THE LAC OPERATOR IS DNA

By Walter Gilbert and Benno Müller-Hill

Departments of Physics and Biology, Harvard University

Communicated by J. D. Watson, October 25, 1967

How repressors act at the molecular level to turn off genes is only now beginning to be worked out. Most vital to this understanding is whether the operator, defined genetically as the site for the action of a repressor, would turn out to be part of a DNA molecule, a region of a messenger RNA molecule, or even a protein. Now that two specific repressors (lactose and λ) are available,[1, 2] it is possible to attack this problem directly. This was first done by Ptashne,[3] who showed that the λ phage repressor, a 30,000-mol-wt protein, binds specifically only to that region of a λ-DNA molecule where the genetic receptors (operators) lie. Here we report experiments, with the lactose repressor, that further show that the operator is DNA. This repressor binds specifically to DNA molecules that carry the lactose operon, attaching only to that unique region of the DNA molecule where the mutations that characterize the operator lie. Furthermore, this repressor is *released* from the operator by inducers, such as IPTG (isopropyl-1-thio-β-D-galactoside).

The Principle of the Experiment.—The assay for the *lac* repressor used the fact that this repressor could bind radioactive IPTG tightly enough to be detected by equilibrium dialysis. Since the relevant affinity is on the order of 10^{-6} M, only repressor concentrations in this range are detectable. This assay cannot be used immediately to study the interaction of the repressor with the operator because attainable gene concentrations are so small. Even if one uses *lac* genes carried on the DNA isolated from a defective phage, one set of genes for each 3×10^7 mol wt, a 3 mg/ml solution of DNA is only 10^{-7} M. The binding of repressor to such DNA would only be barely visible by the IPTG binding assay. An alternative approach is to prepare radioactive repressor, to follow the molecule directly. The IPTG binding assay has been used to guide a several thousandfold purification of unlabeled repressor. With this knowledge one could try to mimic this purification on a small scale with very highly labeled proteins—a blind purification, since the specific labeling is so high and the physical scale of the radioactive preparation so small that one cannot follow the purification by the IPTG binding assay. A complete purification is unnecessary; all that is required is a sufficient enrichment of the *lac* repressor so that it represents a reasonable fraction of the labeled material, while other proteins that bind to DNA are removed so that specific effects can be observed. By including a sizing step, isolating only 7–8S material that includes the *lac* repressor, one can easily distinguish later a small fraction of the label binding to and sedimenting with 35S *dlac* phage DNA.

The details of the purification are given in the experimental methods. Sulfur-labeled proteins from a triploid strain, carrying three copies of the *lac* genes, are fractionated with ammonium sulfate and then run on a DEAE Sephadex column using a step elution. The material is then concentrated and run upon a glycerol gradient. Since the repressor, as determined by its binding to IPTG, sediments near 7.6S, samples are taken from this region of the gradient, determined by an aldolase marker. When this radioactive material is mixed with phage DNA carry-

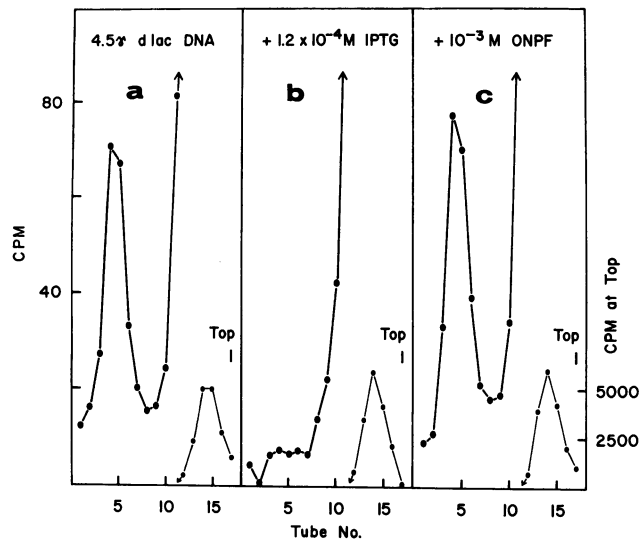

FIG. 1.—The binding of the *lac* repressor to *dlac* phage DNA and its release by inducer. Three identical mixtures of 4.5 γ of *dlac* phage DNA and 8*S* radioactive protein were sedimented on glycerol gradients containing TMEM buffer and run for 2 hr and 20 min at 65,000 rpm as described in the *Experimental Details*. The DNA by itself would form a sharp peak at tubes 4 and 5. The left panel (*a*) shows that a distinct peak of label sticks to the DNA and sediments down the gradient. The DNA is in at least tenfold excess; all the radioactivity that can bind to DNA at this concentration has bound to it. The center panel (*b*) shows that if the gradient solution contains 1.2×10^{-4} *M* IPTG, this binding is abolished. The right-hand panel (*c*) shows that if the gradient contains 10^{-3} *M* ONPF, there is no effect on the binding.

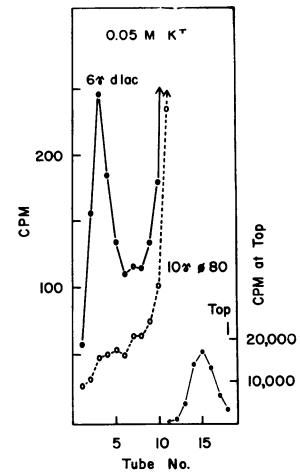

FIG. 2.—The specificity of the binding. In parallel gradients the same repressor preparation was run with two different DNA's. One reaction mixture contained 6 γ of pure *dlac* phage DNA; the other contained 10 γ of the parental ϕ80–λ hybrid. The gradients contained 0.05 *M* KCl in TMEM.

ing the *lac* region, the mixture incubated and then sedimented on a glycerol gradient, a small peak of radioactivity moves out of the 8*S* region and sediments with the DNA around 35 to 40*S*. Figure 1*a* shows such a gradient pattern in 0.01 *M* Mg^{++}. Only 1 per cent of the radioactivity moves with the DNA, even though the DNA is in excess. That the label binding to DNA represents the *lac* repressor, and not just sticky proteins, is shown by the material being released by IPTG. Figure 1*b* shows that if 10^{-4} *M* IPTG is put throughout the gradient, no binding is observed. This effect of IPTG is specific: ONPF (*o*-nitrophenyl-β-D-fucoside), a substance which binds to the repressor but does not induce, has no effect. Figure 1*c* shows that even 10^{-3} *M* ONPF does not interfere with the binding.

RNase has no effect on this binding. When 75 γ/ml of RNase was added to the binding mixture, for a 20-minute incubation at 30°C, no effect on the binding to DNA was observed. Unlabeled, purified *lac* repressor competes for this binding.

The Specificity of the Binding to DNA.—If the repressor interacts with the operator region, the repressor should bind only to DNA carrying the lactose operon itself and, specifically, only to that region at the beginning of the lactose operon which is characterized through mutations as the genetic operator. In fact, no binding is found with phage DNA not carrying the *lac* genes. Figure 2 shows such an experiment in 0.05 *M* KCl. Furthermore, one can ask the more specific question, Does the repressor bind to the operator region by using operator-constitutive (o^c) mu-

tants carried on the phage DNA? We have examined two such mutants: one has a level of enzyme activity in the absence of inducer 20 per cent of that attainable in the presence of inducer; the other, an extremely low-level constitutive, has an enzyme level in the absence of inducer only 1 per cent of the full level. If the active operator region itself is a region of the DNA molecule, the affinity of the repressor for DNA would be changed in both of these o^c mutants. Since the basal level of enzyme is only 0.1 per cent of the fully induced level, the affinity should be at least a factor of 10 weaker for the 1 per cent o^c and a factor of 200 weaker for the 20 per cent o^c.

When DNA isolated from purified defective phages carrying these o^c mutations is used in the experiment, one observes the patterns shown in Figure 3. Figure 3*a* shows the control binding of the radioactive repressor to wild-type DNA, while Figure 3*b* shows the binding to the 20 per cent o^c. No peak is visible, but some radioactive material has been pulled down from the top of the gradient. The residual affinity of the repressor for the DNA is still detectable. Figure 3*c* shows the affinity of the repressor for the 1 per cent o^c. In this case, still more label moves down the gradient, but the affinity of the repressor for this mutant DNA is less than the affinity for the wild-type DNA.

These experiments demonstrate that the repressor binds to a unique sequence on this DNA molecule, the operator region. Furthermore, all attempts to demonstrate binding to denatured DNA have failed. One infers that the binding is to double-stranded DNA.

The Magnitude of the Binding Constants.—One can obtain rough estimates for the affinity of the repressor for the operator region by observing the shape of the peaks riding on the DNA. These experiments have all been done with an excess of DNA, but the DNA concentration falls as the band moves down the gradient, dropping by a factor of 6 from its initial value in the reaction mixture to its final value when the gradient is collected. If the DNA is run separately in a parallel gradient, the recovery is about 70 per cent and the peak concentration, with 4.5 μg of DNA as an input, is only 2.5 μg/ml (only 8×10^{-11} M). From the sharpness of the peaks shown in Figure 1, one would infer that the DNA concentration must be at least a factor of 10 higher than the binding constant. Since a peak of similar sharpness is obtained when only 0.9 μg of DNA is used, one would estimate that the affinity is on the order of 2×10^{-12} M in 0.01 M Mg^{++}.

The tightness of the binding is influenced by the salt concentration. As Ptashne has observed for the λ repressor, when the salt concentration rises, the affinity for the DNA weakens. Since the DNA concentrations that are used are close to the affinity constants, dissociation plays a role, and a slight change in the salt alters the experimental picture. The profile shown in Figure 2, taken in 0.05 M KCl, can be interpreted as showing that 20 per cent of the bound material trails immediately behind the DNA peak due to a weakened affinity. At 0.15 M KCl, the binding to our standard amount of DNA has been essentially abolished. Figure 4 shows, however, that the binding can be easily observed again by raising the DNA concentration a factor of 4. Table 1 collects estimates for the binding constants of the wild type and of operator-constitutive mutants.

What affinities does one expect *in vivo*? The affinity of the repressor for the operator can be estimated from the magnitude of the basal level of enzyme synthe-

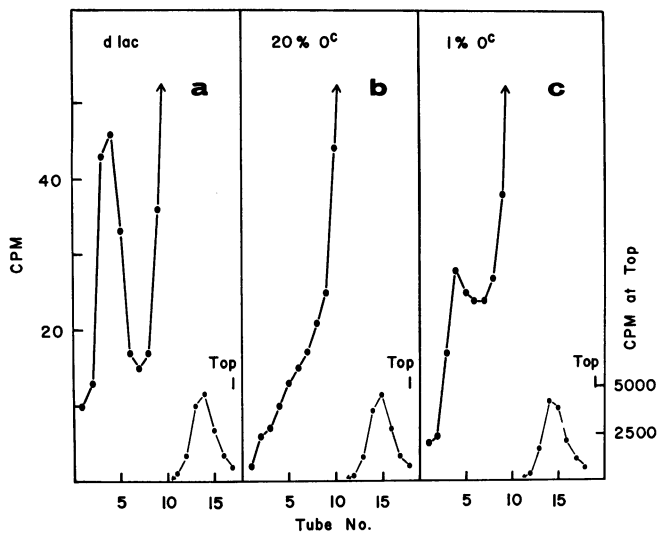

FIG. 3.—The *lac* repressor binds specifically to the *lac* operator. In parallel gradients the same repressor preparation was run with three different purified *dlac* phage DNA's in TMEM. The left-hand panel (*a*) shows the profile obtained with 4.5 γ of wild-type *dlac* DNA. The center panel (*b*) shows that the same amount of DNA carrying an o^c mutation that produces 20% of the full level of enzyme does not bind repressor. The third panel (*c*) shows the binding to an o^c that produces 1% of the full amount (10 times the basal level).

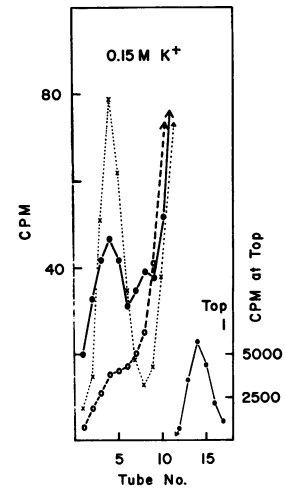

FIG. 4.—The effect of salt. The three superimposed profiles correspond to three parallel gradients: (1) repressor bound to 4.5 γ of *dlac* DNA run in 0.15 *M* KCl and TMEM, O---O---O; (2) repressor bound to 18 γ of *dlac* DNA run in 0.15 *M* KCl and TMEM, ●—●—●; and (3) a control of 4.5 γ of *dlac* DNA run in TMEM, x···x···x.

sis. To the extent that the rate of enzyme synthesis is simply proportional to the amount of DNA free of repressor and obeys mass action kinetics,[4] the basal rate's being only one thousandth of the full rate means that the repressor concentration is about one thousand times the dissociation constant for the operator. On the basis of the isolation procedures, our current estimates are of the order of 10–20 repressor molecules per haploid cell. Thus the repressor concentration in the cell is of the order of $1\text{–}2 \times 10^{-8}$ *M*, and the affinity of the repressor for the operator should be $1\text{–}2 \times 10^{-11}$ *M*. One does not know how to duplicate the ionic conditions within the cell, nor does one know that the interpretation of the basal level as simply determined by the repressor concentration is true. Furthermore, the *in vitro* estimates are made at high pressure on a gradient. Nonetheless, the estimates *in vitro* for the affinity, ranging from 2×10^{-12} *M* in low salt to several times 10^{-10} *M* in higher salt, are in reasonable agreement with the *in vivo* estimate. The weakening of the binding with the o^c's is in the right direction and consistent, roughly, with the difference of the basal rates of synthesis of these two mutations.

TABLE 1

ESTIMATED REPRESSOR-OPERATOR DISSOCIATION CONSTANTS *in vitro*

DNA	0.01 *M* Mg^{++}	0.05 *M* K^+ + 0.01 *M* Mg^{++}	0.15 *M* K^+ + 0.01 *M* Mg^{++}
dlac o^+	$2\text{–}4 \times 10^{-12}$ *M*	2×10^{-11} *M*	3×10^{-10} *M*
1% o^c	10^{-10} *M*	$>10^{-9}$ *M*	
20% o^c	4×10^{-10} *M*	$>10^{-9}$ *M*	

Is the magnitude of this binding physically reasonable? A binding of the order of 10^{-11} M requires some 15 or 16 kcal of binding energy. This energy could arise through the formation of many weak or four to five moderately strong bonds. The repressor must be able to recognize a stretch of at least 11 to 12 bases to select a unique site on the *E. coli* chromosome. (A 12-base sequence selects one out of 1.6×10^7 possible locations, while there are about 3×10^6 base pairs in the chromosome.) To recognize this number of bases individually would require at least 11 or 12 bonds, and thus a free energy change easily in the 15-kcal range. The recognition region will span a considerable distance along the DNA molecule, at least one turn of the helix, a 35-Å stretch; however, the *lac* repressor, 150,000 mol wt, is large enough. If the recognition were not as efficient as possible, the region would be larger.

These dissociation constants imply that the repressor takes hours to fall off the operator. Because the forward rate of formation of the operator-repressor complex will be limited by diffusion and steric factors to be only 10^8 M^{-1} sec^{-1} at the most, a dissociation constant in the 10^{-11} to 10^{-12} M range requires that the rate of decay of the complex be 10^{-3} to 10^{-4} sec^{-1}. How then is it possible that enzyme synthesis begins only minutes after the addition of inducer? Clearly the inducer will bind to the repressor-DNA complex in times that are short compared to 10^4 seconds and trigger the release of the repressor from the operator. Since the rate of release will depend on the amount of inducer, one expects a lag in enzyme induction at low levels of inducer. This lag can be estimated on very general principles and shown to depend only on the slow decay rate of the repressor-operator complex (k_s) and on the ratio of the final rate of enzyme synthesis to the basal rate (r_E/r_B). The argument given in the appendix shows that the time delay (T_c) in the induction curve can be written

$$T_c = (1/k_s)(r_B/r_E) + \text{a constant.}$$

Such time delays have been observed for the induction of the lactose enzymes by Boezi and Cowie.[5] Their data fit this formula with a k_s of 2.2×10^{-4} sec^{-1}, an entirely independent estimate for the decay time *in vivo*.

Summary.—The experiments reported here demonstrate that the *lac* repressor binds specifically to the operator region, that its binding to the operator is weakened by mutations in that region which produce o^c's, and that it is released from the operator by the inducer. These experiments completely support the model of repression which proposes that the repressor, on binding to the operator, hinders the transcription of the adjacent genes into RNA and thus prevents their functioning.

Experimental Details.—Buffers and general methods: TMEM buffer is 0.01 M tris, pH 7.4, 0.01 M magnesium acetate, 10^{-4} M ethylenediaminetetraacetate (EDTA), and 0.007 M β-mercaptoethanol. TMS buffer is TMEM with 0.2 M KCl. All tubes and centrifuge tubes were boiled in EDTA. Other methods were described previously.[1]

Labeled repressor: The bacterial strain carried three sets of *lac* genes, one on $\phi 80$ *dlac* carried as a single defective lysogen at the $\phi 80$ attachment site, one at the normal *lac* site, and one on an F *lac* episome. For the sulfur labeling the cells were grown in minimal medium: 0.1 M potassium phosphate buffer, pH 7.4, 2 gm/liter NH_4Cl, 3 gm/liter NaCl, 2×10^{-4} M Mg^{++}, and 10^{-4} M sulfate, to glycerol starvation at a few times 10^8/ml. They were diluted back and grown to sulfur starvation in 6×10^{-5} M radioactive sulfate. Ten ml of cells were labeled with 20 mc of S^{35}, harvested, and diluted with 1 gm of unlabeled cells. The cells were ground with alumina, the extract suspended in TMS buffer to a final volume of 5 ml, and the debris spun out. The extract was brought to 35% saturation with solid ammonium sulfate, and the precipitate collected and back-

extracted with 2 × 1-ml portions of TMS buffer at 28% and 23% of saturation with ammonium sulfate. The two 23% extractions were pooled, and the ammonium sulfate dialyzed out against TMEM containing 0.1 *M* KCl. The sample was applied to a 2-ml DEAE Sephadex column in the same buffer, and a cut was taken between 0.12 *M* and 0.17 *M* KCl in TMEM. Column fractions (0.5 ml) were collected into tubes containing 0.1 mg of aldolase to provide protective protein and a marker during the centrifugation. The material was concentrated by drying down a dialysis sac with G200 Sephadex, layered on a 5–30% glycerol gradient in TMS containing 0.1 mg/ml BSA, and centrifuged for 16 hr at 45,000 rpm, 4°C. The aldolase marker was located by optical density, and samples from the tubes at and immediately following the aldolase peak were used for the binding experiment.

Phage DNA: The *lac* DNA that was used was isolated from a defective *lac* phage made by V. Rybtchine and Ethan Signer. The phage is derived from a ϕ80–λ hybrid,[6] $h_{80}i^{\lambda}$, contains the c_I857 temperature-inducible mutation, and carries the *lac* genes as a replacement of late phage functions. The o^c mutants used were also made in this phage by Signer. They all have a functioning *i* gene. The defective phages are 0.005 gm/cc denser than the parental hybrid. A double lysogen was grown at 34°C in a glucose casein–amino acid medium buffered with 0.1 *M* tris, pH 7.5, and containing more magnesium ion than phosphate. At 5–8 × 10^8 cells/ml the culture was heat-shocked to 42°C for 15 min, chilled to 37°C, and shaken at 37°C until lysis. The pH was maintained at 7.5. Titers were 1–2 × 10^{11}/ml, after chloroforming. The phage were harvested in the Spinco 30 head, purified on a block CsCl gradient, and then banded in an equilibrium CsCl gradient overnight in the 40 head. DNA was prepared from the purified defective phage by rolling the phage stocks with phenol. The DNA was dialyzed overnight against 0.01 *M* tris, pH 7.4, 0.05 *M* KCl, and 10^{-4} *M* EDTA, and stored at about 100 γ/ml at 4°C.

DNA binding assay: The DNA, handled in 0.1-ml pipettes, was heated to 70°C for 5 min and chilled quickly, to break aggregates. The reaction mixture contained, in a final volume of 0.25–0.3 ml, TMEM buffer with 0.05 *M* KCl, 150 γ of BSA, generally about 5 γ of DNA, and radioactive repressor. After a 20-min incubation at 30°C, the mixture was layered on a 5-ml, 5–30% glycerol gradient containing 500 γ/ml BSA and the specified buffer, either TMEM or TMEM with KCl. After a 2-hr 20-min spin at 65,000 rpm, 8°C, 4-drop samples were collected into scintillation vials and counted with Bray's solution.

Mathematical Appendix.—One can calculate how rapidly the repressor will be driven off the operator by the inducer in a variety of specific models for the induction process. The results are identical so we shall give a general, rough argument. The rate of change of the amount of the repressor-operator complex ([*D-R*]) after the addition of inducer will be given by

$$\frac{d}{dt}[D\text{-}R] = k_s[D\text{-}R] + k_f[D\text{-}R\text{-}I],$$

where [*D-R-I*] is the amount of the repressor-inducer-operator complex and k_s and k_f are the slow and fast decay rates of the two complexes, respectively. One expects k_f to be at least 10^4 times larger than k_s. Furthermore, the equilibration of the inducer, a small molecule, with the various complexes should be very rapid. As we shall see, [*D-R-I*] itself is in general negligible compared to [*D-R*]. One does not know directly what the affinity of the inducer for the DNA-repressor complex would be, but each of these complexes has a dissociation constant

$$\frac{[D][R]}{[D\text{-}R]} = K_0 \quad \text{and} \quad \frac{[D][R\text{-}I]}{[D\text{-}R\text{-}I]} = \tilde{K}_0.$$

Each of these dissociation constants is the ratio of a decay rate to an association rate. To the extent that the shape of the repressor is not greatly changed by complexing with the inducer, the diffusion constant and steric factors will not be changed and the two association rates, k_1 and k_2, will be equal. Thus

$$\frac{[D\text{-}R\text{-}I]}{[D\text{-}R]} = \frac{K_0[R\text{-}I]}{\tilde{K}_0[R]} = \frac{k_s/k_1}{k_f/k_2}\frac{[R\text{-}I]}{[R]} \approx \frac{k_s}{k_f}\frac{[R\text{-}I]}{[R]}.$$

This last is the crucial statement: the amount of DNA-repressor-inducer complex is small just in proportion to its instability. The decay of the repressor-operator complex can be rewritten as

$$\frac{d}{dt}[D\text{-}R] = [D\text{-}R]\, k_s \left(1 + \frac{[R\text{-}I]}{[R]}\right).$$

Thus the complex decays exponentially with a time constant

$$T = \frac{1}{k_s}\left(\frac{[R]}{[R] + [R\text{-}I]}\right).$$

Since there is a reasonable excess of repressor over operators, $[R] + [R\text{-}I]$ is approximately the total amount of repressor in the cell, and the quantity in parentheses is just $(r_B/r_E) - r_B$ because the rate of enzyme synthesis is given by

$$r_E = 1/(1 + [R]/K_0),$$

if one assumes that this rate is proportional to the amount of DNA free of repressor.

We wish to thank Christine Weiss for her excellent technical assistance, Drs. Ethan Signer and Jonathan Beckwith for providing bacteria and phage strains, and the National Institutes of Health (GM 09541-06) for their support of this work.

[1] Gilbert, W., and B. Müller-Hill, these PROCEEDINGS, **56**, 1891 (1966).
[2] Ptashne, M., these PROCEEDINGS, **57**, 306 (1967).
[3] Ptashne, M., *Nature*, **214**, 232 (1967).
[4] Sadler, J. R., and A. Novick, *J. Mol. Biol.*, **12**, 305 (1965).
[5] Boezi, J. A., and D. B. Cowie, *Biophys. J.*, **1**, 639 (1961).
[6] Signer, E., *Virology*, **22**, 650 (1964).