J. Mol. Biol. (1979) **132**, 709–728

DNA-binding Site of *lac* Repressor Probed by Dimethylsulfate Methylation of *lac* Operator

RONALD T. OGATA† AND WALTER GILBERT

Department of Biochemistry and Molecular Biology
Harvard University
Cambridge, Mass. 02138, U.S.A.

(Received 17 April 1979)

In order to compare the structures of the DNA-binding sites on variants of the *lac* repressor, we have studied the influence of these variants on the dimethylsulfate methylation of the *lac* operator. Since a bound protein changes the availability of specific purines in the operator to this chemical attack, comparisons of the methylation patterns will show similarities or differences in the protein–DNA contacts. We compared *lac* repressor, induced *lac* repressor (repressor bound to the gratuitous inducer isopropyl-β-D-thiogalactoside), mutant repressors having increased operator affinities (X86, I12 and the X86-I12 double mutant) and repressor peptides (long headpiece, residues 1 to 59 and short headpiece, residues 1 to 51). All of these repressors and repressor peptides exhibit the same general pattern of protection and enhancement in the operator; however, the short headpiece pattern differs most from that of the repressor while the induced repressor and the long headpiece show intermediate patterns that are strikingly similar to each other. The mutant repressors do not show an isopropyl-β-D-thiogalactoside effect but otherwise are almost indistinguishable from wild-type repressor. These results demonstrate that all molecules bind to the operator using basically the same protein–DNA contacts; they imply that (1) most and possibly all repressor contacts to operator lie within amino acids 1 to 51, (2) inducer weakens many contacts rather than totally disrupting one or even a few and (3) the tight-binding mutants do not make additional contacts to the DNA.

These results are consistent with a model in which the amino-terminal portions of two repressor monomers make the DNA contacts. We show that one can understand the affinity of binding as related to the accuracy of the register of the two amino-terminal portions along the DNA. Furthermore, the action of inducer and the behaviour of the tight binding mutants can be accomodated within a two-state model in which the strongly or weakly binding states correspond to structures in which the amino-terminal regions are rigidly or loosely held with respect to each other.

1. Introduction

Dimethylsulfate methylates the purines in double-stranded DNA at the N-7 of guanines and at the N-3 of adenines (for a review, see Singer, 1975). DNA-binding proteins which recognize specific nucleotide sequences can perturb the methylation reaction at purines which lie within these sequences (Gilbert *et al.*, 1976); the effect at each purine is highly characteristic and the pattern of enhancement or inhibition

† Present address: Department of Genetics, Box 8031, Washington University School of Medicine, St Louis, Mo. 63110, U.S.A.

0022 2836/79/240709–20 $02.00/0

of methylation, across the recognition sequence, reflects the pattern of amino acid–nucleotide contacts between the protein and DNA. Thus purine methylation has been used to characterize the DNA sequences which bind to *lac* repressor (Gilbert *et al.*, 1976), λ repressor (Humayan *et al.*, 1977), *cro* protein (Johnson *et al.*, 1978), RNA polymerase (Johnsrud, 1978) and CAP protein (Majors, 1977). Conversely, the same method can probe the DNA-binding site of the protein. Since the methylation reaction monitors the environments of purines in the recognition sequence, structural changes in the DNA-binding site of the protein can appear as changes in the reactivities of those purines. Here we attempt to define the size and structure of the operator-binding site of *lac* repressor by comparing the effects on the methylation of the *lac* operator of repressor variants, including induced repressor, mutant repressors, and repressor peptides which retain the operator-binding activity.

Which parts of the *lac* repressor constitute the operator binding site? Several lines of evidence stress the importance of the amino-terminal region in operator binding (for reviews see Müller-Hill, 1975; Bourgeois & Pfahl, 1976; Miller & Reznikoff, 1978). Geisler & Weber (1977) and Müller-Hill *et al.* (1976) showed that the amino-terminal region has a strong, non-specific affinity for DNA. We recently reported that a small piece of the repressor, long "headpiece" (residues 1 to 59) (Geisler & Weber, 1977), affects methylation specifically in the operator and that the pattern of protection and enhancement exhibited by long headpiece is almost identical to that found with intact repressor. We concluded from these results that the long headpiece binds specifically to the operator, that the mechanism of binding closely resembles that of intact repressor, and that amino acids 1 to 59 form a structural domain which makes up most and possibly all of the repressor's operator-binding site. In the present study, we define further the operator-binding region of repressor by studying the effect on methylation of an even smaller repressor fragment, short headpiece (residues 1 to 51).

2. Materials and Methods

(a) *Operator fragments*

The 55-base-pair long restriction fragment containing the *lac* operator was isolated from a cloned 203-base-pair long fragment (F. Fuller, unpublished results) containing the entire *lac* control region and 5′-end-labeled with ^{32}P as described previously (Ogata & Gilbert, 1977). Partially (~50%) BrdUrd-substituted operator was prepared from cells grown in BrdUrd-containing medium (Ogata & Gilbert, 1977). Normal thymidine-containing operator was prepared by growing the strain in YT medium supplemented with 10 μg thymidine/ml and isolating the plasmid on an ethidium bromide/cesium chloride gradient as described by Tanaka & Weisblum (1975).

(b) *Repressors*

Wild-type and X86 repressors were prepared as described by Platt *et al.* (1973). I12 repressor was prepared as described by Schmitz *et al.* (1978) and the double X86-I12 mutant repressor was a gift from A. Schmitz.

(c) *Headpieces*

Long and short headpieces were prepared as described earlier (Geisler & Weber, 1977; Ogata & Gilbert, 1978), substituting Sephacryl S200 for Sephadex G150: intact repressor was cleaved with trypsin and the resulting headpieces purified by passage through Sephacryl S200 and Sephadex G25. Digestion for 45 min yielded long headpiece (contaminated ~20% with short headpiece) and digestion for 3 h yielded almost pure short headpiece (Ogata &

Gilbert, 1978). I12 long headpiece showed a slightly greater sensitivity to trypsin than wild-type and thus for I12 long headpiece the digestion was shortened to 30 min. Headpiece concentrations were measured spectrophotometrically (Jovin *et al.*, 1977) with an estimated correction (Wetlaufer, 1962) for the additional tyrosine in the I12 mutant.

(d) *Methylation, cleavage and electrophoresis*

Methylation of DNA with dimethylsulfate was carried out as described (Gilbert *et al.*, 1976; Ogata & Gilbert, 1978) at room temperature and on ice for 4 to 5 min and 60 to 90 min, respectively. These reaction times resulted in methylation of fewer than 1 purine/DNA strand. Cleavage of methylated DNA with heat and alkali, electrophoresis on 20% acrylamide/7 M-urea gels and autoradiography were carried out essentially as described by Gilbert *et al.* (1976) and Maxam & Gilbert (1977). Gels measured 38 cm × 15 cm × 0·8 mm and had 0·5-cm sample wells. A standard reaction contained 1 pmol of labeled operator fragment, 1 μg sonicated calf thymus DNA and 20 μg bovine serum albumin in 100 μl.

(e) *Measurement of protection and enhancement*

The level of methylation at each purine is proportional to the intensity of the band in the autoradiogram which corresponds to cleavage at that base. Autoradiograms were scanned with an Ortec model 4310 densitometer to measure the intensity of each band. The level of protection or enhancement at each purine is expressed as the ratio C/P, where C is the intensity of the corresponding band in the control lane (no added protein) and P is the intensity of that band in the presence of repressor or headpiece.

(f) *Headpiece–operator affinities*

Dissociation constants for headpiece–operator interactions were estimated by measuring the level of protection at guanine 5 and guanine 17 in the sequence as a function of headpiece concentration. Because the total amount of methylation is quite low, the ratio of methylation in the absence and presence of protein, C/P, for a particular purine, is approximately equal to the ratio of methylation rates at that purine in the absence (M_o) and presence (M) of protein: $C/P \simeq M_o/M$. These rates are given by

$$M_o(i) = k_i P_o^i \tag{1}$$

and

$$M(i) = k_i P_f^i + k_i' P_b^i, \tag{2}$$

where k_i and k_i' are the rate constants for methylation of purine i (in this case G5 or G17) at free and protein-bound operator sites, P_f^i and P_b^i are the concentrations of purine i in free and protein-bound sites and P_o^i is the total concentration of purine i. We compare methylation in the presence and absence of protein in 2 separate reactions, both containing the same concentration of operator; therefore, $P_o^i = P_f^i + P_b^i$. Rearranging these terms and substituting P/C for M/M_o, we obtain

$$\frac{1}{1 - P/C} = \frac{1}{1 - \lambda(i)} (1 + K_{HP}/[HP]), \tag{3}$$

where K_{HP} is the headpiece–operator dissociation constant and $\lambda(i) = k_i'/k_i$. We plotted $1/(1 - P/C)$ against $1/[HP]$ and approximated the data with a straight line. A typical plot is illustrated in Fig. 1.

3. Results

(a) *The method*

These studies probe contacts between protein and DNA by comparing the extent of methylation with dimethylsulfate in the presence and the absence of protein. The methylation of an individual purine makes the DNA sensitive to scission at the

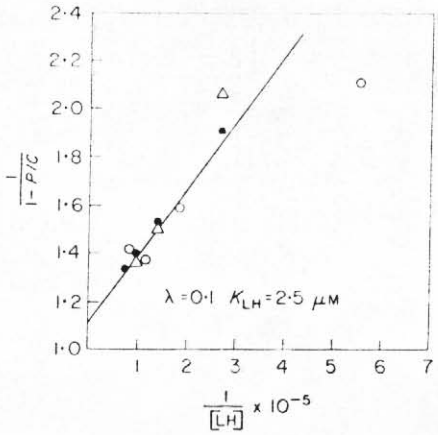

FIG. 1. Plot describing the decrease in methylation at guanine 5 as a function of long headpiece (LH) concentration. P/C is defined in the text. (○,△,●) Results of 3 separate experiments at room temperature and with BrdUrd-substituted operator. The intercept (int) of the line drawn through the points gives λ (G5), the ratio of the methylation rates at G5 in the presence and absence of headpiece ($\lambda = 1 - 1/\text{int}$) and the slope (sl) gives the dissociation constant for the headpiece–operator interaction ($K = \text{sl}(1 - \lambda)$).

modified base. We methylate a double-stranded operator-containing DNA fragment labeled at just one 5′ end, cleave at the methylated bases with heat and alkali, and separate the products on a denaturing polyacrylamide gel. Autoradiography reveals the set of bands illustrated in Figure 2. Because only one 5′ end was labeled, each band corresponds to a single-stranded fragment extending from that 5′ end to a purine in the sequence. Since the length of each fragment determines its mobility, the position of each band is correlated with a specific purine in the sequence. The intensity of a band is proportional to the amount of cleavage at its corresponding purine and thus reflects the level of methylation at that base. The low overall level of methylation (less than one purine per DNA strand) ensures a close relationship between the rate of methylation and the amount of cleavage.

Figure 2 illustrates the low level of methylation (the bulk of the radioactivity remains in uncut strands) and also the variable reactivity of purines to dimethylsulfate even in the absence of DNA-binding protein. In general, the N-7 of guanine, which lies in the major groove of double-helical DNA, reacts about five times faster than the N-3 of adenine, which lies in the minor groove. There is also some neighbor-dependent variability (guanines between thymines and adenines between cytosines appear to be especially reactive). The accuracy of the present results varies somewhat for each purine, decreasing as resolution and reactivity decrease.

(b) *Comparison of repressor, induced repressor and repressor headpieces*

Figure 2 illustrates the effect of *lac* repressor and induced *lac* repressor (repressor bound to the gratuitous inducer isopropyl-β-D-thiogalactoside) on methylation in the 55-base-pair operator-containing restriction fragment. IPTG† weakens the repressor–operator affinity about 1000-fold (Riggs *et al.*, 1970; Barkley *et al.*, 1975). We probe the effect of IPTG by working at concentrations above the dissociation constant of the induced repressor from the operator, where the ternary complex is

† Abbreviations used: IPTG, isopropyl-β-D-thiogalactoside.

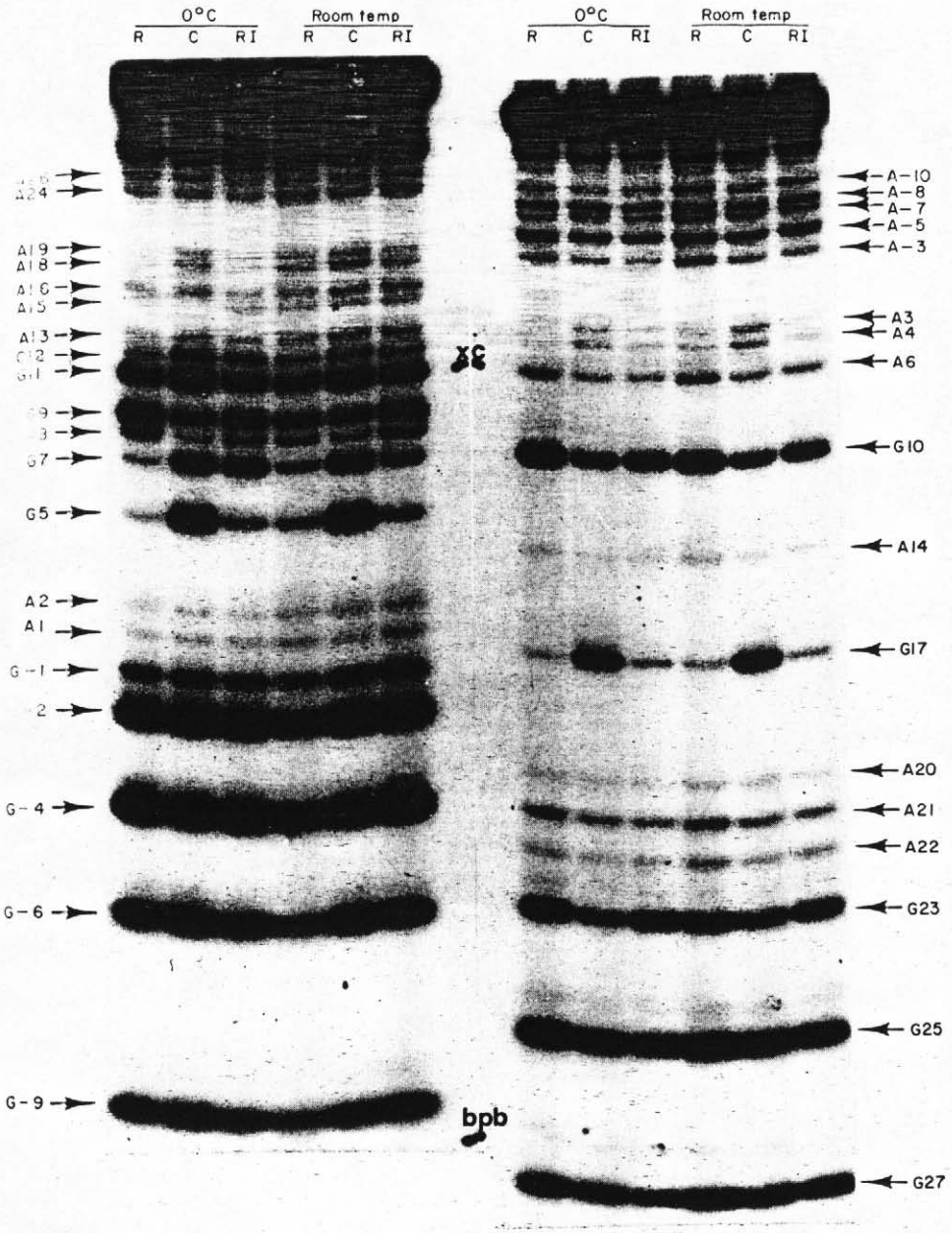

FIG. 2. Autoradiogram of an acrylamide gel showing the effects of 80 nM-repressor (R) and 80 nM-repressor with 2 mM-IPTG (RI) on methylation of the operator-containing restriction fragment (10 nM) at room temperature and 0°C. C is a control sample methylated in the absence of repressor. The nucleotide sequence of the fragment is given at the top of Fig. 4. The 5′ end at the *Hpa*II cutting site is labeled with ^{32}P in the left half of the Figure and the 5′ end at the *Alu*I site is labeled in the right half. Each band represents cleavage at the purine listed on either side of the Figure. Note that cleavage (methylation) is highly purine- and sequence-dependent. xc and bpb mark the positions of xylene cyanol and bromophenol blue on the gel, respectively. The intense, topmost bands represent unmethylated DNA fragments; the bottom strand (labeled at the *Alu*I end) is slightly smaller and thus migrates slightly faster than the top strand.

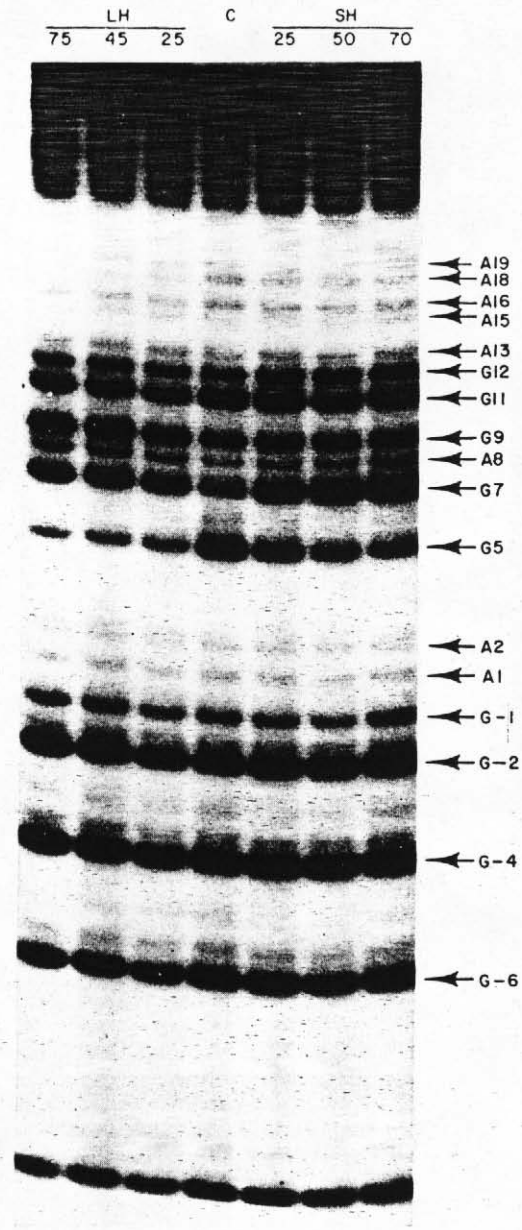

FIG. 3. Autoradiogram of an acrylamide gel showing the effects of long (LH) and short (SH) headpieces on methylation of the operator fragment at 0°C. C is a control sample methylated in the absence of headpiece. The numbers at the top of each lane give the headpiece concentration in μg/ml. Here the 5′ end at the *Hpa*II cutting site is labeled and the bands represent cleavage at purines in the top strand; these are listed to the right of each set of bands.

stable. Figure 3 shows the effects of the two amino-terminal fragments, the two headpieces. Methylation is enhanced at some purines (bands increase in intensity with added protein) and inhibited at others (a decrease in intensity). We quantitated these effects at each purine in the vicinity of the operator by measuring the intensity of each band with a densitometer. We use the ratio (C/P) of band intensities in the absence (C) and presence (P) of repressor or headpieces to describe the effect of protein at that purine in the sequence. These effects were reproducible and the C/P values agreed to within 10 to 20% among several separate experiments.

Figure 4 summarizes our results with repressor (R), induced repressor (RI), long

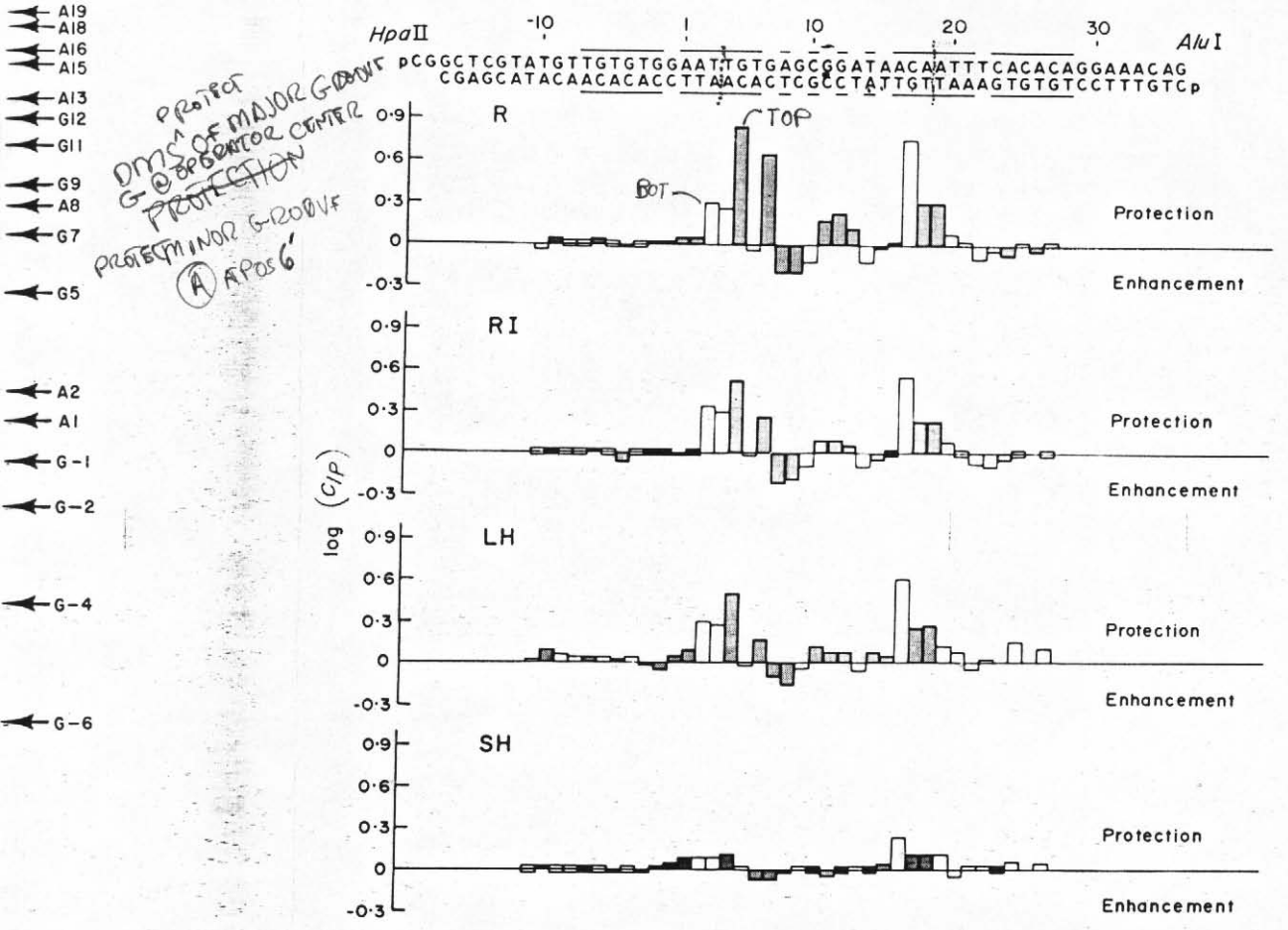

FIG. 4. Summary of the effects of 80 nM-repressor (R), 80 nM-repressor and 2 mM-IPTG (RI), 10 μM-long headpiece (LH) and 12 μM-short headpiece (SH) on methylation at room temperature of the operator-containing restriction fragment (~10 nM). The nucleotide sequence of the fragment is given at the top of the Figure (hyphens omitted for clarity). Each bar is aligned with a purine in the sequence and gives the effect of repressor or headpiece on methylation at that purine. Shaded bars refer to purines in the top strand, and clear bars to purines in the bottom strand. The change in methylation is given as the logarithm of C/P (defined in the text). Methylation is enhanced when $\log (C/P) < 0$, inhibited when $\log (C/P) > 0$ and unaffected when $\log (C/P) = 0$. These data represent the average of 2 to 4 different experiments; C/P values were reproducible to within 10 to 20%.

headpiece (LH, residues 1 to 59) and short headpiece (SH, 1 to 51) at room temperature. Each bar is aligned with a purine in the sequence and gives the value of log (C/P) for that purine. Clearly, induced repressor differs only slightly from repressor: protection is about twofold weaker at guanine 5 (G5), G7 and G17 and about 50% weaker at G12. Otherwise, the methylation does not distinguish between repressor and induced repressor under these conditions. That inducer affects the environments of guanines most strongly suggests that its influence is primarily felt by major groove contacts. In these experiments the repressor and induced repressor concentrations were about fivefold greater than the operator concentration; varying the protein concentration twofold in this range gave the same results, suggesting a total saturation of the operator. Similarly, the effects seen with induced repressor were invariant from 1 to 4 mM-IPTG.

As previously reported, repressor and long headpiece also show very similar effects on methylation (Ogata & Gilbert, 1978). Although the operator is not completely saturated with long headpiece, the experiment in Figure 4 is done at 70% of saturation and the effects are clear. Note that while R and LH are very similar, their differences closely parallel the differences between R and RI. Thus *LH and RI have almost identical effects on methylation*. We show later that this similarity is even more striking with experiments carried out at 0°C. Small headpiece binds poorly to the operator (about 20% saturation under these conditions) and thus only weakly perturbs methylation. Nevertheless, the effect is unmistakable and characteristic of specific binding. SH is unique in that it weakly enhances methylation at G7; this effect is significant, as it also appears with long headpiece and induced repressor under other conditions. Finally, the amount of protection in the middle of the operator, at G11, G12 and adenine 13 (A13), gradually decreases to zero in the order R→RI→LH→SH, suggesting a withdrawal of the protein or peptide from this region of the operator.

Figure 5 summarizes the results of the same experiments carried out at 0°C. With intact repressor, reducing the temperature abolishes the enhancement at A14 and enhances methylation at G23. IPTG has much the same effect on the pattern at both temperatures with one striking exception: the threefold protection at G7 is reduced to no effect at 0°C. Smaller differences are seen at G12 and A13. Again, RI and LH show almost identical patterns. LH, however, shows an even greater change at G7 (methylation is actually *enhanced* here as it was with SH at room temperature) and the protections at G12 and A13 are essentially zero. SH binds more strongly to operator at this temperature and generates a much clearer pattern (we estimate 85% saturation by LH and 60% saturation with SH). It is similar to LH except that the protections are generally weaker and enhancements at G7 and G12 are appreciably stronger. In summary: most contacts are unaffected by the temperature change; intact repressor changes only at A14 and G23, and induced repressor shows similar changes in addition to a large change at G7; and the two headpieces show noticeable temperature effects only at G7, G10 and G23.

Figures 6 and 7 show the protection-enhancement patterns at room temperature and 0°C for DNA substituted about 50% with 5-bromodeoxyuridine (BrdUrd). BrdUrd-substituted operator binds to *lac* repressor about tenfold more tightly than does unsubstituted operator (Lin & Riggs, 1972). Tighter binding is reflected in the methylation patterns, where both protection and enhancement are stronger; this is most evident with small headpiece. However, the patterns are essentially unchanged. Prominent responses to BrdUrd substitution occur at only three purines in the

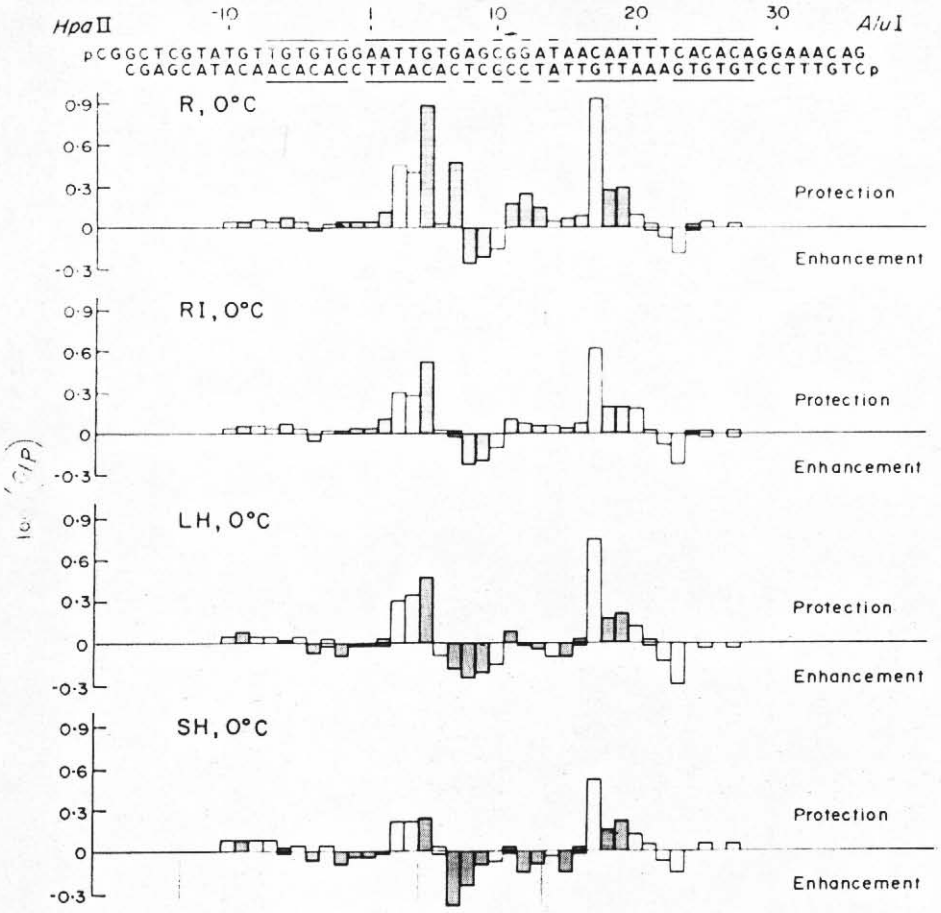

FIG. 5. The same as Fig. 4 except that methylation was carried out at 0°C.

sequence: G7, where *both* RI and LH show temperature-dependent protection-enhancement inversions and at A14 and A22, where enhancement is very pronounced with BrdUrd operator.

What is going on at G7? At this position, repressor always blocks methylation and short headpiece always enhances it, while induced repressor and long headpiece have no effect on, inhibit, or enhance the reaction depending on conditions. We conjecture that a segment of intact repressor covers G7 (protecting it) and that with short headpiece this segment is pulled away a short distance, no longer blocking the site but actually enhancing methylation there by providing a pocket for the dimethylsulfate. Induced repressor and long headpiece represent intermediate states in this scheme. Obviously, the amino acids involved must be those between residues 1 and 51.

(c) *Mutant repressors*

How do amino acid substitutions change the repressor–operator affinity? We examined the methylation patterns of three mutant repressors which exhibit increased

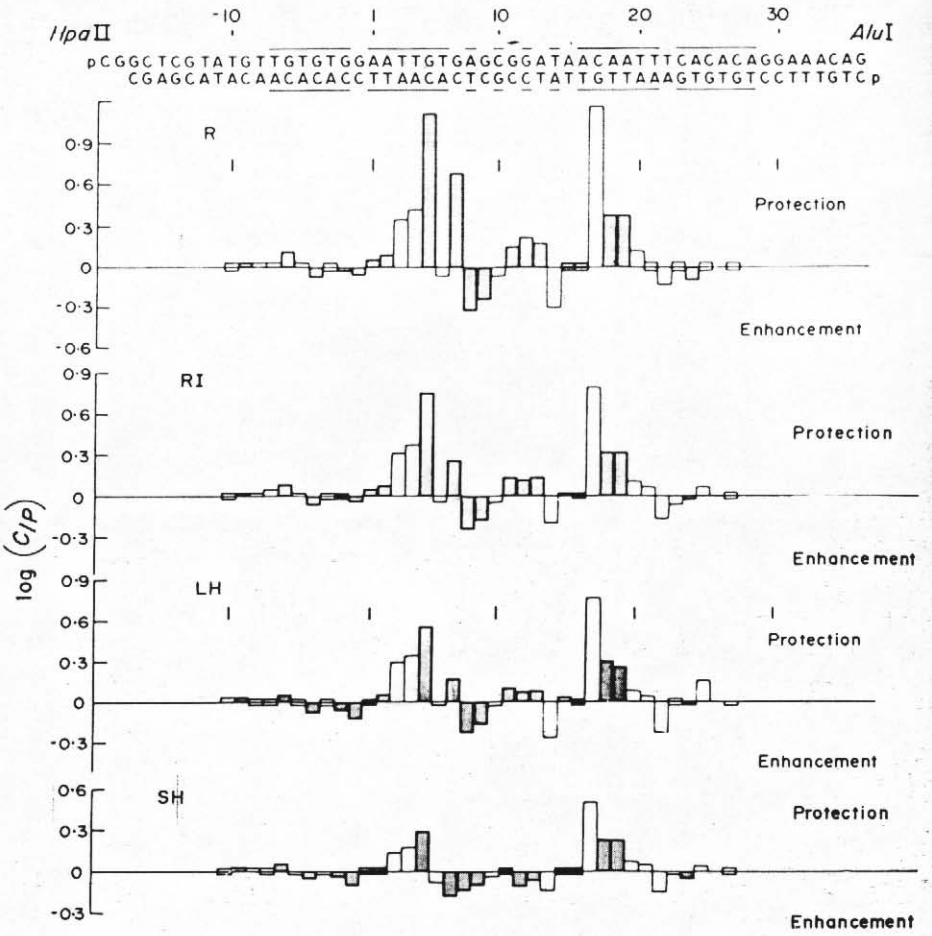

FIG. 6. The same as Fig. 4 except that the operator-containing fragment was ~50% substituted with BrdUrd.

operator affinitities: X86 (Chamness & Willson, 1970), I12 and a double X86-I12 mutant (Schmitz *et al.*, 1978). X86 repressor differs from wild-type repressor by a single serine to leucine change at amino acid 61 (Files & Weber. 1978); it binds normally to IPTG but 50 to 100 times tighter to operator than wild-type repressor (Jobe & Bourgeois, 1972). Like wild-type, X86 binds about 1000-fold less well to operator when complexed with IPTG. I12 repressor has a tyrosine instead of a proline at position 3; its operator- and inducer-binding properties are very similar to those of X86 repressor (Schmitz *et al.*, 1978). The X86-I12 double mutant repressor has both substitutions; it also binds IPTG normally but binds about 10,000-fold tighter to operator than wild-type (Schmitz *et al.*, 1978). IPTG weakens the operator affinity of the double mutant but this effect has not been quantified.

Figure 8 shows the effects of those mutants on methylation at 0°C. Clearly, the patterns are essentially the same as wild-type repressor (though some protections are stronger). They do not even show the relatively small differences seen with induced

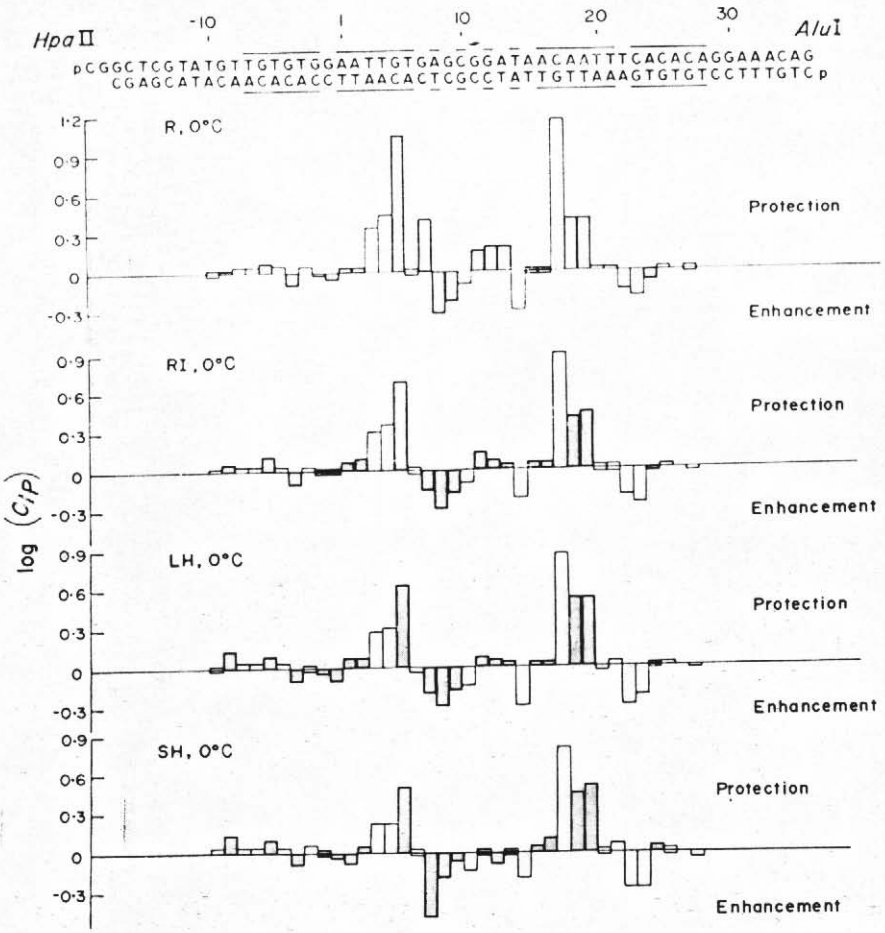

FIG. 7. The same as Fig. 5 except that the operator-containing fragment was ~50% substituted with BrdUrd.

repressor. Methylation at room temperature and with BrdUrd-substituted operators yields the same results. Unlike wild-type repressor, however, the patterns are identical in the absence and presence of IPTG. We do not regard as significant the protections at G-9 and G-6 seen with X86 and I12 repressors. These results are from a single experiment and we attribute those protections to experimental error or to repressor molecules non-specifically bound to the DNA but favoring a position adjacent to the operator-bound protein.

We also examined the effect of I12 short and long headpieces on methylation. Like intact I12 repressor the peptides gave patterns indistinguishable from wild-type headpieces. However, unlike intact I12 repressor, I12 headpieces did not show an increased operator affinity (see the following section).

(d) *Headpiece–operator affinities*

Table 1 lists the dissociation constants describing the binding of long and short headpiece to unsubstituted and BrdUrd-substituted operators, determined by

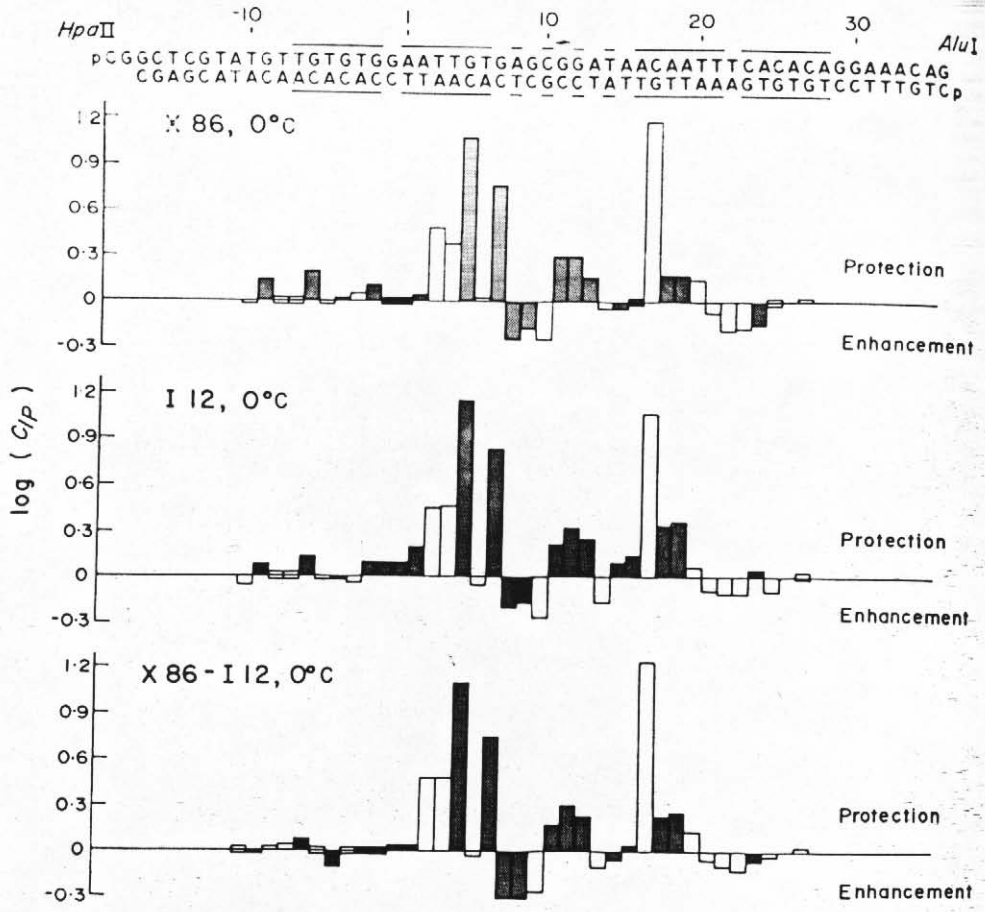

FIG. 8. Summary of the effects of X86, I12 and X86-I12 mutant repressors (all ~50 nM) on methylation at 0°C. Results with X86 and I12 repressors are from a single experiment; the X86-I12 result is the average of 2 separate experiments.

TABLE 1

Dissociation constants (μM) describing headpiece–operator interaction

	dThd		BrdUrd	
	Room temp.	0°C	Room temp.	0°C
LH				
G5	5 (ND)†	1·7 (2)	2·5 (5)	1·5 (ND)
G17	3·5 (ND)	1·4 (2)	1·5 (1·5)	0·8 (ND)
SH				
G5	65 (40)	8·5 (8·5)	7 (14)	3 (4)
G17	25 (12)	3·5 (3·5)	5 (2·5)	2 (2)

† Dissociation constants for I12 headpieces are given in parentheses; ND, not determined.

measuring the extent of methylation at G5 and G17 as a function of headpiece concentration. G5 and G17 were chosen because they show the strongest effects and are symmetrically placed on opposite sides of the operator. Similar measurements at A3, A18 and G9 gave consistent results (arguing against artifacts unique to G5 and G7). These dissociation constants were reproducible to within 20 to 30% (SH affinities are less accurate). Therefore the large differences in affinities of LH and SH for thymidine-containing operator are doubtless significant while other differences are close to experimental limits and therefore suspect. However, some consistent trends are notable: binding is stronger at G17 than at G5, at the lower temperature, and with BrdUrd substitution.

Table 1 also gives the dissociation constants of I12 (Pro3→Tyr) headpieces. Differences between I12 and wild-type are within experimental errors and do not show any definite pattern. Since short and long I12 headpieces also exhibit the wild-type pattern of protection and enhancement, we conclude that I12 and wild-type headpieces make the same contacts to the operator and probably have very similar affinities for operator under these conditions.

4. Discussion

(a) *Purine methylation as a probe of protein structure*

The methylation method probes the local environment of the N-7 of guanine, which lies in the major groove of double-helical DNA, and that of the N-3 of adenine, which lies in the minor groove, by monitoring the reactivities of these positions toward methylation with dimethylsulfate. We interpret the protein-mediated changes in these reactivities as evidence that the protein is very close to, or touches, these specific sites. Although the actual mechanisms of inhibition or enhancement are not known and do not affect this interpretation, for precision, we think of the inhibition as arising from steric hindrance and the enhancement, resulting from the creation of a hydrophobic pocket close to the methylation site, which increases the local concentration of dimethylsulfate. Other mechanisms are certainly possible, so we have not tried to interpret in detail most changes at individual bases; we have no concise explanation for the temperature effects at A14 and G23 or the affects of BrdUrd substitution at A14 and A22. Rather, we look at the patterns as a whole, interpreting similarities or differences in them as evidence of structural similarities or differences in the DNA-binding site of the proteins involved. In comparing two proteins or peptides, we considered their DNA-binding sites to be structurally similar if they had similar patterns at both temperatures and with dThd- and BrdUrd-operators.

Our probe is limited in that methylation can occur at only two sites on the DNA. We do not, therefore, detect changes in the protein which affect phosphate or pyrimidine contacts while leaving contacts around N-7 and N-3 totally unchanged. In addition, the methylation method demonstrates proximity of groups on the protein and not the presence of chemical bonds; therefore, similarities or differences in the *structures* of the binding sites do not demand similarities or differences in the *energies* of the protein–DNA binding reactions.

The fact that methylation probes the structure of the protein–DNA interface and not the energy of the interaction accounts for the apparent disagreement between the present results and earlier studies relating to the size and symmetry of the operator. The present results show clear effects from A3 to G23 in the sequence. Therefore, we

conclude that the repressor makes specific contacts which extend over at least 21 base-pairs. Our previous study of repressor contacts to thymidines in the operator (Ogata & Gilbert, 1977) suggests that repressor also touches thymidines T1 and T2, giving a total length of 23 base-pairs for the binding site. This size agrees with the studies of Goeddel *et al.* (1978), which suggest that the operator spans at least base pairs 1 to 21, but apparently disagrees with the work of Bahl *et al.* (1977), which confines the binding site to base-pairs 3 to 19. However, while Bahl *et al.* demonstrate that most of the binding energy comes from contacts to bases 3 to 19, the lesser involvement of bases outside of this region cannot be ruled out from their experiments. Neither Bahl *et al.* nor Goeddel *et al.* examined contacts beyond position 21.

The present results and those of Goeddel *et al.* (1978) also suggest symmetrical protein contacts to the symmetrically placed bases 1 to 6 and 16 to 21 and possibly to positions 8 and 14. Previous work has argued that the partial 2-fold symmetry in the operator does not play a critical role in repressor binding (Gilbert *et al.*, 1975,1976; Ogata & Gilbert, 1977). Both sets of results are compatible if we interpret the symmetry in the methylation results as evidence only of a symmetrical disposition of the protein on the operator. Data on mutant operators support this interpretation: while protections at the symmetric bases G5 and G17 are almost identical, G5 and G17 are not equally important in repressor binding. Equivalent mutations at these bases produce very different effects. A G→A change at position 5 decreases binding almost 80-fold while the equivalent change at position 17 decreases binding only about 30-fold. These same mutations actually differ by a factor of four in their effects on transcription; and this may more accurately reflect the difference between the two mutations because a secondary pseudo-operator site probably interfered with the direct binding measurements (Gilbert *et al.*, 1975).

(b) *The operator-binding site of* lac *repressor*

(i) *Size*

The clear effect of short headpiece on methylation demonstrates that this peptide binds specifically to the operator and that the binding mechanism is like that of intact repressor. Furthermore, the strong similarity between the methylation patterns of LH and SH and the relatively small differences in their affinities for operator argue against direct contact of amino acids 52 to 59 to the operator. Rather, these results suggest that residues 1 to 51 make all operator contacts and residues 52 to 59 do not touch the DNA but function in maintaining the operator-binding structure. Genetic evidence is consistent with this interpretation: mutations damaging operator binding located within residues 50 to 58 do not affect non-specific DNA binding while mutations earlier in the sequence damage both activities (Müller-Hill, 1975).

(ii) *Effect of inducer binding and amino acid substitutions*

Repressor, induced repressor and the mutant repressors all share essentially the same pattern of methylation protection and enhancement. This demonstrates that repressor molecules differing over seven orders of magnitude in their affinities for operator have very similar structures. The similar patterns of repressor and induced repressor argue in favor of an induction mechanism that weakens many bonds rather than one that totally disrupts one or a few bonds. While the striking change at G7 seems to contradict this interpretation, that change occurs only at 0°C; at room temperature G7 is protected with and without IPTG. Since the 1000-fold effect of

IPTG is essentially the same at both temperatures (Barkley *et al.*, 1975) we conclude that the inversion at G7 is not related to the large change in operator affinity. Instead it apparently reflects a structural change in the binding site which strongly influences the environment of G7 without greatly altering the overall operator affinity. This illustrates the point that methylation probes structure, not bond energies, and also illustrates the sensitivity of the method. Of course we cannot rule out the possibility that IPTG breaks only bonds between the protein and phosphates or pyrimidines in the operator.

The near identity of the results with wild-type, X86, I12 and X86-I12 repressors suggest that these mutations stabilize the DNA-binding conformation rather than provide additional repressor–DNA contacts. Thus the first few residues, as well as residues 52 to at least 61, stabilize the operator-binding form but do not actually touch the DNA. Furthermore, the normal operator affinity of the I12 headpieces implies that the amino terminus interacts with amino acids in the core, beyond residue 59.

(c) *Headpiece–operator binding*

(i) *Stoichiometry and affinity*

Geisler & Weber (1976) and Kania & Brown (1976) have shown that two active repressor subunits per tetramer are sufficient for near-normal repressor–operator binding, suggesting that only two subunits at a time touch the operator. The 2-fold symmetry in the methylation patterns is consistent with this idea; so we expect that two headpieces bind per operator, one on each side of the center of symmetry. Two features of the headpiece–operator affinities support this conclusion. The affinity for the left side of the operator (measured at G5) differs from the affinity for the right side (measured at G17) and the affinities are in the range of the square root of the repressor–operator affinity. We argue that the latter observation implies a 2:1 stoichiometry by reasoning as follows.

That the affinity of a dimer should be about the square of the affinities of the monomers is only approximately true. The enthalpic contributions should add, giving a square term for the affinity, but one must take entropic effects into account separately.

Let us model the binding of a dimer structure by thinking of it as two monomer elements held together by some flexible linker. The linker will constrain one monomer with respect to the other, and we will take this constraint into account by thinking of the second monomer as sweeping out some volume in space, and so being maintained at some effective concentration with respect to some point in that volume.

The interaction between headpiece and operator is defined by

$$K_{HO}[\text{HP-O}] = [\text{HP}][\text{O}], \tag{4}$$

where K_{HO} is the headpiece–operator dissociation constant, and [HP-O], [HP], and [O] are the concentrations of the complex, free headpiece and free operator, respectively. If we link two headpieces, by a flexible linker, without disturbing their other properties, then their binding to operator can be described in two steps: binding of one headpiece to half the operator, described by

$$K_{HO}[\text{DO}'] = 4[\text{D}][\text{O}], \tag{5}$$

where DO′ represents the idealized dimer bound to half the operator (for simplicity

we assume two identical sites for both the dimer and the operator) and followed by binding of the second headpiece to the other half of the operator, described by

$$K_{HO}[DO] = [DO'/C_{eff}. \quad (6)$$

where C_{eff} is the effective concentration of the second headpiece relative to the unoccupied half-operator. This gives an overall binding reaction described by a dissociation constant, K_{DO}, where

$$K_{DO} = K_{HO}^2/4C_{eff}. \quad (7)$$

The C_{eff} term in equation (7) provides an estimate of the entropic correction in the monomer–dimer relationship. Table 2 lists estimates of C_{eff} for various degrees of

TABLE 2

Effective concentrations for a monomer free to move within a sphere of radius r

	r(Å)					
	20	10	3	2	1	0·2
C_{eff}(M)	0·05	0·4	15	50	400	5×10^5
K_{dimer}	$K^2/0{\cdot}2$	$K^2/1{\cdot}6$	$K^2/60$	$K^2/200$	$K^2/1600$	$K^2/2\times10^6$

K refers to the monomer dissociation constant. K_{dimer} is calculated from eqn (7).

constraint between the two monomer elements, and shows that $K_{dimer} \simeq (K_{monomer})^2$ if the second monomer is held in a 10 Å sphere around its binding site once the first has bound. Thus this calculation supports the idea that two repressor monomers, held together by the rest of the protein, make all the DNA contacts. If we try to model the interaction as proceeding *only* through two headpiece–DNA contacts, and assume that the headpiece elements have exactly the same structure on the repressor as when free, then K_{DO} should be the repressor–operator dissociation constant K_{RO}. Using $K_{HO} = 4\times10^{-6}$ M from Table 1 (an average value for long headpiece) and $K_{RO} = 10^{-13}$, C_{eff} would have to be 40 M. This would require that the structure of the protein be such as to hold a second headpiece within 2 Å of its correct position after the first headpiece is bound. This is not unreasonable. (The molecule could be more floppy, if some binding energy is contributed by a core contact, or if the headpiece, resting against the core, is better shaped to bind to the half-operator.)

We conclude that most, if not all, of the repressor–DNA contacts are mediated by two headpieces binding to the operator. Folding a molecular model of the long headpiece in the manner suggested by Chou *et al.* (1975) gives a structure having two relatively rigid components, one α-helical (residues 26 to 45) and the other having predominantly the β-structure (residues 4 to 24), linked by a flexible hinge. Both components have distinct hydrophilic and hydrophobic sides, suggesting that they lie against each other shielding their hydrophobic surfaces. Folding the model in this way gives a globular structure about half the size of the operator.

(ii) *Other properties*

Headpieces and intact repressor show increased binding to BrdUrd-operator but they show inverse responses to temperature: while repressor–operator binding is

about fourfold weaker at 0°C than at room temperature (Riggs *et al.*, 1970). LH and SH bind two- to sevenfold more tightly at 0°C. Therefore the temperature dependence of headpiece–operator binding resembles that of repressor binding to non-specific DNA sequences (Revzin & von Hippel, 1977). The headpieces also appear to differ from an intact repressor in their preference for the right side of the operator. Mutations on the left side of the operator damage repressor binding more than mutations on the right side (Jobe *et al.*, 1974; Gilbert *et al.*, 1975). We have no simple explanation for these differences; however, the temperature dependence of headpiece binding may reflect an instability of the peptide conformation at higher temperatures.

(d) *Possible mechanisms for induction and increased affinities*

How do inducers weaken the interaction between repressor and operator? The similarity of the methylation patterns of repressor and induced repressor argues against an induction mechanism involving the total disruption of one or a few repressor–operator contacts. Instead, IPTG seems to act by perturbing the overall structure of the operator-binding site. What is the nature of this change? The almost total identity of the methylation patterns of long headpiece and induced repressor, especially in their dependence on temperature, suggests a transition to a structure resembling that of long headpiece. This suggests a simple mechanism for induction: the amino-terminal region of *lac* repressor is normally constrained by the rest of the molecule in a high-operator-affinity form. Inducer releases these constraints and the amino terminus then relaxes to a low-affinity form having increased motional freedom and a structure similar to that of long headpiece. In terms of our earlier discussion of headpiece binding, we imagine that inducer simply weakens the linkage between the two binding subunits (headpieces, represent the completely unlinked case) such that C_{eff} decreases to about 0·06 M, equivalent to holding the unbound repressor subunit within 19 Å of the unbound half-operator. (Alternatively, but within the same motional framework, we might imagine the headpieces held rigidly (±2 Å) but moved out of register, in some Gaussian fashion, by some 8 to 10 Å.)

Extending the same analysis to the tight-binding mutants gives C_{eff} equal to $\times 10^3$ M for X86 and I12 and 4×10^5 M for the double mutant. These are equivalent to holding the second repressor subunit within 0·5 and 0·1 Å of the binding site, respectively. Effective concentrations this high are probably unrealistic and illustrate the limitations of this simple analysis.

How *can* wild-type repressor and the three tight-binding mutants, which differ by up to 10,000 in their affinities for operator, have the same structure, while induced repressor, which binds *lac* operator only about 1000-fold weaker than wild-type repressor, has an altered structure? The basic two-state model of Monod *et al.* (1965) offers a simple explanation of this behavior.

Assume that *lac* repressor exists in a strong-binding (t) and weak-binding (r) form. Let K_t and K_r be operator dissociation constants for the t and r states; and let L describe the equilibrium between the two states: $L = [R_t]/[R_r]$, where $[R_t]$ and $[R_r]$ are the concentrations of repressor in the t and r states. Since the methylation reaction senses the structure bound to the operator our results reflect the ratio of bound R_t and R_r structures ($[R_tO]/[R_rO]$). This is given by

$$\frac{[R_tO]}{[R_rO]} = Lc, \tag{8}$$

where $c = K_r/K_t$. With this relationship we can predict the appearance of the methylation pattern from the parameters of the model. When $Lc \gg 1$, the pattern is characteristic of the t state and when $Lc \ll 1$, the pattern reflects the r state.

Model parameters which are consistent with experimental data can be selected as follows. K_r and K_t are invariant, since we assume only two structures. Therefore differences in the operator dissociation constants of the repressor types must arise only from perturbations in the equilibrium between the two states. This is described by

$$K_{obs} = K_r \frac{L + 1}{Lc + 1}, \tag{9}$$

where K_{obs} is the observed operator dissociation constant.

If we assume that the X86-I12 mutant represents the strongest binding repressor possible, then the operator dissociation constant for this mutant is equal to K_t. And if we assume that the induced repressor represents the weak binding form, then K_r is the dissociation constant for the binding of induced, wild-type repressor to operator. This gives:

$$K_{X86\text{-}I12} = K_t = 10^{-17} \text{ M}$$

$$K_{RI} = K_r = 10^{-10} \text{ M}$$

and

$$c = 10^7.$$

With equation (2) we can calculate L for wild-type, X86 and I12 repressors from the known operator dissociation constants for these proteins: K_{obs} (wild-type) $= 10^{-13}$ M, giving $L_{wt} = 10^{-4}$; $K_{obs}(X86) = K_{obs}(I12) = 10^{-15}$ M, giving $L_{X86} = L_{I12} = 10^{-2}$. With these parameters we calculate the ratio of $[R_tO]$ to $[R_rO]$ for wild-type repressor

$$L_{wt}c = 10^3$$

and for the two mutants

$$L_{X86}c = L_{I12}c = 10^5.$$

For both wild-type and the two mutants, the model predicts that the methylation pattern should predominantly reflect the same (t) form; this is consistent with our results. However, Lc becomes much smaller than one as K_{obs} approaches 10^{-10} M. For example, when $K_{obs} = 9 \times 10^{-11}$ M, $L = 10^8$ and $Lc = 0{\cdot}1$ (or 90% the r form). Since we arbitrarily set K_{obs} (induced) $= K_r$, induced repressor exhibits the r state. Thus the simple two-state model fully accounts for our results. These parameters predict that wild-type repressor and the tight-binding mutants X86 and I12 normally exist almost totally in the r form. The greater operator affinity of the t form drives this equilibrium to favor that state when repressor is bound to operator.

Why does IPTG affect only the methylation pattern of wild-type repressor? If we assume that the effect of IPTG is limited to decreasing L by a constant factor, the results with wild-type repressor require that inducer decrease L by at least 10^4. For this minimum effect, the model predicts that in the presence of inducer $[R_tO]/[R_rO] = 10$ for X86 and I12 and much more for the double mutant. Again, this agrees with our observations. Increasing the effect of inducer on L leads of course to a greater proportion of the r state, and thus our experimental results do not allow a change in L much larger than 10^4.

This two-state analysis is doubtless oversimplified; nevertheless it demonstrates that our results can easily be explained in terms of a physical model. The two-state

model and the analysis in terms of C_{eff} are not necessarily contradictory if we make the not unreasonable assumption that the experimentally determined dissociation constant for binding of headpieces to operator, K_{HO}, *underestimates* the affinity of a repressor subunit for the operator. If we assume that K_{HO} is 100-fold higher than the actual subunit dissociation constant, i.e. we assume that the subunit constrains the headpiece so that it binds more precisely, more favorably, to a half-operator, then we calculate from equation (7) that when $C_{eff} = 40$ M, $K_{DO} = 10^{-17}$ M, which is equal to $K_{X86\text{-}I12}$. We can then assign the t form in the two-state model to a structure in which $C_{eff} = 40$ M, the r form to one in which $C_{eff} = 0{\cdot}06$ M and assume an equilibrium between the two forms described by L.

Thus we view the action of inducer, and mutations to tight-binding repressors as changing the balance between two states, from one in which the critical elements on two subunits are held rigidly, matching the DNA structure to within a few ångström units, to the other in which the amino-terminal regions are held loosely or out of register.

We thank Albert Schmitz for his gift of the X86-I12 mutant repressor and Judith Howard for expert technical assistance. This work was supported by National Institutes of Health grant GM 21514-05. One of the authors (W.G.) is an American Cancer Society Professor of Molecular Biology.

REFERENCES

Bahl, C. P., Wu, R., Stawinsky, J. & Narang, S. A. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 966–970.
Barkley, M. D., Riggs, A. D., Jobe, A. & Bourgeois, S. (1975). *Biochemistry*, **14**, 1700–1712.
Bourgeois, S. & Pfahl, M. (1976). *Advan. Protein Chem.* **30**, 1–99.
Chamness, G. C. & Willson, C. D. (1970). *J. Mol. Biol.* **53**, 561–565.
Chou, P. Y., Adler, A. J. & Fasman, G. D. (1975). *J. Mol. Biol.* **96**, 29–45.
Files, J. G. & Weber, K. (1978). *J. Mol. Biol.* **123**, 454–456.
Geisler, N. & Weber, K. (1976). *Proc. Nat. Acad. Sci., U.S.A.* **73**, 3103–3106.
Geisler, N. & Weber, K. (1977). *Biochemistry*, **16**, 938–943.
Gilbert, W., Gralla, J., Majors, J. & Maxam, A. (1975). In *Protein–Ligand Interactions* (Sund, H. & Blauer, G., eds), pp. 193–210, Walter de Gruyter and Co., Berlin.
Gilbert, W., Maxam, A. & Mirzabekov, A. (1976). In *Control of Ribosome Synthesis* (Kjeldgaard, N. C. & Maaløe, O., eds), pp. 139–148, Munksgaard, Copenhagen.
Goeddel, D. V., Yansura, D. G., Winston, C. & Carruthers, M. H. (1978). *J. Mol. Biol.* **123**, 661–687.
Gumayan, Z., Kleid, D. & Ptashne, M. (1977). *Nucl. Acids Res.* **4**, 1595–1607.
Jobe, A. & Bourgeois, S. (1972). *J. Mol. Biol.* **72**, 139–152.
Jobe, A., Sadler, J. R. & Bourgeois, S. (1974). *J. Mol. Biol.* **85**, 231–248.
Johnson, A., Meyer, B. J. & Ptashne, M. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 1783–1787.
Johnsrud, L. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 5314–5318.
Jovin, T. M., Geisler, N. & Weber, K. (1977). *Nature (London)*, **269**, 668–672.
Kania, J. & Brown, D. T. (1976). *Proc. Nat. Acad. Sci., U.S.A.* **73**, 3529–3533.
Lin, S.-Y. & Riggs, A. D. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 2574–2576.
Majors, J. (1977). Dissertation, Harvard University, Cambridge, Mass.
Maxam, A. M. & Gilbert, W. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 560–564.
Monod, J., Wyman, J. & Changeux, J.-P. (1965). *J. Mol. Biol.* **12**, 84–118.
Miller, J. H. & Reznikoff, W. S. (1978). *The Operon*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
Müller-Hill, B. (1975). *Progr. Biophys. Mol. Biol.* **30**, 227–252.

Müller-Hill, B., Heidecker, G. & Kania, J. (1976). In *Proc. Third John Innes Symp: Structure Function Relationships of Proteins* (Markham, R. & Horne, R. W., eds), pp. 167–179. North-Holland, New York.
Ogata, R. T. & Gilbert, W. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 4973–4976.
Ogata, R. T. & Gilbert, W. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 5851–5854.
Platt, T., Files, J. G. & Weber, K. (1973). *J. Biol. Chem.* **248**, 110–121.
Revzin, A. & von Hippel, P. H. (1977). *Biochemistry*, **16**, 4769–4776.
Riggs, A. D., Bourgeois, S. & Cohn, M. (1970). *J. Mol. Biol.* **53**, 401–417.
Schmitz, A., Coulondre, C. & Miller, J. H. (1978). *J. Mol. Biol.* **123**, 431–456.
Singer, B. (1975). *Progr. Nucl. Acid Res. Mol. Biol.* **15**, 219–284.
Tanaka, T. & Weisblum, B. (1975). *J. Bacteriol.* **121**, 354–362.
Wetlaufer, D. B. (1962). *Advan. Protein Chem.* **17**, 303–390.