

---

**Studies on gene control regions XII. The functional significance of a *Lac* operator constitutive mutation**

---

E.F. Fisher and M.H. Caruthers

---

Department of Chemistry, University of Colorado, Boulder, CO 80309, USA

---

Received 14 August 1979

---

**ABSTRACT**

The functional significance of a *lac* operator constitutive mutation has been determined. The transition adenine-thymine to guanine-cytosine was shown to be a constitutive mutation simply because thymine contains the functionally important 5-methyl group whereas cytosine does not. The remainder of the base pair is of no consequence. The experimental approach was to synthesize various modified operators containing cytosine, 5-methylcytosine, and 5-bromocytosine. The synthetic operator containing a guanine-cytosine base pair displays an eightfold reduction in stability with *lac* repressor whereas the operator containing 5-methylcytosine binds repressor at least as tightly as does the wild type sequence. Results published previously have shown that a similar decrease in stability of the repressor-operator complex can be obtained simply by substituting uracil for thymine or by inverting the base pair to thymine-adenine. All these results taken together implicate the thymine 5-methyl as the only important functional group recognized by the *lac* repressor at this base pair. Further confirmation of this conclusion was obtained by substitution of 5-bromocytosine and 5-bromouracil at this base pair. Both altered the stability of the repressor-operator complex by about the same percent suggesting that the bromine atom was the important determinant of complex stability for 5-bromopyrimidine analogs.

**INTRODUCTION**

Proteins involved in the regulation of gene expression presumably function by interacting with sequence specific recognition sites on DNA. Since any base pair contains many potential recognition sites (1, 2, 3), the problem therefore is to determine those functional groups on the DNA that interact with control proteins. A corollary to this question is to probe the biochemical significance of genetic mutations in gene control regions. Presumably deciphering these mutations should contribute



was indeed the 5-methyl group. If this is the case, simply inserting 5-methylcytosine for cytosine in the constitutive mutation should lead to a tight binding *lac* repressor-*lac* operator complex. The experimental approach was the following: (1) Prepare *lac* operators containing cytosine and the cytosine analogs 5-methylcytosine and 5-bromocytosine at position 13. (2) Measure the stability of these *lac* repressor-*lac* operator complexes. These experiments have shown that the 5-methyl group is the important functional group leading to the O<sup>c</sup> phenotype at this position.

#### MATERIALS AND METHODS

Procedures for the preparation and purification of various chemicals, reagents and enzymes have been published (6). T4-ligase, T4-kinase, and *E. coli* DNA Polymerase I were isolated as described previously (8). Wild type (SQ) repressor was purified by a published procedure (8). Tight binding (QX86) repressor was prepared from a strain (P90-86) made by J. Miller (9) and was a gift from Dr. J. Sadler. 5-Methyldeoxycytidine 5'-triphosphate was chemically synthesized from 5-methyldeoxycytidine 5'-phosphate and pyrophosphate (10). The triphosphate was purified by column chromatography on A-25 Sephadex and paper chromatography.

Unmodified deoxyoligonucleotides were prepared by a combination of chemical and enzymatic procedures. These syntheses have been reported previously (11, 12, 13). Two unmodified *lac* operator duplexes whose syntheses have been reported (8, 13, 14) were also used in this research (duplexes I and II). Duplex I is base pairs 1-26 and duplex II is base pairs 6-26 (Figure 2).

Enzymatic reactions using T4-ligase, T4-kinase, *E. coli* DNA Polymerase I, and deoxynucleotidyl terminal transferase were carried out according to published procedures (6, 7, 8). The membrane filter assays were performed basically as described previously (15, 16). Adaptations of these methods to our system have been published (8). Additionally  $\lambda$ d*lac* DNA was replaced as competitor by a plasmid HOE 101. This plasmid contains 12 tandem repeat copies of the *lac* operator (17). Plasmid DNA was added to a final concentration of 90  $\mu$ g per milliliter (10,000 fold

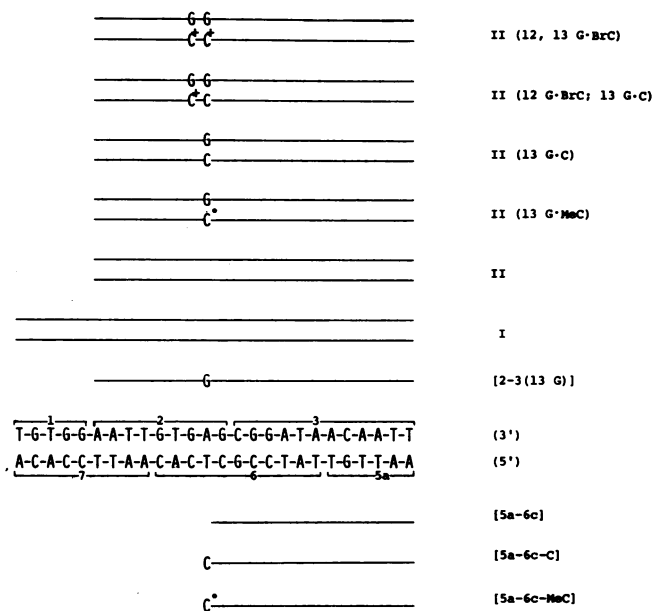


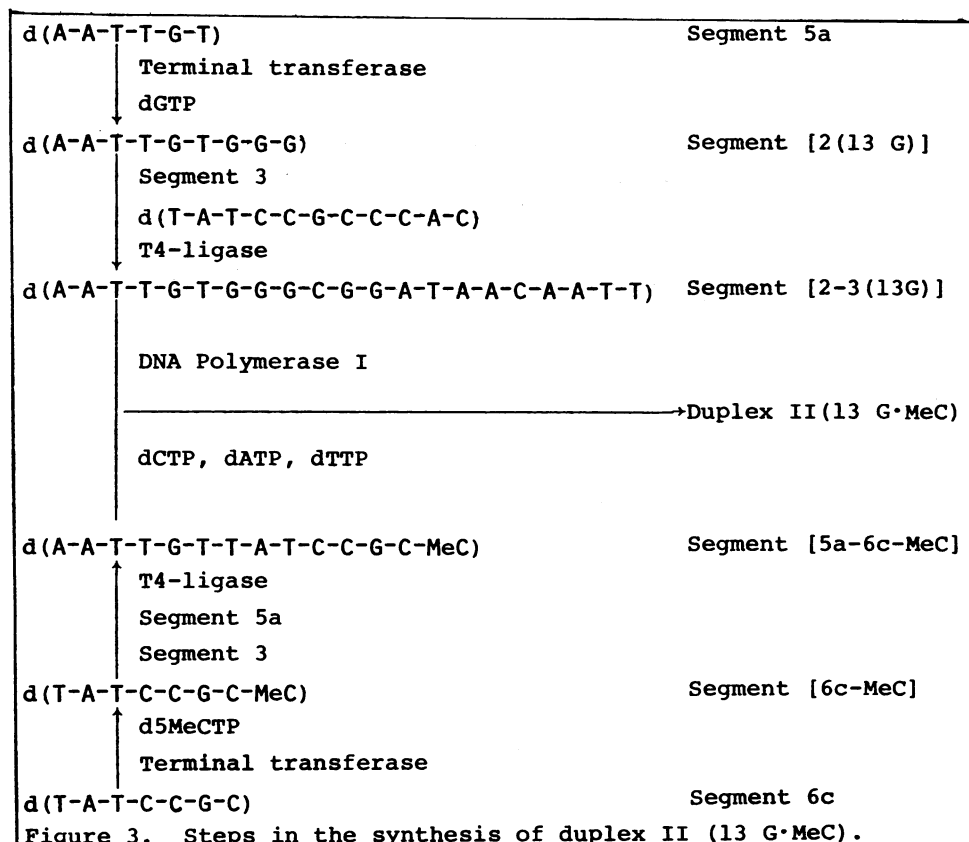
Figure 2. Chemically synthesized *lac* operator segments are partitioned and numbered. Lines above and below the operator are partial sequences. Bracketed, hyphenated numbers define the segments. Parentheses enclosing a number and symbol define a base change. For example segment [2-3 (13G)] is segments 2 and 3 covalently joined where site 13 has adenine replaced by guanine. Numbers followed by lower case letters refer to sequences that are part of the original segments. Sequences modified by nucleotide additions are described. An example is segment [5a-6c-C] which defines d(A-A-T-T-G-T-T-A-T-C-C-G-C-C). Roman Numerals define duplexes. Parentheses indicate modifications within duplexes. For example duplex II (13 G·C) refers to duplex II where site 13 has been altered to contain guanine and cytosine. The symbols C<sup>+</sup> and C<sup>°</sup> refer to 5-bromo and 5-methylcytosine respectively. In the text and as part of the duplex names BrC and MeC refer to these analogs.

excess of unlabeled *lac* operator). The duplexes used in the filter binding assays were labeled using a published procedure (18).

#### EXPERIMENTAL

Duplex II (13 G·C), duplex II (13 G·MeC), duplex II (12

G·BrC; 13 G·C) and duplex II (12, 13 G·BrC) were synthesized by a combination of chemical and enzymatic procedures. These duplexes and the deoxyoligonucleotides used in their synthesis are shown in Figure 2. An outline for the synthesis of duplex II (13 G·MeC) is shown in Figure 3. Initially segment [2 (13G)] and segment [6c·MeC] were prepared using the deoxynucleotide primers shown in Figure 3, deoxytriphosphates and deoxynucleotidyl terminal transferase. The next step was synthesis of segment [2-3 (13G)] and segment [5a-6c·MeC] using T4 ligase and the appropriate deoxyoligonucleotides as primers and templates. The final step was an *E. coli* DNA Polymerase I catalyzed repair reaction. The same synthetic steps using other modified deoxyoligonucleotides were used to prepare the remaining three duplexes. A detailed presentation of these synthetic steps fol-



lows.

Multistep synthesis of duplex II (13 G·C) and duplex II (13 G·MeC): (a) The first synthetic step was terminal deoxynucleotidyl transferase catalyzed extension of two deoxyoligonucleotides. [5'-<sup>32</sup>P] segment 5a (2860 pmoles), dGTP (7280 pmoles), and terminal deoxynucleotidyl transferase (15 units) were allowed to react in the cacodylate buffer system, including 10 mM MgCl<sub>2</sub>, as outlined previously (6). An analysis of the reaction after 6 hr indicated that the product, [5'-<sup>32</sup>P] segment [2(13G)], and several other homologues had formed. The analytical pattern is shown in Figure 4, gel 1. The product (301 pmole) was isolated by preparative gel electrophoresis. In a similar manner, [5'-<sup>32</sup>P] segment [6c-MeC] was prepared from [5'-<sup>32</sup>P] segment 6c and 5-methyldeoxycytidine 5'-triphosphate. Gel 2 in Figure 4 displays the analytical gel pattern following reaction with terminal deoxynucleotidyl transferase. The isolated yield of [5'-<sup>32</sup>P] segment [6c-MeC] was 41 pmole (12%). (b) The second step was two T4-ligase catalyzed joining reactions. The first ligation was the joining of segment [2(13G)] (301 pmole) to [5'-<sup>32</sup>P] segment 3 (350 pmole) using d(T-A-T-C-C-G-C-C-A-C) as template (380 pmole). Before use in the ligation step, [5'-<sup>32</sup>P] segment [2(13G)] was treated with alkaline phosphatase in order to remove the 5'-phosphate. In this way the dimer of the product duplex was avoided. As shown by gel electrophoretic analysis (gel 3, Figure 4), the formation of segment [2-3(13G)] was approximately 20% complete after 48 hrs. The isolated yield of this segment was 50 pmole. The second ligation was the joining of segment 5a (136 pmole) to [5'-<sup>32</sup>P] segment [6c-MeC] (41 pmole) using segment 3 (77 pmole) as template. The isolated yield of the product, segment [5a-6c-MeC] was 20 pmole. (c) The final step was two DNA polymerase I catalyzed repair reactions. Duplex II (13 G·C) was prepared using DNA polymerase I (3 units), segment [2-3 (13G)] as template (8.3 pmole), and segment [5a-6c] as primer (8.1 pmole). The reaction mixture also contained potassium phosphate buffer (pH 6.9, 0.12 mM), MgCl<sub>2</sub> (8mM), DTT (5 mM), and dTTP, dCTP, and dATP (0.1 mM each). Gel electrophoretic analysis of the reaction mixture is shown in Figure 5, channels a and b. The yield isolated by preparative gel electro-

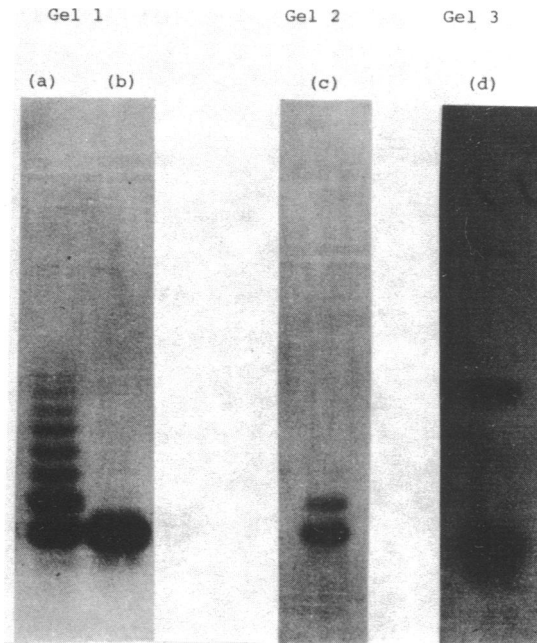


Figure 4: Synthesis of [5'-<sup>32</sup>P] segment [2 (13G)], [5'-<sup>32</sup>P] segment [6c-MeC], and segment [2-3 (13G)]. Extent of reaction was monitored by gel electrophoresis. The denaturing gels were 20% acrylamide, 1% N, N-methylenebisacrylamide in 89 mM Tris-borate (pH 8.3), 2.2 mM EDTA (TBE) and 7 M urea. Gel 1 shows the results from the synthesis of [5'-<sup>32</sup>P] segment [2 (13G)]. Channel a contained the products of a deoxynucleotidyl terminal transferase reaction using dGTP and [5'-<sup>32</sup>P] segment 5a. Channel b contained [5'-<sup>32</sup>P] segment 5a. Gel 2, channel c, shows a similar enzyme catalyzed reaction using 5-methyldeoxycytidine 5'-triphosphate and [5'-<sup>32</sup>P] segment 6c as primer. Gel 3, channel d, shows the analysis of the ligation reaction (see text) used to prepare segment [2-3 (13G)]. The slow migrating band is the desired product. The other band is [5'-<sup>32</sup>P] segment 3.

phoresis was 1.9 pmole. Duplex II (13 G·MeC) was prepared similarly except that the primer was segment [5a-6c-MeC]. Analysis of the repair reaction is shown in channels c and d of the gel displayed in Figure 5. Isolation by preparative gel electrophoresis gave 2.2 pmole of the duplex.

Duplex II (12, 13 G·BrC) and duplex II (12 G·BrC; 13 G·C)

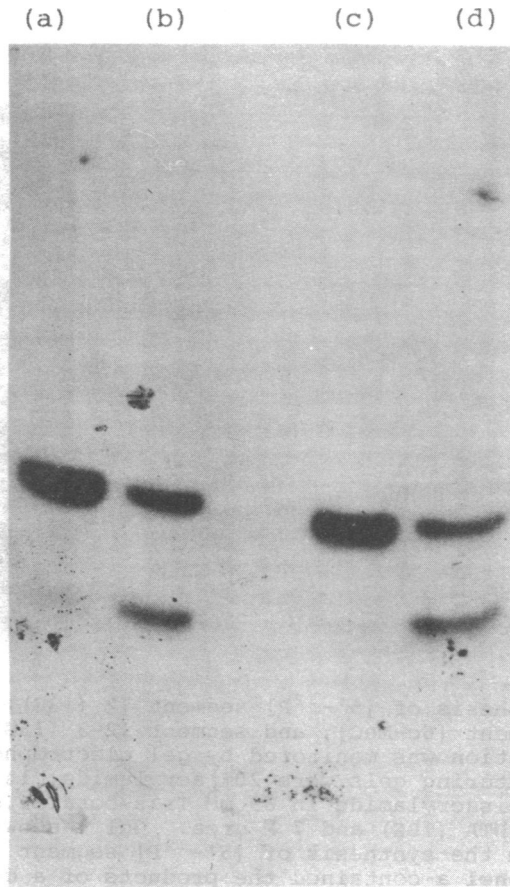


Figure 5: Synthesis of duplex II (13 G·C) and duplex II (13 G·MeC). Analysis was by gel electrophoresis using denaturing conditions as outlined in Figure 4. The synthesis of duplex II (13 G·C) is shown in channels a and b. Channel b contained the reaction mixture including the primer, segment [5a-6c], and the template, segment [2-3 (13G)]. Channel a contained the reaction mixture after addition of DNA Polymerase I. The synthesis of duplex II (13 G·MeC) is shown in channels c and d. Channel d contained the primer, segment [5a-6c-MeC], and template, segment [2-3 (13G)]. Channel c shows the extent of reaction after 5 hrs. All deoxy-oligonucleotides were internally labeled.

were prepared by two series of similar reaction steps. Duplex II (12, 13 G·BrC) was prepared using two successive repair reactions



with *E. coli* DNA Polymerase I. The first reaction used segment [2-3 (13G)] whose synthesis is described above, as template (8 pmole), segment [5a-6c] as primer (8 pmole) and d5BrCTP. Gel electrophoresis analysis of this reaction (Figure 6, channels a and b) indicated that the expected two nucleotides had added to

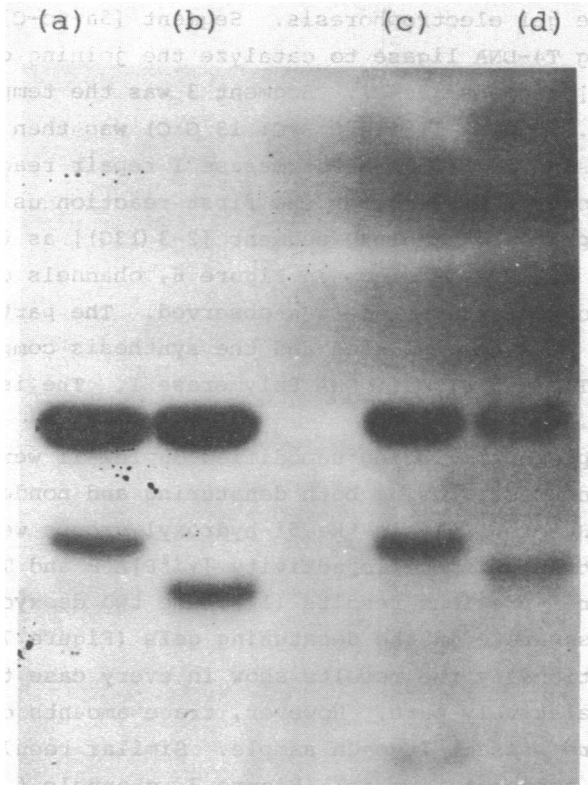


Figure 6: Initial steps in the synthesis of duplex II (12, 13 G·BrC) and duplex II (12 G·BrC; 13 G·C). The first repair reaction for the synthesis of duplex II (12, 13 G·BrC) is shown in channels a and b. Channel b shows the reaction mixture containing d5BrCTP, segment [5a-6c], and segment [2-3 (13G)]. Channel a shows the same reaction mixture 5 hrs after addition of DNA Pol I. The first repair reaction for the synthesis of duplex II (12 G·BrC; 13 G·C) is shown in channels c and d. Channel d shows the reaction mixture containing d5BrCTP, segment [5a-6c-C], and segment [2-3 (13G)]. Channel c shows the same reaction mixture 5 hrs after addition of DNA Pol I. All deoxyoligonucleotides were internally labeled.

the primer. This partially repaired duplex was then isolated and the synthesis completed using dCTP, dATP, dTTP and *E. coli* DNA Polymerase I. The isolated yield was 1.5 pmole. Duplex II (12 G·BrC; 13 G·C) was synthesized in the following manner. Segment [6c-C] was first prepared from segment 6c using dCTP and terminal deoxynucleotidyl transferase. The product was isolated by preparative gel electrophoresis. Segment [5a-6c-C] was synthesized using T4-DNA ligase to catalyze the joining of segment 5a to [5'-<sup>32</sup>P] segment [6c-C]. Segment 3 was the template for this reaction. Duplex II (12 G·BrC; 13 G·C) was then prepared by two successive *E. coli* DNA Polymerase I repair reactions. The gel electrophoresis pattern for the first reaction using segment [5a-6c-C] as primer (8 pmole), segment [2-3 (13G)] as template (8 pmole) and d5BrCTP is shown in Figure 6, channels c and d. The addition of one nucleotide was observed. The partially repaired duplex was then isolated and the synthesis completed using dCTP, dATP, dTTP and *E. coli* DNA Polymerase I. The isolated yield was 1.1 pmole.

These duplexes including unmodified duplex II were analyzed by gel electrophoresis using both denaturing and nondenaturing conditions. Before analysis the 5'-hydroxyl groups were phosphorylated with high specific activity [ $\gamma$ <sup>32</sup>P]ATP and T4 kinase. As expected from previous results (14), the two deoxyoligonucleotide strands separate on the denaturing gels (Figure 7, channels a - e). Additionally the results show in every case that the two strands are relatively pure. However, trace amounts of various impurities were present in each sample. Similar results were observed on the nondenaturing gel (Figure 7, channels f - i).

## RESULTS

The dissociation kinetics of preformed repressor-operator (RO) complexes were measured using the nitrocellulose filter assay. For each analysis, [<sup>32</sup>P] operator and repressor were allowed to form a complex. A large molar excess of unlabeled *lac* operator as plasmid HOE 101 was then added and the dissociation kinetics were monitored using the nitrocellulose filter assay. The dissociation kinetics of *lac* operator modified by insertion

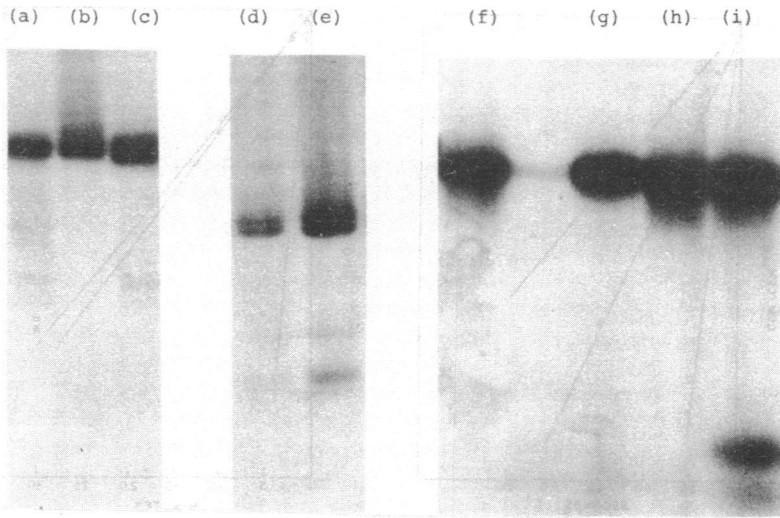


Figure 7. Analysis of [5'- $^{32}$ P] phosphate labelled duplexes. Analyses were by gel electrophoresis using denaturing conditions (channels a - e) as outlined in Figure 4 and nondenaturing conditions (TBE only) for channels (f - i). (a) Duplex II (12 G·BrC; 13 G·C); (b) duplex II (12, 13 G·BrC); (c) duplex II (13 G·C); (d) duplex II (13 G·MeC); (e) duplex II; (f) duplex II (12 G·BrC; 13 G·C); (g) duplex II (12, 13 G·BrC); (h) duplex II (13 G·MeC); (i) duplex II.

of 5-methylcytosine at site 13 are shown in Figure 8 using SQ repressor (panel A) and QX86 repressor (panel B). Control operators were duplex II, the unmodified operator, and duplex II (13 G·C), an operator containing the  $O^C$  mutation at site 13. Introduction of the G-C base pair dramatically reduced the stability of the RO complex. With SQ repressor, the half life was 10 sec as compared to 52 sec for duplex II. With QX86 repressor, the same operators had half-lives of 1.3 minutes and 13.7 minutes respectively. By substitution of 5-methylcytosine for cytosine, the stability of the RO complex was increased several fold. The half-life with SQ repressor increased to 72 seconds and the half-life with QX86 repressor increased to 12.8 minutes. The RO complexes containing 5-methylcytosine are therefore at least as stable as RO complexes containing an A-T base pair at this posi-

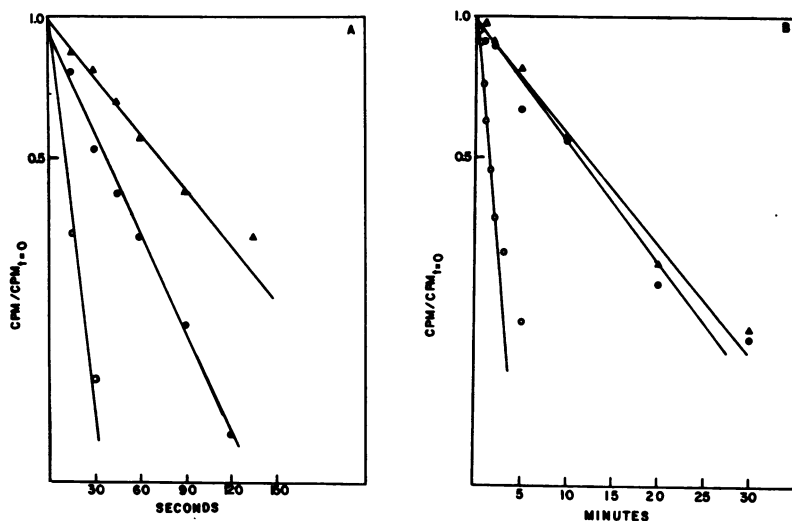


Figure 8. Dissociation kinetics of the duplex II operator containing 5-methylcytosine and of related duplexes from SQ (panel A) and QX86 (panel B) repressors. The ionic strength was 0.05 M. ●, duplex II; Δ, duplex II (13 G·MeC); o, duplex II (13 G·C).

tion. Results obtained with 5-bromocytosine are presented in Figure 9. The measured half-lives with SQ repressor and 5-bromocytosine containing operators were as follows: duplex II (12 G·BrC; 13 G·C), 9 sec; duplex II (12, 13 G·BrC), 21 sec. With QX86 repressor the same operators had half-lives of 1.8 and 6.6 min respectively. Insertion of bromine at site 12 in the constitutive operator therefore had little effect on the stability of the RO complex. However, insertion of bromine at the constitutive base pair dramatically increased the stability of the RO complex. All these results are summarized in Table 1.

#### DISCUSSION

These results show how a mutation located in a gene control region functions. In particular the A-T to G-C transition at position 13 in the *lac* operator (Figure 1) is a constitutive mutation simply because thymine has a 5-methyl group which interacts with *lac* repressor whereas cytosine does not. The remainder

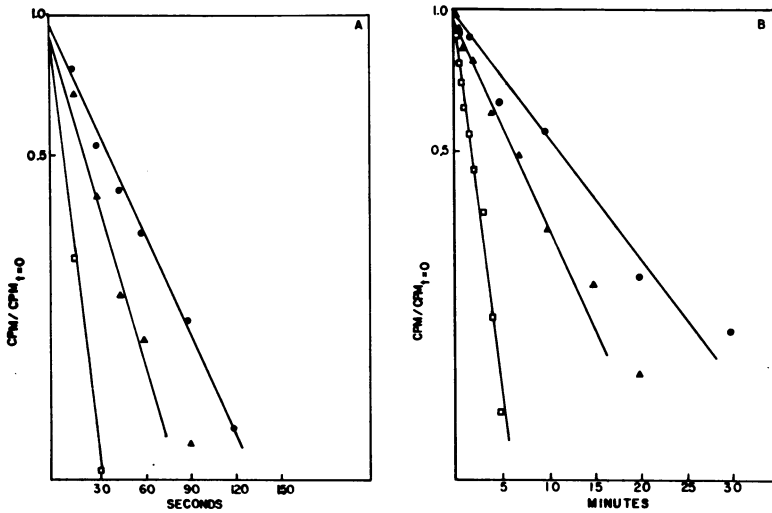


Figure 9. Dissociation kinetics of duplex II and duplex II operators containing 5-bromocytosine from SQ (panel A) and QX86 (panel B) repressors. The ionic strength was 0.05 M. ●, duplex II; □, duplex II (12 G-BrC; 13 G-C); ▲, duplex II (12, 13 G-BrC).

of these base pairs (A-T or G-C) contribute nothing to the specificity of the RO interaction. The evidence for these conclusions is as follows. Previous results included in Table 1 have shown that removal of the 5-methyl group from thymine (via insertion of uracil) dramatically reduces the RO complex stability. The increase in binding free energy was 5.26 kJ/mol with QX86 repressor. An identical increase in binding free energy (5.31 kJ/mol) was observed for the  $O^C$  mutation (GC in place of AT; duplex II (13 G-C) in Table 1). Similar results were obtained with SQ repressor. Worth emphasizing is that the remainder of the base pair is unmodified when adenine-uracil replaces adenine-thymine. The converse occurs when guanine-5-methylcytosine replaces adenine-thymine at this position. The 5-methyl group is held constant whereas the remainder of the base pair has been changed to the operator constitutive mutation. And yet, the duplex containing 5-methylcytosine forms an RO complex which is slightly more stable than the wild type operator. The decrease in binding free

Table 1. Dissociation half-lives, equilibrium dissociation constants, and binding free energy changes for modified *lac* operators.<sup>a</sup>

Operator	SQ Repressor			QX86 Repressor		
	$t_{1/2}$ (sec)	$K_d$ ( $\times 10^{12}M$ )	$\Delta\Delta G^b$ (kJ/mol)	$t_{1/2}$ (min)	$K_d$ ( $\times 10^{12}M$ )	$\Delta\Delta G^b$ (kJ/mol)
II <sup>C</sup>	52	7.4		13.7	21.1	
II (13 G·C)	10	38.5	(-)3.73	1.6	218.5	(-)5.31
II (13 G·MeC)	72	5.3	(+)0.77	12.8	22.5	(+)0.15
II (12, 13 G·BrC)	21	18.6	(-)2.08	6.6	43.7	(-)1.65
II (12 G·BrC; 13 G·C)	9	44.6	(-)4.07	1.8	159.4	(-)4.59
I <sup>C</sup>	57	6.1		19.9	14.5	
I (13 A·U) <sup>C</sup>	<10 <sup>d</sup>			2.0	147.1	(-)5.26
I (13 A·BrU) <sup>C</sup>	42	8.2	(-)0.70	12.6	23.0	(-)1.04

- All the data was analyzed by the least squares fit method.
- All  $\Delta\Delta G$  values were determined relative to the wild type, which was arbitrarily assigned a  $\Delta\Delta G$  value of zero.
- Previously reported results (5, 6) were modified slightly by reanalysis of the data using the least squares fit method.
- The dissociation of this RO complex was complete by 15 sec.

energy was 0.15 kJ/mol with QX86 repressor and 0.77 kJ/mol with SQ repressor. Insertion of a bromine atom for a hydrogen at the 5-pyrimidine position also appears to probe this specific *lac* repressor-*lac* operator contact site. With QX86 repressor, the relative effects can be seen by comparing the binding free energy changes for duplex I (13 A·U) and duplex I (13 A·BrU), 5.26 and 1.04 respectively, or for duplex II (12 G·BrC; 13 G·C) and duplex II (12, 13 G·BrC), 4.59 and 1.65 respectively. Insertion of bromine for hydrogen at the 5 position of either uracil or cytosine therefore stabilizes the RO complexes. However, the 5-bromopyrimidine complexes were less stable than those derived from operators containing 5-methylpyrimidines. This was the expected result if the interaction were hydrophobic. A bromine atom is less hydrophobic than a methyl group but is certainly superior to a hydrogen atom (19, 20). Therefore all the evidence is con-

sistent with the interpretation that the transition A-T to G-C is an operator constitutive mutation simply because cytosine does not contain a 5-methyl group.

\*This is paper XII in a series on "Studies of Gene Control Regions". Paper XI is reference 21. This work was supported by grants from the National Institutes of Health (GM21120), and the National Science Foundation (PCM76-01489). E. F. was supported by an American Cancer Society Institutional Research Grant. M. H. C. was supported by a Career Development Award from the National Institutes of Health (1 KO4 GM00076).

#### REFERENCES

1. Seeman, N. C., Rosenberg, J. M. and Rich, A. (1976) *Proc. Nat. Acad. Sci. U.S.A.* 73, 804-808
2. von Hippel, P. H. (1969) *J. Cell Physiol.* 74, Suppl. 1, 235-238
3. Yarus, M. (1969) *Ann. Rev. Biochem.* 38, 841-880
4. Gilbert, W., Gralla, J., Majors, J. and Maxam, A. M. (1975) in *Protein-Ligand Interactions*, Sund, H. and Blauer, G., Eds., pp 193-210. DeGruyter, Berlin, West Germany
5. Goeddel, D. V., Yansura, D. G. and Caruthers, M. H. (1977) *Nucleic Acids Research* 4, 3039-3054
6. Goeddel, D. V., Yansura, D. G., Winston, C. and Caruthers, M. H. (1978) *J. Mol. Biol.* 123, 661-687
7. Sista, H. S., Loder, R. T. and Caruthers, M. H. (1979) *Nucleic Acids Research* 6, 2583-2599
8. Yansura, D. G., Goeddel, D. V., Cribbs, D. L. and Caruthers, M. H. (1977) *Nucleic Acids Research* 4, 723-737
9. Jobe, A., and Bourgeois, S. (1972) *J. Mol. Biol.* 72, 139-152
10. Smith M., and Khorana, H. G., (1958) *J. Am. Chem. Soc.* 80, 1141-1145
11. Goeddel, D. V., Yansura, D. G., Kawashima, E., Gadek, T. and Caruthers, M. H. (1976) in *Molecular Mechanisms in the Control of Gene Expression*, Nierlich, D. P., Rutter, W. J. and Fox, C. F., Eds., pp. 159-164, Academic Press Inc., New York
12. Goeddel, D. V., Yansura, D. G. and Caruthers, M. H. (1977) *Biochemistry* 16, 1765-1772
13. Yansura, D. G., Goeddel, D. V., and Caruthers, M. H. (1977) *Biochemistry* 16, 1772-1780
14. Goeddel, D. V., Yansura, D. G. and Caruthers, M. H. (1977) *Proc. Nat. Acad. Sci. U.S.A.* 74, 3292-3296
15. Riggs, A. D., Suzuki, H., and Bourgeois, S. (1970) *J. Mol. Biol.* 48, 67-83
16. Riggs, A. D., Bourgeois, S., and Cohn, M. (1970) *J. Mol. Biol.* 53, 401-417
17. Sadler, J., unpublished results

18. Lillehaug, J. R., Kleppe, R. K., and Kleppe, K. (1976) *Biochemistry* 15, 1858-1865
19. Wilhelm, E., Battino, R., and Wilcox, R. (1977) *Chem. Rev.* 77, 219-262
20. Gill, S. J. and Wadso, I (1976) *Proc. Nat. Acad. Sci. U.S.A.* 73, 2955-2958
21. Yansura, D. G., Goeddel, D. V., and Caruthers, M. H. (1979) *Biochemistry*, submitted for publication.