

Contacts between the *lac* repressor and thymines in the *lac* operator

(*E. coli* genetic control/bromodeoxyuridine substitution/ultraviolet light crosslinking/DNA sequencing)

RONALD OGATA AND WALTER GILBERT

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

Contributed by Walter Gilbert, September 8, 1977

ABSTRACT We have identified important points of contact between the *lac* repressor and the *lac* operator by crosslinking the repressor to bromouracil-substituted operator. We substituted bromouracils for thymines in a 55-base-long restriction fragment containing the *lac* operator and labeled one or the other 5' end with ³²P. Ultraviolet irradiation of this fragment produced single-strand breaks at the bromouracils. We examined breakage at each bromouracil in the sequence by denaturing the DNA and displaying the UV-generated fragments on a polyacrylamide gel. In the presence of *lac* repressor, UV radiation failed to break at specific sites. We attribute this to a competing reaction in which the DNA crosslinks to the repressor rather than breaking. These crosslinkable sites thus define positions at which the *lac* repressor protein lies close to the methyl group of a thymine in the major groove of DNA.

How do macromolecules recognize each other? When two large molecules interact, groups on each make close contact. If we introduce short crosslinks into this complex, they will bind together regions that neighbor each other; the positions of these crosslinks thus will identify points of contact. Consider control proteins that interact with DNA. Irradiation with UV light can induce zero-range crosslinks between these proteins and 5-bromodeoxyuridine (BrdUrd)-substituted DNA (1-3). Here we shall describe an experiment that identifies the position of such UV-induced crosslinks on a DNA molecule.

When bromine is substituted for the methyl group of thymine, by incorporating bromouracil into DNA, the DNA remains functional, *in vivo* and *in vitro*, although some changes occur in the binding of proteins. For example, Lin and Riggs (4) observed that the *lac* repressor binds 10-fold more tightly to bromouracil-substituted *lac* operator than to unsubstituted operator. Irradiation of substituted DNA with UV light displaces the bromine and generates a free radical at the 5-position on the deoxyuridine. If a protein, bound to this DNA, has an appropriate amino acid side chain close to the free radical, a covalent crosslink between the protein and the DNA can form (1-3). This zero-range crosslink obviously identifies a close contact between the protein and its binding site on DNA. In the absence of a reactive group, the light-induced free radical on deoxyuridine usually extracts a hydrogen atom from the neighboring sugar on the 5' side; that sugar then decomposes to give a single-strand break and a 5'-terminal deoxyuridine (1). We use this ability to break BrdUrd-substituted DNA with UV light to identify the positions of the bromouracils in a DNA sequence. A bound protein quenches the breakage of DNA as crosslinks are formed; this identifies the points of crosslinking.

We isolate a 55-base-long double-stranded DNA fragment that contains the *lac* operator from BrdUrd-substituted DNA and label the 5' end of one strand only with ³²P. This can be

done because the fragment has been cut out of a longer DNA molecule by two different restriction enzymes. We irradiate this DNA fragment with UV light, denature it in alkali, and electrophorese the strands on a polyacrylamide gel. The autoradiograph of this gel shows a pattern of lengths corresponding to breaks at the BrdUrds in the labeled strand of the fragment. If we irradiate this fragment in the presence of the *lac* repressor, at those points at which the protein is close to a bromouracil, the free radical can interact with the protein rather than breaking the DNA. Electrophoresis and autoradiography then show a decrease in the intensity of specific bands corresponding to some of the bromouracils. Associated with this suppression of breakage is the appearance of radioactivity at the top of the gel. We conclude that these effects are due to the formation of crosslinks between repressor and DNA at the suppressed sites. Thus, we identify those thymines in the operator whose methyl groups are close to the repressor.

MATERIALS AND METHODS

lac repressor was purified from strain M96 by previously published procedures (5). *Hpa* II was prepared by the method of Sharp *et al.* (6). *Alu* I and *Eco*RI were gifts from U. Siebenlist and M. Pasek, respectively. *Hpa* II and *Alu* I digestions were carried out at 37° in 10 mM Tris/10 mM MgCl₂/1 mM dithiothreitol/1 mM EDTA, pH 7.4. *Eco*RI digestions were carried out at 37° in 0.1 M Tris/50 mM NaCl/5 mM MgCl₂/1 mM dithiothreitol/0.2% Triton X-100, pH 7.5.

BrdUrd-Substituted *lac* Operator-Containing DNA. The DNA was labeled with ³²P in the top (bottom) strand by first cutting a 203-base-pair-long restriction fragment containing the *lac* control region (7) with *Hpa* II (*Alu* I), treating with *Escherichia coli* alkaline phosphatase (Worthington, BAPF), labeling the 5' ends (8), and recutting with *Alu* I (*Hpa* II). The final mixture was separated on a polyacrylamide gel (8% acrylamide, 0.27% bisacrylamide, 50 mM Tris-borate, 1 mM EDTA, pH 8.3), and the 55-base-long operator fragment labeled in the top or bottom strand was cut out and eluted (8).

The 203-Base-Pair Fragment. This was isolated from a recombinant plasmid that was constructed by F. Fuller, using methods described in ref. 9. P1 level containment was used in experiments involving this plasmid. The plasmid was placed in a thymine-requiring host (FMA10 from F. Ausubel; W3102, r⁻, TdR⁻). Cells were grown in M9 plus casamino acids (10) supplemented with thymidine (2 µg/ml). At OD₅₅₀ of 1.0-1.2, chloramphenicol (35 mg/ml in ethanol) and BrdUrd (10 mg/ml in water) were added to a final concentration of 170 µg/ml and 20 µg/ml, respectively. Incubation at 37° was continued overnight. Cells were harvested and the BrdUrd-substituted plasmid DNA was extracted by the method of Guerry *et al.*

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

Abbreviations: BrdUrd, 5-bromodeoxyuridine; PEG, polyethylene glycol; TE buffer, 10mM Tris/mM EDTA, pH 7.5.

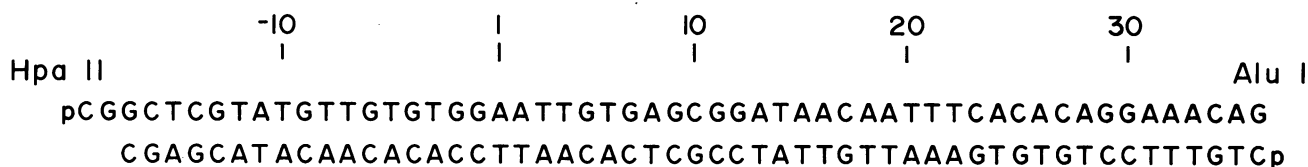


FIG. 1. Sequence of the 55-base fragment generated by cutting the *lac* control region with *Hpa* II and *Alu* I (17). The numbering system refers to the start of transcription (18).

(11). The plasmid-containing supernatant was precipitated by adding 0.6 volume of isopropanol or 0.75 volume of 20% polyethylene glycol (PEG) (12). After centrifugation, the pellet was resuspended in 10 mM Tris/1 mM EDTA, pH 7.5 (2–3 ml per liter of cells). Solid cesium chloride was added (1.35 g/ml of DNA solution) and the solution was centrifuged at 38,000 rpm for 40 hr at 15° in a Beckman type 40 rotor. The bottom of the centrifuge tube was punctured and fractions were collected. Aliquots of each fraction were assayed on a 1% agarose gel in E buffer (13). Plasmid-containing fractions were pooled and dialyzed for 1 hr against 0.5 M ammonium acetate, and the DNA was precipitated with ethanol (8). The pellet was rinsed with 80% ethanol and redissolved in 10 mM Tris/1 mM EDTA, pH 7.5 (TE buffer) to a concentration of about 1 mg/ml. At this point, the plasmid was either cut with *Eco*RI or first reprecipitated in the cold with 5.5% PEG (12) to remove contaminating RNA and redissolved in TE buffer.

Cutting with *Eco*RI releases the 203-base-pair fragment and leaves the rest of the plasmid intact. After cutting, the DNA was precipitated with ethanol and redissolved in TE buffer to a concentration of about 1 mg/ml. The solution was made 4.5% in PEG and 0.5 M in NaCl and allowed to stand at room temperature overnight. Under these conditions, most of the plasmid DNA precipitates while about 90% of the 203-base-pair fragment stays in solution. The mixture was spun in a clinical centrifuge at room temperature and the supernatant was decanted. The 203-base-pair fragment in the PEG supernatant was precipitated with ethanol. The PEG pellet was resuspended and reprecipitated with 4.5% PEG to recover the rest of the fragment. The DNA in the second supernatant was precipitated with ethanol and combined with the material from the first PEG precipitation.

The combined material from the fractional PEG precipitations was run on a preparative polyacrylamide gel (8% acrylamide, 0.27% bisacrylamide in Tris-borate/EDTA as described above). Loading of the gel did not exceed 0.3 mg of DNA per cm² of cross-sectional area in the sample well. The DNA was visualized by supporting the gel on SaranWrap, placing it on a fluorescent silica gel plate, and illuminating with UV light (14). Minimum background absorption is obtained when the gel is aged for 24 hr. The 203-base-pair fragment was eluted from the gel (8), precipitated with ethanol, and resuspended in TE buffer. The yield of this fragment is generally about 15 μg/liter of cells. We estimate (15) that the fragment is 50% substituted with BrdUrd.

UV Irradiation. This procedure was carried out with a germicidal lamp positioned about 1.5 cm above a 20-μl sample placed on a sheet of Parafilm. Reactions were carried out in binding buffer (10 mM Tris/10 mM magnesium acetate/10 mM KCl/0.1 mM EDTA/0.1 mM dithiothreitol/5% dimethyl sulfoxide/bovine serum albumin (50 μg/ml), pH 7.4) (16). After irradiation, 1 μg of sonicated calf thymus DNA in 20 μl of 5 M ammonium acetate was added to each sample. The DNA was precipitated with ethanol, washed, dried, resuspended in electrophoresis sample buffer (0.05 M NaOH/0.5 mM EDTA/5 M urea/0.025% xylene cyanol/0.025% bromophenol blue), and

run on a "sequencing gel"; the gel was autoradiographed (see ref. 8 for detailed procedure).

RESULTS AND DISCUSSION

Fig. 1 shows the sequence of the 55-base-long fragment that spans the distance between an *Alu* I (from *Arthrobacter luteus*) cut and a *Hpa* II (from *Haemophilus parainfluenzae*) cut. We incorporated BrdUrd biosynthetically by growing this DNA in the presence of BrdUrd. The DNA fragment was labeled in

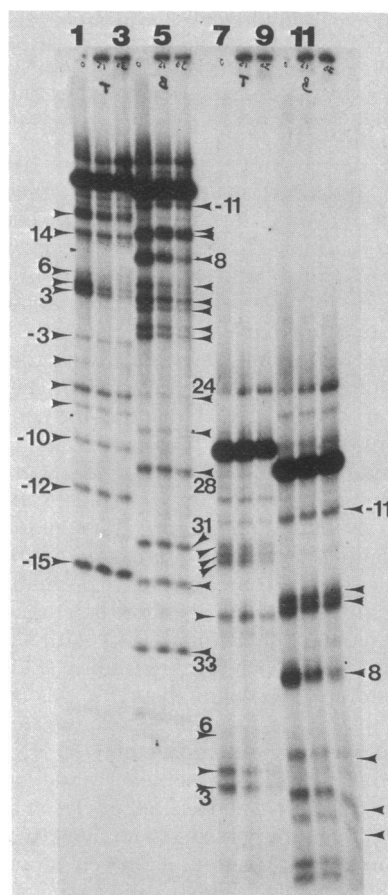


FIG. 2. Autoradiograph of an acrylamide gel showing UV-induced breakage of the 55-base fragment labeled in the top (lanes 1–3, 7–9) and bottom (lanes 4–6, 10–12) strands and the effect of *lac* repressor. Lanes 1–6 were loaded after the xylene cyanol in lanes 7–12 had migrated halfway down the gel. Here, about 33 ng (0.9 pmol) of labeled fragment in 20 μl of binding buffer was irradiated for 90 sec in the absence (lanes 1, 4, 7, 10) or presence of 225 ng (lanes 2, 5, 8, 11) or 500 ng (lanes 3, 6, 9, 12) of *lac* repressor. The arrows point to the bands generated by UV-induced breakage of the DNA and the numbers next to the arrows refer to the corresponding positions in the DNA sequence (see Fig. 1). Clearly, UV breakage is very nonuniform. The very heavy band in each lane is due to uncut strands; the bottom strand migrates slightly faster than the top strand. At the top of the gel (immediately below the lane numbers), crosslinked label appears in those samples irradiated in the presence of repressor.

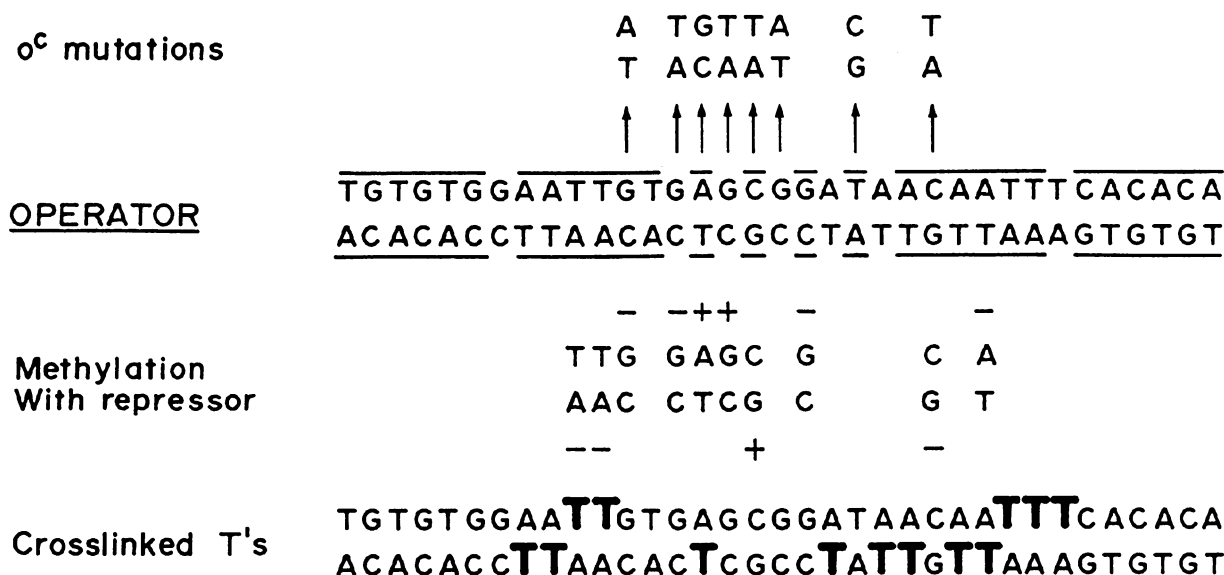


FIG. 3. *lac* operator sequence. There are two hyphenated symmetry regions on the DNA extending over 35 base pairs. The positions of the o^c mutations that have been sequenced (7) and the positions at which the *lac* repressor inhibits or enhances DNA methylation (17) are shown. The positions of the thymines that crosslink are shown in bold face.

the top strand by first cutting with the *Hpa* II enzyme, labeling with ³²P with polynucleotide kinase, and then cutting secondarily with the *Alu* I enzyme. The bottom strand was labeled by cutting with *Alu* I, labeling, and then cutting with *Hpa* II.

Irradiation with a germicidal lamp for 1–2 min breaks the DNA at the positions of each of the bromouracils. This amount of irradiation produces fewer than one break per DNA molecule. Fig. 2 shows an autoradiograph of irradiated DNA, denatured in alkali and electrophoresed on a 20% polyacrylamide gel in 7 M urea. Each band represents breakage at a bromouracil in the labeled strand. The positions of the BrdUrd can be determined by comparison either with the known sequence of this DNA or with a parallel display of this DNA broken by our chemical sequencing method (8). The fragments broken by UV light are shorter by one base than those broken chemically because UV breakage is at the sugar 5' to the BrdUrd while chemical breakage occurs at the sugar attached to the bromouracil. The breakage is not uniform: the molecule breaks more readily at some bromouracils than at others. The effect is primarily sequence-dependent (unpublished data); we do not know if this is a consequence of the local structure of DNA influencing the breakage or changing the target size for the UV hit.

Fig. 2 also shows the result of the crosslinking experiment. When the irradiation was carried out in the presence of repressor, certain bands lost intensity in the UV-induced breakage pattern. Label also appeared at the top of the gel; we believe that this is due to crosslinking to the protein, rendering the labeled DNA immobile in the gel. Thus, the bands that lose intensity identify for us the thymines close enough to protein to crosslink. In the top strand, five bands show a decrease in intensity. These correspond to positions 3, 4, 20, 21, and 22 in the sequence. Of these, positions 3 and 4 are most strongly affected. In the bottom strand, eight bands corresponding to positions 1, 2, 8, 13, 15, 16, 18, and 19 show a decrease in intensity. Of these, position 8 and to a lesser extent positions 13, 15, 18, and 19 are strongly affected. Although we assume, for simplicity, that suppression is the result of crosslinking, it may also occur without crosslinking: the protein could donate a hydrogen atom to the uracil free radical. We have not distin-

guished between these possibilities here because both mechanisms require an appropriate amino acid in close proximity to the free radical.

Fig. 3 shows the positions of the crosslinkable thymines in the operator sequence. The major points of contact outlined by this experiment lie within the central region of the hyphenated symmetries on the DNA molecule. All but one of the contacts shown by this experiment lie within the 21-base stretch across the center of the operator region. We do not see contacts in the outermost wings of the symmetrical sequence; thus, this experiment further supports the idea that such regions have a minor, if any, role in repressor binding (7, 17, 19, 20). In addition, the three strongest crosslinking sites—3, 4, and 8—are all on the left side of the symmetry region. This result, the location of operator constitutive mutations (7), and the results of methylation studies (17) demonstrate that the left side of the symmetry is most important for repressor binding.

The crosslinks are most likely in the major groove, at the 5-position on the uracil. Because the breakage reaction represents the abstraction of a neighboring hydrogen atom, a few angstroms away, and the crosslinking reaction interferes with or is more rapid than this, we argue that the crosslinking reaction must be a reaction with a group on the protein that is in intimate contact with the 5-position. It is unlikely that the free radical can wander around the uracil, or to the adjoining sugar, so as to crosslink the protein elsewhere on the DNA.

This experiment suggests that there are regions on DNA at which the repressor lies in the major groove. Fig. 3 shows points of contact revealed in a dimethyl sulfate methylation experiment (17); these experiments agree in delineating the same region of DNA as an area of contact. Both experiments suggest points of contact in the major groove, but the contacts do not intersperse in detail in a simple fashion: the experiments suggest that the repressor protein does not completely fill the major groove.

The situation that we see on the *lac* operator is entirely different from that seen in the interaction of the *lac* repressor with alternating poly[d(A,T)]. In that case, two lines of experiments suggest that the contacts are not in the major groove. Richmond and Steitz (21) filled the major groove of a d(A,T) analog, d(A,U-HgX), with various sulfhydryl reagents bound to the

mercuries. That blockage of the groove had no effect on the binding of the repressor. Second, we expect that the binding of the repressor on alternating d(A,T) mimics the background nonspecific binding of the repressor to non-operator-containing DNA. Kolchinsky *et al.* (22) have argued, based on methylation experiments, that the background binding involves the minor groove rather than the major groove. Furthermore, we do not observe crosslinking of the repressor to alternating poly[d(A-BrdUrd)]. Thus, we infer again that the binding of the repressor to poly[d(A,T)] does not involve the major groove and second, that the free radical does not migrate to the minor groove.

It is unlikely that our results simply reflect properties peculiar to BrdUrd-substituted *lac* operator. The mechanism of repressor-operator recognition must be basically the same because the pattern of enhancement and inhibition of purine methylation (17) is identical for thymine and bromouracil-substituted operators. Furthermore, while the affinity of repressor for substituted operator is 10-fold greater than for unsubstituted operator, this increase is also observed when comparing repressor binding to poly[d(A,T)] and poly[d(A-BrdUrd)] (23). Therefore, the increased affinity for BrdUrd operator probably involves changes in minor-groove contacts and is most likely a nonspecific effect.

The failure of the crosslinking at some of the thymines does not prove that the protein does not closely touch those points but only that there is no appropriately reactive amino acid in the protein close enough to that group on the DNA. These experiments thus reveal only some of the points of contact between the protein and the major groove of DNA.

We thank Judy Howard for excellent technical assistance. This work was supported by National Institutes of Health Grant GM21514-3. W.G. is an American Cancer Society Professor of Molecular Biology.

1. Hutchinson, F. (1973) *Qt. Rev. Biophys.* **6**, 201-246.
2. Lin, S. Y. & Riggs, A. D. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 947-951.

3. Weintraub, H. (1973) *Cold Spring Harbor Symp. Quant. Biol.* **38**, 247-256.
4. Lin, S. Y. & Riggs, A. D. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 2574-2576.
5. Platt, T., Files, J. G. & Weber, K. (1973) *J. Biol. Chem.* **248**, 110-121.
6. Sharp, P. A., Sugden, B. & Sambrook, J. (1973) *Biochemistry* **12**, 3055-3063.
7. Gilbert, W., Gralla, J., Majors, J. & Maxam, A. (1975) in *Protein-Ligand Interactions*, eds. Sund, H. & Blauer, G. (Walter de Gruyter and Co., Berlin), pp. 193-210.
8. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560-564.
9. Backman, K., Ptashne, M. & Gilbert, W. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 4174-4178.
10. Tanaka, T. & Weisblum, B. (1975) *J. Bacteriol.* **121**, 354-362.
11. Guerry, P., LeBlanc, D. J. & Falkow, S. (1973) *J. Bacteriol.* **116**, 1064-1066.
12. Lis, J. T. & Schleif, R. (1975) *Nucleic Acids Res.* **2**, 383-389.
13. Loening, U. E. (1967) *Biochem. J.* **102**, 251-257.
14. Hassur, S. M. & Whitlock, H. W. (1974) *Anal. Biochem.* **59**, 162-164.
15. Couch, J. L. & Hanawalt, P. C. (1971) *Anal. Biochem.* **41**, 51-56.
16. Jobe, A., Riggs, A. D. & Bourgeois, S. (1972) *J. Mol. Biol.* **64**, 181-199.
17. Gilbert, W., Maxam, A. & Mirzabekov, A. (1976) in *Control of Ribosome Synthesis*, eds. Kjeldgaard, N. O. & Maaloe, O. (Munksgaard, Copenhagen), pp. 139-148.
18. Maizels, N. M. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 3585-3589.
19. Bahl, C. P., Wu, R., Stawinsky, J. & Narang, S. A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 966-970.
20. Yansura, D. G., Goeddel, D. V., Cribbs, D. L. & Caruthers, M. H. (1977) *Nucleic Acids Res.* **4**, 723-737.
21. Richmond, T. J. & Steitz, T. A. (1976) *J. Mol. Biol.* **103**, 25-38.
22. Kolchinsky, A. M., Mirzabekov, A. D., Gilbert, W. & Li, L. (1976) *Nucleic Acids Res.* **3**, 11-18.
23. Lin, S. Y. & Riggs, A. D. (1971) *Biochem. Biophys. Res. Commun.* **45**, 1542-1547.