DNA-DIRECTED PEPTIDE SYNTHESIS, VI. REGULATING THE EXPRESSION OF THE LAC OPERON IN A CELL-FREE SYSTEM*

By Geoffrey Zubay and Muriel Lederman†

DEPARTMENT OF BIOLOGICAL SCIENCES, COLUMBIA UNIVERSITY

Communicated by Erwin Chargaff, November 15, 1968

Abstract.—The operon model was proposed by Jacob and Monod in 1961 to explain the regulation of enzyme synthesis in bacteria. The model requires two genetic elements: a locus *i* which directs the synthesis of a repressor, and a locus called the operon which contains an operator gene and associated structural genes. The repressor is believed to stop RNA and enzyme synthesis by combining with the operator site. The inhibiting effect of the repressor is reversed by an inducer which interacts directly with the repressor. A DNA-dependent cell-free system has been developed for studying the regulation of the operon containing the gene for β-galactosidase enzyme. In this cell-free system, gene activity is indirectly measured by the synthesis of part of the enzyme. Qualitatively, the regulation of the operon in the cell-free system is similar to its regulation in intact cells. The cell-free system is used to advantage in experiments where it is desirable to use predetermined levels of repressor, operon, and inducer. Such quantitative studies suggest that two inducer molecules are most effective in removing the repressor from the gene.

We have described a DNA-dependent cell-free system in which it is possible to synthesize a part of the polypeptide chain of β-galactosidase.[1] This peptide, referred to as α, is believed to correspond to about the first fifth of the polypeptide chain. Its presence is quantitatively estimated by complementation with a defective β-galactosidase to give active enzyme. The synthesis of α is partially inhibited by the *lac* repressor. The effect of repressor is eliminated in the presence of IPTG (isopropyl-1-thio-β-D-galactopyranoside), a strong inducer of the *lac* operon. These observations are in qualitative agreement with the regulatory behavior of the *lac* operon in whole cells.[2] It seems likely that the inhibiting effect of *lac* repressor on the synthesis of α is due to the binding of the repressor to the *lac* operator. The cell-free system is used to advantage in experiments where it is desirable to use predetermined levels of repressor, operon, and inducer. Recent improvements made in the synthetic efficiency of the cell-free system, the purity of the repressor, and the precision of measurements have led us to some new observations concerning the thermodynamics of regulation.

Materials and Methods.—*E. coli strains:* (*a*) 21 F'i^- contains the gene for β-galactosidase with the M-15 deletion described by Ullman, Jacob, and Monod[3] on both chromosome and episome. (*b*) 21 F'i^+ is identical to 21 F'i^-, except that it carries the *lac* repressor gene on both chromosome and episome. (*c*) Z 19i^q contains the mutant repressor gene i^q and the M-15 modified β-galactosidase gene on both chromosome and episome. This strain was obtained from strain BMH 292 (*lac* i^{+qTSS}, gal^-, SMR) by suitable crosses and selection. From the first cross of BMH 292 × 21 F'i^-, lac^- males containing i^q and M-15 modified β-*gal* gene on the episome were isolated. From the same cross, a set of lac^- males were isolated; these were cured in pen broth with 1% acridine orange, and an F^- strain with mutant *i* and β-*gal* gene on the chromosome was isolated. The male and

female $i^q lac^-$ types selected above were mated to give strain Z $19i^q$. (4) Z-$19i^-$ is identical to Z $19i^q$ except that it contains no *lac* repressor gene. The strains described above are referred to in the text as $21i^+$, $21i^-$, $19i^q$, and $19i^-$, respectively.

Incubation conditions for synthesis: The incubation mixture used is similar to that described by Lederman and Zubay[4] except that the concentration of all ingredients other than the S-30 and DNA is 10% higher. The DNA concentration is specified in the various experiments.

Purification of repressor: Repressor was prepared from strain $19i^q$. Crude repressor called fraction IV was used for determining the formation constant for the repressor-inducer complex; it was prepared according to a method described elsewhere.[1] Highly purified repressor was obtained from fraction IV by gradient elution on phosphocellulose according to the procedure of Riggs and Bourgeois.[5] The repressor eluted similarly to peak B of Riggs and Bourgeois and was about 300-fold purified over crude *E. coli* extract by the IPTG binding assay described below.

Determination of repressor concentration: The concentration of repressor was determined by a modification of the method of Gilbert and Müller-Hill,[6] who measured the binding of ^{14}C-IPTG by dialysis equilibrium in TMS buffer. We have found that, in this buffer, the binding curve is not interpretable in terms of a single association constant, suggesting that the binding of one IPTG molecule by repressor influenced the binding of the subsequent IPTG molecule(s). Most of our binding studies were done in a more concentrated mixture of salts: 40 mM Tris-Ac, pH 8.2, 1.25 mM dithiothreitol, 50 mM KAc, 25 mM NH_4Ac, 13.3 mM $Mg(Ac_2)$, 6.7 mM $CaCl_2$, and 60 mM NaAc. Since this buffer is closer in ionic strength to the synthetic system (and the binding data were interpretable in terms of one association constant), over a broad range of IPTG concentration a plot of IPTG (bound)/IPTG (free) versus IPTG (free) gives a constant slope from which an intrinsic association constant $K_{R'I}$ of 1.8×10^6 moles/liter is calculated (symbols defined in section (1) of *Results and Discussion*). It was possible to measure the repressor concentration in the $19i^q$ S-30 directly because, in this buffer, there is little or no binding of IPTG to substances other than repressor as determined by the absence of binding on $19i^-$ S-30; the concentration of repressor in the $19i^q$ strain is high. The repressor concentration could not be determined directly on $21i^+$ S-30's but had to be estimated from measurements on partially purified repressor. It was estimated to contain about one ninth as much repressor as the $19i^q$ S-30; this is in reasonable agreement with Müller-Hill, Crapo, and Gilbert.[7]

Results and Discussion.—(1) *Description of the cell-free system for synthesis of* α: A detailed description of the cell-free system is given elsewhere.[1, 4, 8–10] The DNA used is isolated from a ϕ80 *dlac* virus preparation containing approximately equal amounts of ϕ80DNA and ϕ80*dlac* DNA. Since the molecular weight of the viral DNA is 30×10^6 gm, there is approximately one mole of *lac* operon per 60×10^6 gm of DNA. The total operon concentration is calculated from this and the known amount of added DNA. The bacterial extract used is a preincubated S-30 preparation made according to a modification of the method of Nirenberg.[4] Four different strains have been used in the preparation of S-30 extracts: $19i^q$, $19i^-$, $21i^+$, $21i^-$. All of the strains contain a large excess of α complement. The two i^- strains do not make *lac* repressor. The i^+ strain makes wild-type repressor, whereas the i^q strain makes about nine times as much of a mutant-type repressor. In addition to the DNA and the S-30 extract, the cell-free system contains salts and substrates required for transcription and translation (see *Materials and Methods*). The reaction mixture is incubated for one hour at 37°C. This is sufficient time for maximum α synthesis and its complementation. The β-galactosidase activity of an aliquot is measured by standard techniques and is reported here as $OD_{420\ m\mu}$ after 20 hours of assay.

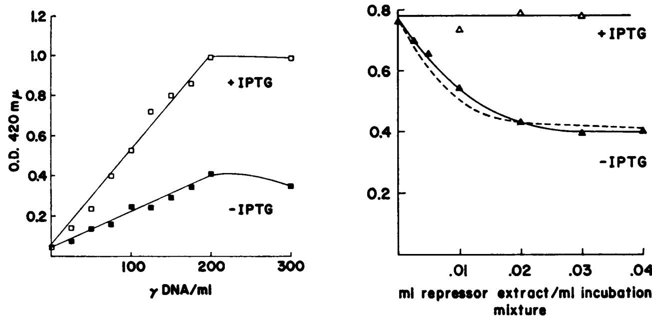
In some experiments, repressor made from strain $19i^q$ is added to the incubation mixture. The amount of repressor in the S-30 or purified repressor extract is determined by the binding of inducer (see *Materials and Methods*). The binding of inducer to repressor in a salt solution comparable in strength to the buffer used in the cell-free system obeys the following relationship:

$$K_{R'I} = \frac{(R'I)}{(R')\ (I)} = 1.8 \times 10^6,$$

where $R'I$ is the molar concentration of bound inducer, I is the molar concentration of free inducer, R' is the concentration of available inducer binding sites, and $K_{R'I}$ is the intrinsic association constant for the repressor-inducer complex. For purposes of discussion, we shall express the repressor concentration, R, as one half the molar concentration of inducer binding sites, R'.

(2) *Dependence of α synthesis on DNA concentration:* If α synthesis is to be used as a measure of gene activity, we must first know the relationship between the amount of α synthesized and DNA concentration. The cell-free system contains a large excess of α complement. Evidence has been presented elsewhere[11] that the active enzyme formed in the incubation mixture contains one α polypeptide attached to a molecule of the complementing protein and that the amount of enzyme activity formed is proportional to the amount of α present. In Figure 1, it is shown that the amount of enzyme activity formed with a $19i^q$ S-30 is also directly proportional to the amount of $\phi 80dlac$ DNA added, up to about 200 γ/ml of DNA. The enzyme activity obtained is linearly related to the amount of DNA added either in the presence or absence of IPTG. The activity obtained in the presence of IPTG is about 2.5 times that obtained in its absence. Incubation mixtures containing $21i^+$ S-30 give qualitatively similar results. $19i^-$ or $21i^-$ S-30's which do not contain repressor are unaffected by the presence of inducer, as would be expected. The total activity as well as the ratio of activity obtained under derepressed conditions (+ IPTG) to that obtained under repressed conditions (− IPTG) varies considerably with different S-30's, even from the same strain. Highly reproducible results are only obtained when the same S-30 is used. In spite of these variations, any set of data obtained from different S-30 preparations leads to the same conclusions regarding the thermodynamics.

(3) *Dependence of α synthesis on repressor concentration:* Figure 1 shows that 40 per cent of the maximum enzyme activity is obtained under repressed conditions. The amount of enzyme activity obtained under repressed conditions *in vivo* is less than 1 per cent of that obtained under derepressed conditions. The relatively high basal activity shown by the repressed cell-free system could result from: (*a*) incomplete binding of repressor to operator resulting from insufficient repressor, (*b*) incomplete binding of repressor to operator resulting from a lowered affinity of repressor for operator, (*c*) factors allowing α synthesis despite the presence of the repressor-operator complex. In order to assess these possibilities, the dependence of enzyme activity on amount of repressor was determined. For this purpose, an S-30 was prepared from strain $21i^-$ which contains no repressor. Partially purified repressor was prepared from strain $19i^q$ (see *Materials and*



(*Left*) Fig. 1.—Enzyme activity as a function of ϕ80*dlac* DNA concentration with and without added inducer. Incubation and assay carried out as described in *Materials and Methods*. In experiments where IPTG is present, the final concentration is 5×10^{-5} *M*. The S-30 is prepared from strain $19i^q$.

(*Right*) Fig. 2.—Enzyme activity as a function of repressor concentration with and without 5×10^{-5} *M* IPTG added. The S-30 is prepared from strain $21i^-$. The repressor is prepared from strain $19i^q$, as described in *Materials and Methods*. The repressor is added to the S-30, and the combination is added to the incubation mixture in the usual way. 145 γ/ml of ϕ80*dlac* DNA is present in all incubations. The dashed line represents a calculated curve discussed in section (5) of *Results and Discussion*. The dashed line is calculated from the equation $K_{RO} = (RO)/[(R)\ (O)]$, using a value for the repressor-operator association constant K of 2.9×10^9 moles^{-1} liter. The OD_{420} reading of 0.78 in the absence of repressor is used as the value for RO equals zero. The OD_{420} reading of 0.40 is used as the value for O equals zero. The total operator concentration, O_T, is equal to the sum of O and RO and was calculated to be 2.4×10^{-9} *M*. The total repressor concentration, R_T, is equal to the sum of R and RO. R_T was calculated from the concentration of repressor in the purified extract, 2.7×10^{-7} *M*, and the known amount added.

Methods). The amount of enzyme activity obtained in the presence of varying amounts of the added repressor is shown in Figure 2. In the absence of inducer, there is a decrease in activity with increasing repressor until a plateau is reached (lower curve). Beyond this point, there is no further decrease in activity, demonstrating that a saturating level of repressor is present and ruling out possibilities (*a*) and (*b*) mentioned above. As a control, parallel observations (upper curve) were made to show that, if sufficient inducer is present during synthesis, the enzyme activity obtained is not lowered by added repressor. Addition of partially purified repressor to an S-30 prepared from strains $19i^q$ or $21i^+$ does not increase the amount of repression. This shows that a saturating amount of repressor is already present when these repressor-containing S-30's are used. Our current belief is that the high level of irrepressible synthesis in the cell-free system is caused by conditions that allow transcription and translation to start at abnormal locations insensitive to repressor binding. For the purposes of the present study, it is sufficient to keep in mind that synthesis is the sum of repressible and irrepressible parts, and that we distinguish between the two. Repressible synthesis is sensitive to the levels of repressor and inducer present; irrepressible synthesis is not. For a given S-30, the amount of irrepressible synthesis is proportional to that enzyme activity obtained at a saturating level of repressor in the absence of inducer. This amount must be subtracted from the total en-

zyme activity observed at varying degrees of induction to obtain a measure of the repressible synthesis.

(4) *Dependence of α synthesis on inducer concentration:* The derepression of synthesis by inducer was studied at otherwise saturating levels of repressor by using S-30's from $19i^q$ or $21i^+$. The results with a $19i^q$ S-30 are presented in Figure 3. Irrepressible synthesis has the value of 0.2 OD unit, and maximum synthesis has a value of 0.5. If the value for irrepressible synthesis is taken into account, the value for half-maximum induction occurs at 0.35, which has an IPTG concentration of 0.41×10^{-5} *M*. Five determinations derived from different S-30's give qualitatively similar results to those presented in Figure 4 with the value for IPTG at 50 per cent induction ranging from 0.33 to 0.50^{-5} *M* IPTG. Similar studies with $21i^+$ S-30's show that an IPTG concentration of 0.3×10^{-5} *M* is necessary to achieve half-maximum induction. In view of the much lower repressor concentration in strain $21i^+$, such a value suggests a higher affinity between operator and wild-type repressor.

(5) *Thermodynamic calculations:* The object of this section is to give a quantitative interpretation to the previous results. The following assumptions, which we have already attempted to justify, will be used: (*a*) Only part of the synthesis is repressible and this can be determined by the procedure described in section (3). (*b*) If we subtract the irrepressible synthesis, the amount of enzyme formed is directly proportional to the ϕ80*dlac* DNA free of repressor. We shall designate this as *O* and use *RO* to designate repressor-complexed DNA. $O + RO$ is calculated from the total amount of DNA added. In setting up thermodynamic expressions to represent the equilibrium between repressor, operon, and inducer, we considered the possibilities that one or more molecules of repressor and one or more molecules of inducer, respectively, might interact with a single operon. It was found that the simplest acceptable fit of the results was obtained with the following expressions:

In the absence of inducer,

$$R + O \rightleftharpoons RO. \qquad (1)$$

In the presence of excess repressor and varying amounts of inducer,

$$RO + 2I \rightleftharpoons RI_2 + O, \qquad (2)$$

where *O*, *R*, and *RO* are as defined; *I* represents IPTG, and RI_2 represents repressor complexed with two IPTG molecules.

The corresponding expressions for the association constants are

$$K_{RO} = \frac{(RO)}{(R)(O)} \qquad (3)$$

and

$$K_{RO} = (K_{R'I})^2 \left(\frac{RO}{O}\right)\left(\frac{1}{RI_2}\right)I^2 \qquad (4)$$

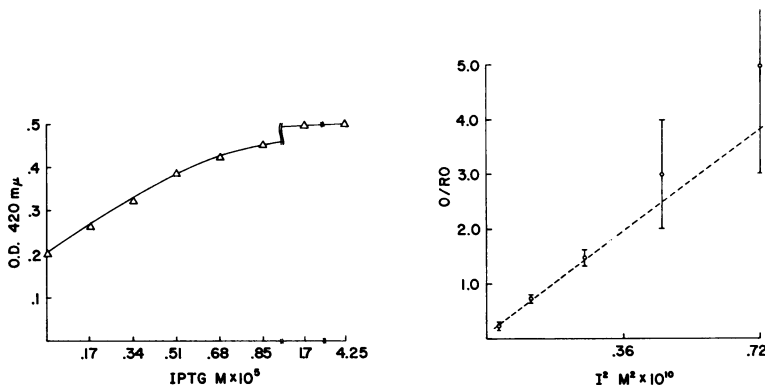
respectively, where $K_{R'I}$ is the intrinsic association constant for the repressor-

inducer complex as previously defined, and K_{RO} is the intrinsic association constant for the repressor-operator complex.

Equation (4) can be transposed to give

$$\frac{O}{RO} = \frac{(K_{R'I})^2}{K_{RO}}\left(\frac{1}{RI_2}\right)I^2 \qquad (5)$$

The data of Figure 3 have been used to obtain Figure 4, in which O/RO is plotted against I^2. A close approximation to a straight line is obtained, as would be expected, if everything besides I^2 on the right-hand side of equation (5) were a constant. Actually, RI_2 varies somewhat over the range of concentrations studied. It can be calculated that RI_2 varies from 14.2 to 21.2 $\times$ 10^{-9} M, while I varies from 1.6 to 9.7 $\times$ 10^{-6} M. At the point where O/RO has a value of unity, a value for K_{RO} of 2.9 $\times$ 10^9 is calculated from the observed results and equation (5). The dashed-line plot in Figure 4 has been obtained by use of equation (5) and its value for K_{RO}. The agreement between the calculated curve and the experimental points is within experimental error. A similar analysis



(*Left*) FIG. 3.—Enzyme activity as a function of IPTG concentration. The S-30 used was prepared from strain 19i^q. 145 γ/ml of φ80*dlac* DNA was added in each experiment. The total repressor concentration, R_T, is 23 $\times$ 10^{-9} M in all experiments. The lowest total inducer concentration used is 1.6 $\times$ 10^{-6} M so that free inducer is very nearly equal to total inducer. The inducer is usually added to the S-30 20 min before the S-30 is added to the incubation mixture, although addition as little as 5 sec before gives the same results. The DNA is usually added to the incubation mixture before the S-30, although the same results are obtained if the DNA is added to the S-30 first. Each point on this curve is the average of four determinations with an average error of ±2%.

(*Right*) FIG 4.—O/RO versus $(\text{IPTG})^2$. The assumptions used in the calculations of O and RO have been explained in the text. O is proportional to the difference $OD_{420}^{I} - OD_{420}^{I=0}$, where the first term is the optical density obtained in the assay at any given value of IPTG, and the second term is the same quantity measured when at zero IPTG. RO is proportional to the difference $OD_{420}^{I\,max} - OD_{420}^{I}$, where the first term is the optical density obtained in the assay at the highest concentration of IPTG in the incubation mixture of 5 $\times$ 10^{-3} M. The total repressor concentration, R_T, is 23 $\times$ 10^{-9} M. R_T is equal to the sum of R, RO, RI, and RI_2, where RI represents a functional unit of repressor complexed with one inducer molecule, and RI_2 represents a functional unit of repressor complexed with two molecules of repressor. The concentration of RI_2 is determined through use of the above equality and the expression for the intrinsic formation constant for the repressor-inducer complex. The assumption is made that all sites for binding inducer have the same affinity, as indicated by inducer binding studies (see *Materials and Methods*).

for wild-type repressor gives an approximate value for K_{RO} of 2–4 × 10^{10}. The proportional relationship observed between O/RO and I^2 makes it highly likely, but does not prove that at this high level of repressor the main route to derepression involves two inducer molecules. Thermodynamic analysis is limited to describing the major species present in a reaction system. A rigorous demonstration of the mechanism of derepression would require a kinetic analysis which is beyond the scope of this study. Herzenberg's data[12] on derepression *in vivo* can be replotted to give a similar proportionality between O/RO and I^2.

The value of K_{RO} determined from induction studies should also be compatible with the data presented in Figure 2, which shows the decrease in enzyme activity resulting from increasing repressor concentration in the absence of inducer. The dashed-line plot in Figure 2 is calculated from the estimated value of K_{RO}, the known values for added repressor, and equation (3). The calculated curve gives the best agreement with the experimental points when the ratio of 2 is used for inducer binding sites to repressor binding sites.

Gilbert and Müller-Hill[13] and Riggs *et al.*[14] have attempted to measure directly the binding of repressor protein to DNA. These determinations showed that the *RO* complex is very sensitive to its environment. A meaningful comparison is not possible now but should be possible once the DNA-repressor binding has been measured under conditions where gene expression occurs. Such conditions for the cell-free state have been described in this paper. The association constant for the repressor-operator complex *in vivo* has been estimated as 5–10 × 10^{10} for wild-type repressor. This is in reasonable agreement with the value of 2–4 × 10^{10} obtained here. The considerably lower value for the mutant i^q repressor is consistent with the observations of Müller-Hill, Crapo, and Gilbert.[7]

(6) *A physical picture of the repressor:* *Lac* repressor is believed to be a tetramer with a mole weight of about 1.5 × 10^5 gm.[7, 14] Riggs *et al.* have isolated monomers that bind to inducer[5] but not to operator.[15] This suggests that there are four inducer binding sites per tetramer. The results reported here suggest that two inducer molecules catalyze the release of repressor and that the ratio of inducer binding sites to operator binding sites is 2. It is not difficult to arrange subunits within a tetramer consistent with these properties. An example of such a structure is a tetramer in which the asymmetric monomers are placed at the apexes of a tetrahedron so as to give three dyad axes of symmetry. If the operator binding site is located on a dyad axis, two predictions can be made from consideration of symmetry alone: (*a*) Two operator binding sites would be expected, one on either side of the molecule. This gives the ratio of 2 for inducer binding sites to operator binding sites. (*b*) The monomers are arranged symmetrically in pairs about each operator binding site. Without knowing the detailed effects of inducer binding, it seems likely that each operator binding site would only be responsive to the binding of inducer to one of the pairs of monomers. This would be in agreement with the data on derepression which suggest that the binding of two inducer molecules catalyzes the dissociation of the repressor-operator complex.

The thermodynamic arguments presented in section (5) treat the repressor as a functional unit containing two inducer binding sites and one operator binding

site. The difference between this treatment and the treatment for a tetramer similar to the one just described should be negligible, with the possible exception of the region of low levels of repressor where the binding of two DNA molecules to one repressor molecule can become a significant factor. Further quantitative studies are needed to determine whether two repressor-operator association constants are required to fit the data in this region.

Note added in proof: The addition of 1 mM adenosine 3′,5′-monophosphate to the cell-free system herein described has improved the yield of β-galactosidase enzymatic activity by about tenfold and the efficiency of repression from 50 to 95%. The detailed effects of 3′,-5′ AMP will be the subject of subsequent communications from this laboratory.

We are most grateful to Dr. Benno Müller-Hill for giving us strain BMH 292.

* This work was supported by grants from the Damon Runyon Research Foundation (no. DRG-989) and from the National Science Foundation (grants 4824 and 7597).

† Predoctoral fellow of the National Institutes of Health during the period of this work.

[1] Zubay, G., M. Lederman, and J. K. DeVries, these PROCEEDINGS, **58**, 1669 (1967).

[2] For a recent review of the subject, see Epstein, W., and J. R. Beckwith, *Ann. Rev. Biochem.*, **37**, 673 (1968).

[3] Ullman, A., F. Jacob, and J. Monod, *J. Mol. Biol.*, **32**, 1 (1968).

[4] Lederman, M., and G. Zubay, *Biochem. Biophys. Res. Commun.*, **32**, 710 (1968).

[5] Riggs, A. D., and S. Bourgeois, *J. Mol. Biol.*, **34**, 361 (1968).

[6] Gilbert, W., and B. Müller-Hill, these PROCEEDINGS, **56**, 1891 (1966).

[7] Müller-Hill, B., L. Crapo, and W. Gilbert, these PROCEEDINGS, **59**, 1259 (1968).

[8] Lederman, M., and G. Zubay, *Biochim. Biophys. Acta*, **149**, 253 (1967).

[9] DeVries, J. K., and G. Zubay, these PROCEEDINGS, **57**, 1010 (1967).

[10] Hercules, K. D., R. Knacht, and G. Zubay, *Virology*, **34**, 337 (1968).

[11] DeVries, J. K., and G. Zubay, *J. Bacteriol.*, in press.

[12] Herzenberg, L., *Biochim. Biophys. Acta*, **31**, 525 (1959).

[13] Gilbert, W., and B. Müller-Hill, these PROCEEDINGS, **58**, 2415 (1967).

[14] Riggs, A. D., S. Bourgeois, R. F. Newly, and M. Cohn, *J. Mol. Biol.*, **34**, 365 (1968).

[15] Bourgeois, S., personal communication.