUG→GUG, AUG→ACG, and AUG→AUA mutations in the initiation codon. In the region of overlap between the P_{RE} promoter and the NH_2-terminal region of the *c*II gene, most amino acid substitutions in the *c*II protein do not result in a loss of *c*II function, indicating that this region of the gene does not contain essential information for *c*II function. We suggest that the overlap itself is an evolutionarily conserved structure and that it somehow coordinates the bidirectional transcriptional and translational events that occur in this region.**

The P_{RE} promoter of bacteriophage λ controls the expression of the phage repressor (*c*I) protein during infection of a sensitive cell (Fig. 1; refs. 1–4). This promoter is positively regulated by the phage *c*II protein (2, 3). The *c*II protein is both necessary and sufficient for transcriptional activation of P_{RE} (5, 6).

One striking feature of the *c*II-P_{RE} activation system is that the promoter region overlaps structurally with the DNA sequence encoding both the translation initiation and the NH_2-terminal coding region of the *c*II gene (Fig. 2; refs. 5, 11, and 14). These overlapping elements are oriented in opposite directions. Transcription of the P_{RE} promoter proceeds leftward on the template towards the repressor (*c*I) gene, whereas the *c*II gene itself is transcribed from the P_R promoter as part of the major rightward transcription unit (Fig. 1).

The *c*II-P_{RE} activation system is well suited for both genetic and biochemical analysis. Mutations affecting either P_{RE} promoter function or the *c*II gene product can be readily isolated and characterized (11, 15). Of particular interest are those mutations that map in the region of structural overlap between P_{RE} and the *c*II gene. Within this region mutations can be identified that affect *c*II expression or function. These mutations also lie within the P_{RE} promoter region yet are functionally P^+_{RE}. Similarly, P_{RE} promoter mutations can be identified within this region that also alter the sequence of the NH_2-terminal-coding and ribosome-binding region of the *c*II gene yet retain functional *c*II activity. Thus, the system not only allows isolation and characterization of P_{RE} and *c*II mutations but concomitantly allows examination of phenotypically silent base-pair changes within the P_{RE} region and the NH_2-terminal end of the *c*II gene.

Previously, 8 different P^-_{RE} mutations were isolated and characterized (11). In this study we identify 9 additional P^-_{RE} mutations as well as 10 *c*II$^-$ mutations that occur in the NH_2-terminal region of the *c*II gene. The position and sequence changes associated with these 27 different mutations are defined and discussed with regard to the effects they have on P_{RE} promoter function, translation of the *c*II gene, and *c*II activity.

MATERIALS AND METHODS

Bacterial and phage strains are described in ref. 11. Mutagenesis with either UV light or a *mutD* mutator strain of *Escherichia coli* (16) was done according to ref. 17. Genetic map distances were determined by four-factor crosses as described in ref. 15. DNA sequence analysis was carried out according to the methods described in ref. 18. The appropriate DNA restriction fragments from mutant phage DNAs were isolated, 5′-end-labeled with ^{32}P, and subjected to sequence analysis as described in ref. 8. In each case, the nucleotide sequence was determined for the entire y region and compared with that of λ$^+$.

RESULTS

Mutations that inactivate the P_{RE} promoter or the *c*II gene of λ have readily identifiable, clear plaque phenotypes (11, 19, 20). These mutations can be distinguished from other clear plaque mutations by genetic complementation and mapping procedures (15). We isolated >1,000 clear plaque mutants using either UV light or a *mutD* mutator strain of *E. coli* as mutagenic agent. Thirty-eight of these were found to lie in or near the P_{RE} promoter region by genetic mapping, and these were divided into P_{RE} and *c*II mutations by genetic complementation. The locations of these mutations were

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.

‡Present address: Smith Kline & French Research Laboratories, 709 Swedeland Road, Swedeland, PA 19406.

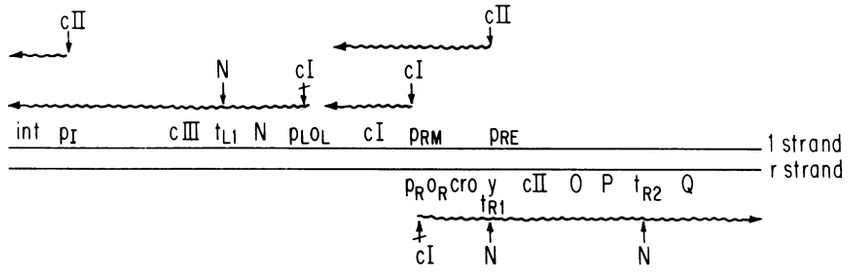

FIG. 1. Partial genetic map of bacteriophage λ. The wavy lines with arrowheads indicate the origin, direction, and extent of various transcripts. Vertical arrows indicate sites of action of the *N*, *c*I, and *c*II gene products. An arrow with a slash indicates that the gene product represses transcription. An arrow without a slash at the start of a wavy line indicates that the gene product activates a promoter. An arrow in the middle of a wavy line indicates that the gene product prevents transcription from terminating at a termination site. (See ref. 1 for a more extensive discussion.)

then precisely determined by genetic fine structure mapping. Our goal was to determine the number of unique mutational events represented in our collection so as to avoid sequence analysis of identical recurrences of the same mutation.

Gussin *et al.* showed that map distance in the P_{RE}-*c*II region is very nearly proportional to nucleotide distance in the range of 3–160 nucleotides (21). However, for very short separations, map distances are significantly depressed. Therefore, it is not enough to establish whether or not two different clear mutations recombine with each other, because any two mutations separated by <3 nucleotides show little or no recombination with each other. Mutations are most accurately mapped by crossing them with mutations that lie 3, 4, or 5 nucleotides away, for in this range map distance increases dramatically with increasing physical distance (11, 21). The mapping data for various *c*II and P_{RE} (*cy*) mutations are tabulated in Table 1.

DNA sequence changes were determined for representative mutations in the set of 38 mutant strains. In Fig. 3, we summarize the positions and sequence changes determined for all the mutations isolated in the P_{RE} region. All strains with single base changes were either *c*II$^-$ P^+_{RE} or *c*II$^+$ P^-_{RE} except λ*c*3073, whose *c*II$^-$ P^-_{RE} phenotype is due to two separate base changes.

DISCUSSION

Distinguishing Mutations by Recombination. Sequence determination of the *c*II and P_{RE} mutations allows us to analyze map distance as a function of physical distance in the range between 1 and 7 base pairs. The mapping data in Table 1 are tabulated as a function of nucleotide separation. The data confirm the conclusion of Gussin *et al.* that map distances are strongly depressed at very short distances (21). For example, we could detect no recombination (<0.01 map unit) for any pair of mutations at 2 nucleotides separation, although recombination frequencies at longer distances (up to 160 nucleotides separation) would have led us to predict 0.1 map unit for a pair of mutations at 2 nucleotides separation (21). Similarly, measured genetic distances were generally much lower than predicted for separations of 3, 4, and 5 nucleotides.

The second striking feature of the data is the variability shown at each distance. At nucleotide separations of 3 and 4, recombination frequencies vary enormously for different pairs of nucleotides. At a distance of 3 nucleotides, the range is <0.01–0.11 map unit (>10-fold difference) and at 4 nucleotides, the range is 0.01–0.18 map unit (about a 20-fold difference). At 5 nucleotides separation, the range is 0.05–0.26 map unit (about a 5-fold difference). This degree of variability is not seen at longer distances (21).

We conclude that recombination frequency is most critically dependent upon the precise nucleotide sequence of the crossover region for distances of 5 nucleotides or less. It is this high variability in recombination frequency at very short distances that has made it possible to distinguish different closely linked nucleotide changes solely on the basis of recombination frequencies. Indeed, each nucleotide change exhibits its own "signature," or characteristic recombination frequencies with nearby mutations, and this allows it to be distinguished from closely spaced mutations.

In some cases different nucleotide changes at the same nucleotide pair yield strikingly different recombination frequencies with nearby mutations. For example, *cy* 3001, a GC→AT change at position 38,380, is separated by 0.08 map unit from *cy*844, which lies 4 nucleotides to the left. However, *cy*3095, a GC→TA change that is also at position 38,380, is separated by only 0.01 map unit from *cy*844. As a second example, *c*II3109, an AT→TA change in position 38,375, exhibits lower recombination frequencies with some closely spaced mutations than does *c*II3114, an AT→GC change that is also at position 38,375. We emphasize that each mutation

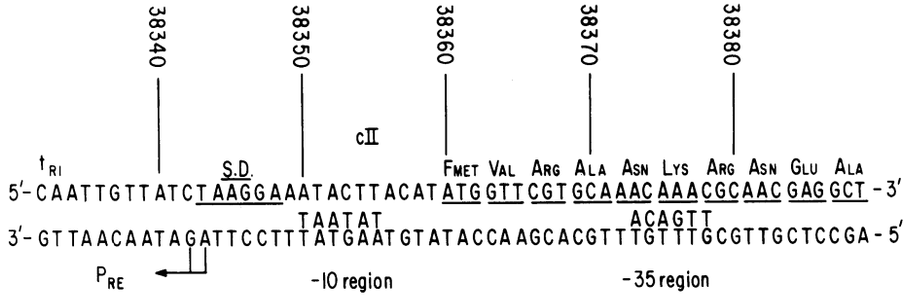

FIG. 2. Nucleotide sequence of the P_{RE} promoter and the NH_2-terminal region of the *c*II gene (7, 8). The numbering system is that of ref. 9 for λ DNA. The six-base consensus sequences for the −10 and −35 regions of prokaryotic promoters (10) are depicted at the appropriate places in the space between strands (1, 11, 12). The P_{RE} message is initiated at either of 2 nucleotides (5), as indicated. The t_{R1} termination site for rightward transcription from P_R (8) is at the extreme left end of the figure. The line labeled "S.D." indicates the Shine and Delgarno homology for the *c*II gene [i.e., the region where P_R mRNA is homologous with the 3′ end of 16S rRNA (13)].

Table 1. Mapping data for new *cy* and *cII* mutations in the P_{RE} region

Left marker	Right marker	Map unit*	Left marker	Right marker	Map unit*
		Two nucleotides apart			
cII3105	*cII3059*	<0.01	*cy3008*	*cy3072*	<0.01
cy3003	*cy3077*	<0.01	*cy3075*	*cy3001*	<0.01
cy3095	*cy3098*	<0.01	*cII3109*	*cy3075*	<0.01
cy3008	*cy3078*	<0.01	*cII3114*	*cy3075*	<0.01
		Three nucleotides apart			
cy3077	*cII3109*	<0.01	*cy3019*	*cII3086*	0.03
cII3109	*cy3107*	0.01	*cy3075*	*cy3072*	0.035
cII3114	*cy3107*	0.01	*cy3107*	*cy3078*	0.05
cy844	*cy3008*	0.01	*cII3067*	*cII3059*	0.06
cy3008	*cy3098*	0.01	*cy3098*	*cII3085*	0.07
cy3107	*cy3072*	0.02	*cy3075*	*cy3078*	0.11
cy3077	*cII3114*	0.025			
		Four nucleotides apart			
cy844	*cy3095*	0.01	*cy844*	*cy3001*	0.08
cy3048	*cy3073a*	0.02	*cy3107*	*cy3098*	0.09
cy3048	*cy3071*	0.03	*cII3086*	*cII3059*	0.10
cy3048	*cy2001*	0.05	*cy42*	*cII3114*	0.10
cy3019	*cII3067*	0.05	*cy3072*	*cII3085*	0.10
cy42	*cII3109*	0.05	*cy3077*	*cy844*	0.11
cII3109	*cy3008*	0.08	*cy3078*	*cII3085*	0.13
cII3114	*cy3008*	0.08	*cy3075*	*cy3098*	0.18
		Five nucleotides apart			
cy3019	*cII3105*	0.05	*cy844*	*cy3078*	0.13
cII3109	*cy3095*	0.09	*cy844*	*cy3072*	0.14
cy42	*cy844*	0.10	*cII3088*	*cII3067*	0.15
cy3003	*cII3109*	0.10	*cy3001*	*cII3085*	0.16
cII3114	*cy3095*	0.11	*cII3114*	*cy3001*	0.26
cII3109	*cy3001*	0.13			
		Six nucleotides apart			
cy42	*cy3075*	0.09	*cy844*	*cy3098*	0.25
cy3048	*cII3088*	0.17			
		Seven nucleotides apart			
cy42	*cy3107*	0.11	*cy2001*	*cII3067*	0.19
cy3048	*cy3019*	0.16	*cII3059*	*cy42*	0.21
cy3107	*cII3085*	0.17			
		Eight nucleotides apart			
cy3073b	*cy42*	0.18			

Data for *cII3625* are not shown. The data have been grouped according to actual nucleotide separation, as later determined by DNA sequence analysis. See Fig. 3.

*Crosses were of the form λimm^{434} *c1 O*am205 X λ*c2*, followed by determination of the percentage of imm^{434} c^+ O^+ recombinants among the total imm^{434} O^+ recombinants. Map distances were obtained from crosses in which imm^{434} c^+ O^+ recombinants may arise from single crossover events (15). One map unit equals 1% recombination. The data are reproducible to about 20%, except that larger fluctuations were obtained for distances of <0.1 because smaller numbers of imm^{434} c^+ O^+ recombinants were counted.

that was subjected to sequence analysis in this study was chosen for analysis because genetic mapping data showed it to be different from previously analyzed mutations.

Mutations at P_{RE}. Prokaryotic promoters generally exhibit two regions of strong sequence homology, positioned ≈10 base pairs and ≈35 base pairs before the transcription start site (i.e., the −10 and −35 regions, respectively) (10, 23, 24). The functional importance of these two regions of conserved sequence has been demonstrated both by mutational analyses and by chemical probe experiments designed to detect the DNA contacts made by RNA polymerase within the promoter region (10). The DNA sequence of the P_{RE} promoter exhibits little homology to these other promoters. The −10 region of P_{RE} is identical in only three positions to the hexamer consensus sequence for prokaryotic promoters (Fig. 2; ref. 5). However, two of these positions coincide with the most strongly conserved bases occurring in the −10 region hexamer: the second position A·T pair and the sixth position T·A pair. The −35 region of P_{RE} has even less similarity to other promoters. The six-base consensus sequence usually found here is absent at P_{RE}. Perhaps the lack of structural homology with other promoters is not surprising because P_{RE} function is totally dependent upon an activator protein. In the absence of *cII* protein, RNA polymerase does not interact with the P_{RE} region (6).

Mutations that inactivate promoter function (P^-_{RE} or *cy* mutations) distinctly cluster in the −10 and −35 regions of P_{RE}, analogous to mutations in other promoters. In contrast, most of the *cII* mutations that occur within the promoter region lie in the 17-base-pair spacer region between the −10 and −35 regions. These mutations do not affect promoter function and support the contention that this region of the promoter does not have a sequence-specific role but serves only to space correctly the −10 and −35 region sequences. A few of the *cII* mutations do occur in the −10 and −35 regions of P_{RE}. Presumably, these mutations indicate specific base changes in these regions that are acceptable to promoter function.

Promoter-down mutations usually alter the more highly conserved bases in the −10 and −35 region hexamer sequences (7). The nature of the −10 region P_{RE} mutations is consistent with this idea. Four of the 5 mutations affect the highly conserved second position A·T pair and the sixth position T·A pair found in P_{RE}. The second position A·T has been changed to all three of the other possible base pairs. All three substitutions are P^-_{RE}, which underscores the importance of the original base pair. The fifth mutation (*cy3019*) is a G·C→A·T change 2 base pairs upstream from the conserved hexamer. Although G·C, C·G, and T·A pairs are found at this position in other promoters, rarely does an A·T pair occur at this site (7). Apparently, an A·T pair at this position has a negative influence on promoter function. It is of interest that, unlike the other P_{RE} mutations, this P_{RE} mutation does not exhibit the typical clear plaque phenotype but rather forms lightly turbid plaques. This phenotype is indicative of a partial defect in function, consistent with its location in a more weakly conserved region of the promoter. G·C→A·T changes at the same position as the *cy3019* mutation have also been identified for the λ P_{RM} and the P22 P_{ant} promoters (25, 26).

One *cII* mutation, *cII3088*, is found in the −10 region. It is a T·A→C·G change 1 base pair upstream from the conserved hexamer. Although T·A is somewhat preferred to C·G at this site in prokaryotic promoters (10), this mutation is without apparent effect on P_{RE} function. No promoter mutations at this site have been reported in other systems.

In contrast to the −10 region mutations, the mutations that occur in the −35 region of P_{RE} do not affect bases that are normally conserved in the promoter. Moreover, these mutations span some 13 base pairs, a region far more extensive than the corresponding RNA polymerase contact region in other promoters. Most importantly, the region encompassed by these P_{RE} mutations shares a striking homology to the corresponding region of the other *cII*-dependent promoter in phage λ, P_I (27–29). Both regions have the sequence T-T-G-C-N_6-T-T-G-C, where N_6 stands for 6 intervening nucleotides between the T-T-G-C repeat sequences. Based on the strong −35 region homology exhibited by these two *cII*-dependent promoters, it has been proposed that this region is involved directly in activation of P_{RE} transcription (5, 11, 27,

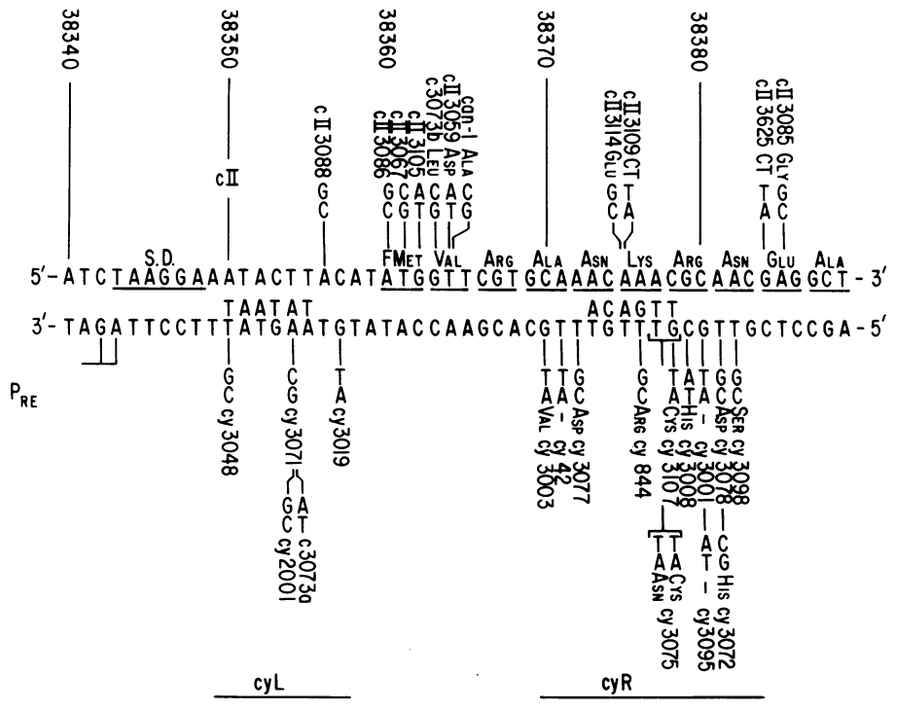

FIG. 3. Nucleotide sequence (7, 8) and sequence changes in the P_{RE} region and NH_2-terminal portion of the *c*II gene of bacteriophage λ. Sites with a *c*II prefix and *c*3073b are *c*II$^-$ mutations. The *can*-1 mutation of Jones and Herskowitz (22) results in a more stable *c*II protein. The cluster of *cy* mutations between nucleotides 38,357 and 38,350 (*cy*L mutations) are in the −10 region of the P_{RE} promoter, and the cluster of *cy* mutations between nucleotides 38,382 and 38,370 (*cy*R mutations) are in the −35 region of the P_{RE} promoter (11). Other conventions are as in Fig. 2. Eight *cy* changes (*cy*3048, *cy*2001, *cy*3019, *cy*3003, *cy*42, *cy*844, *cy*3008, and *cy*3001) and the *can*-1 change are from ref. 11. All other changes were determined in this study.

28, 30), and some evidence for this has been obtained (6). In their determination of DNA sequence elements specifically required for *c*II binding, Ho *et al.* (12) found that *c*II binding is greatly reduced by P_{RE}^- mutations in either of the two T-T-G-C repeat sequences, while *c*II binding is unaffected by P_{RE}^- mutations in the intervening 6 nucleotides.

Alterations in the *c*II Gene. The level of *c*II protein in an infected cell is crucial for determining whether the infection will proceed along the lytic or lysogenic pathway of phage development (see refs. 1 and 4 for reviews). Mutations in the NH_2-terminal region of the *c*II gene may alter these levels by affecting the translation efficiency of *c*II or the activity of the mature protein or both.

Several of the *c*II mutations affect the initiation of protein synthesis. One of these, the *c*II3088 mutation, lies outside the *c*II coding region, 4 base pairs to the left of the AUG initiation codon and 7 base pairs to the right of the region of complementarity with the 3′ end of 16S ribosomal RNA [the Shine and Delgarno, or S.D., sequence (13)] (Fig. 3). The translation efficiency of *c*II3088 mRNA is dramatically reduced in an *in vitro* RNA-dependent cell-free translation system (unpublished data). In an analysis of mRNA sequences near the initiation regions of many genes, Stormo *et al.* found that G is a rather infrequent base in the region between the S.D. sequence and the initiation codon (31). (Note that there are no Gs in this region of the wild-type *c*II gene and that *c*II3088 is an A→G change.) Stormo *et al.* note that, in addition to AUG, GUG and UUG codons also are recognized by initiator tRNAs (32, 33, 34). They propose that XUG codons in the initiation region may interfere with recognition of the proper initiation codon (31). This could explain the *c*II phenotype of the *c*II3088 mutation, because it creates a UUG codon 5 nucleotides to the right of the S.D. sequence. Alternatively, the decreased translation efficiency in λ*c*II3088 may be a consequence of the formation of mRNA secondary structure in the ribosome-binding region. Recent studies have shown that sequence alterations that affect mRNA structure can have dramatic effects on translation efficiency (35). The *c*II3088 mRNA sequence immediately preceding the AUG, 5′ U-U-G-C-A-U 3′, is complementary to the sequence 5′ G-U-G-C-A-A 3′ positioned 9 nucleotides downstream, and these two sequences could form a hydrogen-bonded complex. The *c*II3088 A→G change is central to the formation of this structure, which might interfere with *c*II translation initiation.

λ*c*II3088 lies in the midst of a cluster of five −10 region P_{RE} mutations, all of which lie between the S.D. sequence and the initiation codon. These mutations are all *c*II$^+$ by complementation and show normal translation efficiencies *in vitro* using an RNA-dependent S30 system (unpublished data). Two of these mutations are A→G changes, but neither creates XUG codons or new mRNA secondary structures. The occurrence of these mutations clearly indicates a high degree of flexibility in the sequence of this region of the ribosome-binding site.

Transition mutations have been found for each of the three positions of the initial AUG codon, creating GUG, ACG, and AUA codons. It is surprising that the mutation of GUG (*c*II3086) is *c*II$^-$ because GUG is an active initiation codon in other systems (see ref. 33). By using an *in vitro* RNA-dependent S30 system, little, if any, *c*II expression from this GUG codon could be detected (unpublished data). However, this A→G change also creates an AUG codon two nucleotides upstream from the normal start-site, and perhaps this out-of-phase initiation codon interferes with initiation at the proper site. It is interesting to note that the wild-type *lacI* gene has the same sequence, A-U-G-U-G, at its start-site (36), and, in this case, translation begins at the GUG codon. However, in *lacI* the spacing between the S.D. sequence and the out-of-phase AUG codon may be too small for translation to be initiated at the improper codon (31), whereas in λ*c*II3086 this spacing is sufficient to allow the improper codon to be

used. A pseudorevertant of λ*c*II3086 has been found in which the A-U-G-U-G in *c*II3086 is changed to A-C-G-U-G (unpublished data), an observation that supports the idea that the translation defect in λ*c*II3086 is related to the out-of-phase AUG codon.

The first two amino acid residues of the nascent *c*II protein are removed *in vivo* to form the mature, active protein (37). Because the second amino acid is not present in the normal protein, *c*II mutations in the second codon cannot be ordinary missense mutations. They must affect either the processing step or the efficiency of *c*II translation. The *c*II3059 mutation, a GUU→GAU (Val→Asp) second codon mutation, does result in a large decrease in translation efficiency both *in vitro* and *in vivo* (unpublished data). In contrast, the second codon mutations *can*-1, a GUU→GCU (Val→Ala) change (11, 22), and *c*3073b, a GUU→CUU (Val→Leu) change, have little effect on translation efficiencies (unpublished data). Instead, these mutations presumably affect the processing step. Preliminary experiments on the purified *can*-1 derivative support this idea (unpublished data).

The first six amino acids of the mature *c*II protein (corresponding to codons 3–8 of the *c*II gene) are encoded by DNA that overlaps with the −35 region of the P_{RE} promoter (5, 11). Only two *c*II mutations have been identified in this region, one of which is a chain termination mutation at codon 6 (*c*II3109), and the other a substitution mutation, also at codon 6, which changes a basic amino acid to an acidic amino acid (*c*II3114). Nine of the −35 region P_{RE} mutations cause amino acid substitutions in these same 6 codons, yet none of these results in a cII^- phenotype. Even λ*c*y3075, a strain with a tandem double mutation that results in a Lys-Arg →Asn-Cys double amino acid change at codons 6 and 7, retains partial *c*II activity by complementation analysis. All of the other P_{RE} mutant strains appear to regain full *c*II activity, including λ*c*y3107, which has an Arg→Cys change at codon 7 without the accompanying change at codon 6. We conclude that the NH_2-terminal region of the *c*II protein can tolerate considerable variation of the wild-type amino acid sequence without loss in activity.

In contrast to the single *c*II missense mutation that has been found in codons 3–8, we have identified 7 different *c*II missense mutations in codons 9–15 (unpublished data). We conclude that the relative insensitivity of the *c*II protein to the precise amino acid sequence in codons 3–8 is not a general property of the protein. Why then should the *c*II gene be so flexible to change in the region of overlap with P_{RE}? We believe that the positioning of the overlap itself is important and that evolution has placed a premium on maintaining the overlap and optimal P_{RE} promoter function. We suggest that the precise positioning of the overlap in some way coordinates the bidirectional transcriptional and translational events that occur in this region. It is likely that this mechanism functions to regulate *c*II expression autogenously by coupling the transcription/translation of *c*II to the *c*II-dependent activation of the P_{RE} promoter.

We thank J. Cinque, B. Cohen, and D. Skarecky for technical assistance and Gail Taff and Kathleen Schuff for typing and editing the manuscript. C. Debouck, Y. Ho, C. Queen, and L. Strano also contributed to this work. This work was supported by U.S. Public Health Service Research Grants GM25438 and GM28370.

1. Herskowitz, I. & Hagen, D. A. (1980) *Annu. Rev. Genet.* **14**, 399–445.
2. Reichardt, L. & Kaiser, A. D. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2185–2189.
3. Echols, H. & Green, L. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2190–2194.
4. Wulff, D. L. & Rosenberg, M. (1983) in *Lambda II*, eds. Hendricks, J., Roberts, J., Stahl, F. & Weisberg, R. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 53–73.
5. Schmeissner, U., Court, D., Shimatake, H. & Rosenberg, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3191–3195.
6. Shimatake, H. & Rosenberg, M. (1981) *Nature (London)* **292**, 128–132.
7. Schwarz, E., Scherer, G., Hobom, G. & Kössel, H. (1978) *Nature (London)* **272**, 410–414.
8. Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. L. (1978) *Nature (London)* **272**, 414–423.
9. Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F. & Petersen, G. B. (1982) *J. Mol. Biol.* **162**, 729–773.
10. Rosenberg, M. & Court, D. (1979) *Annu. Rev. Genet.* **13**, 319–353.
11. Wulff, D. L., Beher, M., Izumi, S., Beck, J., Mahoney, M., Shimatake, H., Brady, C., Court, D. & Rosenberg, M. (1980) *J. Mol. Biol.* **138**, 209–230.
12. Ho, Y., Wulff, D. L. & Rosenberg, M. (1983) *Nature (London)* **304**, 703–708.
13. Shine, J. & Delgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
14. Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. L. (1978) in *The Operon*, eds. Miller, J. H. & Reznikoff, W. S. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 304–324.
15. Wulff, D. L. (1976) *Genetics* **82**, 401–416.
16. Fowler, R., Degnen, G. & Cox, E. (1974) *Mol. Gen. Genet.* **133**, 179–191.
17. Enquist, L. W. & Weisberg, R. A. (1977) *J. Mol. Biol.* **111**, 97–120.
18. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
19. Brachet, P. & Thomas, R. (1969) *Mutat. Res.* **7**, 257–260.
20. Kaiser, A. D. (1957) *Virology* **3**, 42–61.
21. Gussin, G. N., Rosen, E. D. & Wulff, D. L. (1980) *Genetics* **96**, 1–24.
22. Jones, M. O. & Herskowitz, I. (1978) *Virology* **88**, 199–212.
23. Pribnow, D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 784–788.
24. Gilbert, W. (1976) in *RNA Polymerase*, eds. Losick, R. & Chamberlin, M. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 193–205.
25. Rosen, E. D., Hartley, J. L., Matz, K., Nichols, B. P., Young, K. M., Donelson, J. E. & Gussin, G. N. (1980) *Gene* **11**, 197–205.
26. Youderian, P., Bouvier, S. & Susskind, M. (1982) *Cell* **30**, 843–853.
27. Abraham, J., Mascarenhas, D., Fischer, R., Benedik, M., Campbell, A. & Echols, H. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2477–2481.
28. Hoess, R. H., Foeller, C., Bidwell, K. & Landy, A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2482–2486.
29. Davies, R. W. (1980) *Nucleic Acids Res.* **8**, 1765–1782.
30. Schmeissner, U., Court, D., McKenney, K. & Rosenberg, M. (1981) *Nature (London)* **292**, 173–175.
31. Stormo, G. D., Schneider, T. D. & Gold, L. M. (1982) *Nucleic Acids Res.* **10**, 2971–2996.
32. Files, J., Weber, K. & Miller, J. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 667–670.
33. Steege, D. A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4163–4167.
34. Napoli, C., Gold, L. & Singer, B. S. (1981) *J. Mol. Biol.* **149**, 433–450.
35. Hall, M., Gabay, J., Debarbouille, M. & Schwartz, M. (1982) *Nature (London)* **295**, 616–618.
36. Farabaugh, P. J. (1978) *Nature (London)* **272**, 765–769.
37. Ho, Y. & Rosenberg, M. (1982) *Ann. Microbiol. (Paris)* **133A**, 215–218.