Proc. Nat. Acad. Sci. USA
Vol. 71, No. 12, pp. 4808–4812, December 1974

Non-specific DNA Binding of Genome Regulating Proteins as a Biological Control Mechanism: I. The *lac* Operon: Equilibrium Aspects

(*lac* repressor/DNA-protein interactions/RNA polymerase/repressor–inducer complexes)

PETER H. VON HIPPEL, ARNOLD REVZIN, CAROL A. GROSS*, AND AMY C. WANG

Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, Oreg. 97403

Communicated by V. Boekelheide, September 3, 1974

ABSTRACT **The regulatory system of the lactose operon has been "modeled" by a set of mass action equations and conservation constraints which describe the system at equilibrium. A "base-set" of values of binding constants and total component concentrations has been assembled from the available experimental data, and the simultaneous equations solved by computer procedures, to yield equilibrium concentrations of all the relevant molecular species. Considering the operator–repressor–inducer system alone, it is shown that the *in vivo* basal and induced (derepressed) levels of *lac* enzyme synthesis in both wild-type and certain mutant *Escherichia coli* can be accounted for only if binding of repressor and repressor–inducer complexes to non-specific DNA sites is included in the calculations as an integral component of the overall control system. A similar approach was applied to the RNA polymerase–promoter system to show that sigma factor may modulate the general level of transcription in the cell by "inducing" polymerase off non-specific DNA binding sites, thus making it available to promoters. Competitive and non-competitive models for the interaction of repressor and polymerase at the *lac* operon can, in principle, be distinguished by these computational procedures, though data sufficient to permit unambiguous differentiation between the models are not available at this time. However, for any competitive binding model the results show that repression in the entire (operator–repressor–RNA polymerase–*lac* promoter) system can occur only because non-specific binding of the regulatory proteins reduces the concentration of free polymerase, relative to that of repressor, to appropriate levels.**

The interaction of genome-regulating proteins (repressors, polymerases, etc.) with their specific target sequences on DNA can be perturbed by the relatively weak binding of such proteins to the non-specific DNA sites which are present in overwhelming preponderance within the cell. The equilibrium distribution of these proteins will depend on their relative affinities for the available sites, and on the relative concentrations of site types. Under physiological conditions the regulatory proteins (which are often present in fairly few copies per cell) are likely to be almost entirely complexed with non-functional sites, i.e., the concentration (activity) of free regulatory proteins in the cell will be very low. As a consequence the kinetics of the association (and dissociation) of such proteins with functional sites on the *Escherichia coli* chromosome will also depend on the relative numbers and spatial distribution of specific and non-specific binding sites on the DNA (1).

In this paper we consider equilibrium aspects of these effects, as illustrated by a computer modeling study of the *lac* repressor–operator–inducer–non-specific DNA binding system, and the coupling of these interactions with the RNA polymerase–promoter–non-specific DNA ensemble. Experimental results required for the calculations are taken largely from the literature, though some recent direct measurements of the binding of *lac* repressor to non-specific DNA made in our laboratory are used as well (ref. 1 and manuscript in preparation).

I. The *lac* repressor–operator–inducer–non-specific DNA system

Any model for the regulation of the *lac* system must be quantitatively consistent with the following facts: (*a*) Repression of the *lac* operon in *E. coli* is a consequence of repressor (R) binding to operator (O), and induction (derepression) is due to inducer (I) binding to R, forming an RI_n complex with a decreased affinity (relative to that of R) for O (2–5). (*b*) The constitutive rate of *lac* enzyme production is approximately the same in the presence of saturating inducer and in i^- (inactive R) mutant *E. coli* cells. Thus, full derepression is possible with attainable inducer concentrations (6). (*c*) The constitutive rate of *lac* enzyme synthesis is about 10^3 times greater than the fully repressed (basal) rate in wild-type cells (6). (*d*) An increase in intra-cellular R level decreases the basal rate of *lac* enzyme synthesis in direct proportion (7, 9, 10). (*e*) Inducer concentrations necessary to achieve constitutive rates range from about 10^{-1} to about 10^{-3} M for inducers with repressor (sub-unit) association constants (K_{RI}) ranging from 10^4 to 10^6 M^{-1}. Repressor mutants exhibiting decreased values of K_{RI} require increased I concentrations to achieve full derepression (8). (*f*) Operator single- (or double-) base-pair mutations (O^c) increase the basal rate by factors of 10- to 500-fold. *In vitro* filter-binding assays show that these mutated operator sequences display corresponding decreases in association constant for wild-type repressor (9, 10). (*g*) The affinity of the RI_n complex for O is about 10^3 less than that of R for O (9–11). (*h*) R and RI_n bind to non-specific DNA with approximately equal affinity (ref. 1 and manuscript in preparation).

Equilibrium Considerations. The following parameters relevant to the wild-type *E. coli* cell have been used to establish a "base-set" of constants and constraints to describe the coupled equilibria which apply *in vivo*. The internal volume of the wild-type cell is taken as 10^{-15} liters. We assume a total operator concentration $[O_T]$ of about 2×10^{-9} M (about 1 per cell), a total repressor concentration $[R_T]$ of about 2×10^{-8} M (about 10 per cell), a total non-specific DNA site concentration $[D_T]$ of about 2×10^{-2} M (about 10^7 per cell; the *E.*

* Present address: McArdle Laboratory, University of Wisconsin, Madison, Wisc. 53706.

coli chromosome contains about 10^7 base pairs, and in principle every base pair represents the beginning of a separate non-specific DNA binding site), and a fully-derepressing total inducer concentration $[I_T]$ of about 10^{-3} M (this applies to the gratuitous inducer, isopropyl-β-D-thiogalactoside, termed IPTG).

To illustrate our approach we consider the system depicted in Fig. 1, which shows the relevant equilibrium constants and molecular species. The definitions of the necessary binding constants and constraints for this model, together with the "base-set" values used, are†‡:

$$\begin{aligned} K_{RO} &= [RO]/[R]\cdot[O] \simeq 10^{14}\ \mathrm{M}^{-1} \\ K_{RD} &= [RD]/[R]\cdot[D] \simeq 10^{5}\ \mathrm{M}^{-1} \\ K_{RI} &= [RI]/[R]\cdot[I] \simeq 10^{6}\ \mathrm{M}^{-1} \\ K_{RIO} &= [RIO]/[RI]\cdot[O] \simeq 10^{-3}\,K_{RO} \\ K_{RID} &= [RID]/[RI]\cdot[D] \simeq K_{RD} \end{aligned} \qquad [1]$$

$$\begin{aligned} [O_T] &= [O] + [RO] + [RIO] \\ [R_T] &= [R] + [RO] + [RD] + [RI] + [RIO] + [RID] \\ [D_T] &= [D] + [RD] + [RID] \\ [I_T] &= [I] + [RI] + [RIO] + [RID] \end{aligned} \qquad [2]$$

Since the exact internal milieu of the *E. coli* cell is unknown, the "base-set" values adopted can only represent best estimates. The most important assumption concerns the parameter K_{RO}. The minimum value possible for K_{RO} in *E. coli* is about 10^{11} M^{-1}; this value is calculated assuming that all R not bound to O is free in solution (4). We estimate from our *in vitro* measurements that K_{RD} is about 10^5 M^{-1} under ionic conditions likely to be found in the bacterium. (It is often stated that the *in vivo* environment contains about 0.2 M KCl and about 0.003 M Mg^{++}, as well as high concentrations of spermidine and other ionic species. In fact, many of these ions must be bound to various macromolecular constituents of the cell, so that the effective ionic strength may be appreciably lower.) A large fraction of R will therefore be bound to non-specific DNA sites, which, in turn, requires that K_{RO} be considerably greater than 10^{11} M^{-1}; our choice of $K_{RO} = 10^{14}$ M^{-1} is based on these considerations and interpolation of *in vitro* measurements of K_{RO} at various ionic strengths to reasonable (assumed) *in vivo* ionic activities (12). The value for K_{RI} (for IPTG) has been measured by several groups (8). The relation between K_{RIO} and K_{RO} is discussed above (fact *g*), as is the equality of K_{RID} and K_{RD} (fact *h*). It should be emphasized that the validity of the conclusions reached below is not critically dependent on the exact values used for the "base-set" parameters; in most instances the *ratios* of binding constants, rather than their absolute values, are the important quantities.

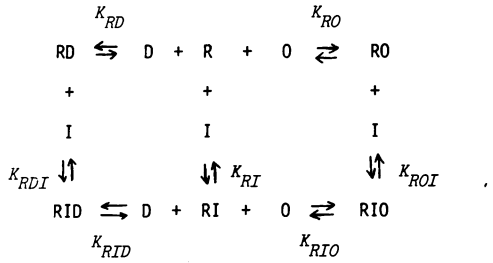

FIG. 1. Model of the repressor–operator–inducer–non-specific DNA system: in which R = repressor, O = operator, I = inducer, D = non-specific DNA sites; *RO*, *RD*, *RI*, *RIO*, and *RID* are the various complex species; and K_{RO}, K_{RD}, etc., represent the indicated association constants.

Calculations. The simultaneous Eqs. 1, subject to the constraints of Eqs. 2, were solved by appropriate iterative procedures (using a PDP-10 computer system) for the concentrations of all molecular species. Calculations were performed using the "base-set" values discussed above, and also with different values of these parameters. This permits us to assess the sensitivity of the results to variations in the values of the "base-set" quantities. Furthermore, we can readily account for the effects of mutations on the equilibrium distribution of *lac* system components; for instance, the computations are made using smaller values of K_{RO} for O^c mutations (9, 10), or with smaller values of K_{RI} for certain mutant "superrepressors" (11). The results are usually expressed as fraction of free operator ($[O]/[O_T]$) present, since this is the parameter which is actually (indirectly) measured *in vivo*§.

The Basal Level. The importance of the non-specific DNA binding process in controlling the magnitude of the observed repression of the *lac* operon is illustrated in Fig. 2, in which we plot the calculated fraction of free operator sites as a function of either non-specific repressor binding constant or total concentration of non-specific sites. No inducer is present, and we consider only one class of non-specific binding sites. At low K_{RD} (or low $[D_T]$), there is no non-specific binding and the system is repressed to a basal rate of about 5×10^{-7} of the constitutive level. As we increase K_{RD} (or $[D_T]$) the extent of repression decreases, approaching asymptotically the totally unrepressed state ($[O]/[O_T] = 1$). Note that the value of

† Note that we consider only a single binding interaction for I with R in these equilibria, although it is likely that binding of more than one inducer per tetrameric repressor is required to bring about complete derepression. However, since we define K_{RIO} (and K_{RID}) as the observed binding constants for the repressor–inducer complex at *saturating* inducer concentrations, for present purposes we can represent the situation in terms of a single inducer binding site. By the above definition of K_{RIO} (and K_{RID}) the concentration of I at which the system is induced to twice the original basal level will be independent of the actual value of n involved in the derepressing RI_n complex. On the other hand the *shape* of the induction curve *will* depend on n. We also list only a single mass action equation for the *RD* interaction, since direct repressor binding experiments (1) show that the predominant class of sites can be characterized by a single binding constant, which is essentially independent of DNA nucleotide composition and sequence and is of about the order of magnitude indicated. However, in principle, we can sub-divide the D sites into different types, each characterized by a separate mass action relation (see Fig. 3). Similarly, repressor–inducer complexes of higher inducer valence (e.g., RI_2, ...; RI_2O, ...; RI_2D, ...; etc.) can be incorporated into the model by adding suitable mass action equations and modified conservation constraints.

‡ Values of K_{ROI} and K_{RDI} need not be assigned, since they are defined by the other equilibrium constants of the model.

§ Actually one measures *in vivo* the ratio of basal (repressed) to constitutive intra-cellular β-galactosidase activity ($Z_{basal}/Z_{const.}$). For the physiological evidence showing that this ratio is a direct reflection of the fraction of free operator, see references relevant to facts (*a*), (*b*), (*d*), and (*f*) above.

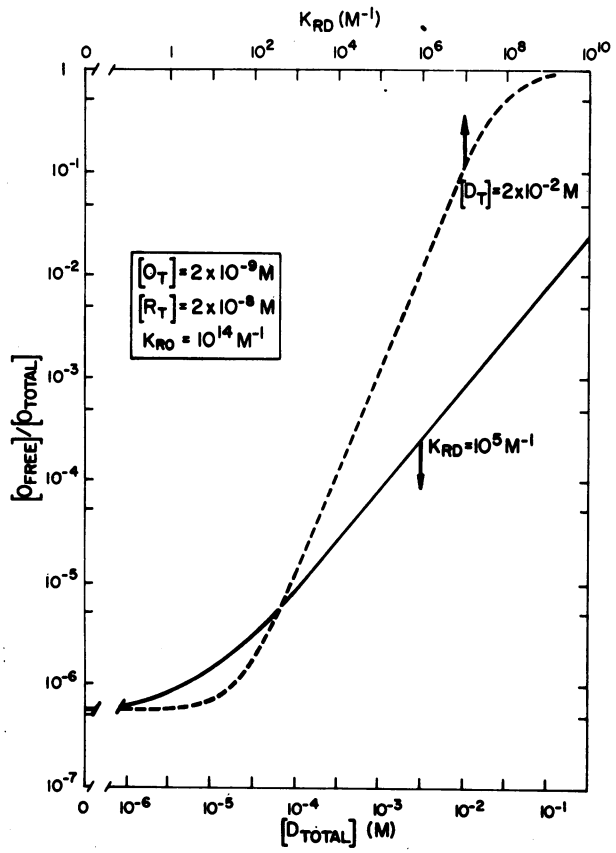

FIG. 2. Control of the fraction of free *lac* operator in *E. coli* by the binding of repressor to non-specific DNA sites. The amount of *RD* complex is altered by varying either K_{RD} (dashed curve and upper abscissa; $[D_T]$ taken as 2×10^{-2} M) or $[D_T]$ (solid curve and lower abscissa; K_{RD} taken as 10^5 M^{-1}).

$[O]/[O_T]$ which these calculations predict at the levels of $[D_T]$ and K_{RD} which presumably apply in the *E. coli* cell is $\simeq 10^{-3}$, which is close to the basal level actually measured *in vivo* (fact *c*, above). Thus, these results show that in the absence of *R* binding to non-specific DNA sites, the basal rate of *lac* enzyme synthesis should be 3 to 4 orders of magnitude smaller than is actually observed, and that the observed basal level is primarily established by non-specific binding of *R* to *D* sites, the latter acting as "sinks" for *R*, in competition with *O*.

Effects of Increasing $[R_T]$. Calculations (data not shown) designed to determine the repressed (basal) level of *lac* enzyme synthesis as a function of $[R_T]$, show that in mutants that overproduce repressor the basal level is expected to decrease linearly with increasing $[R_T]$. This is consistent with physiological measurements (fact *d*, above).

Effects of Additional Non-specific Sites. Fig. 3 shows the calculated effects on $[O]/[O_T]$ of adding 100 stronger binding sites to the cell, in addition to the 10^7 sites/cell having $K_{RD} = 10^5$ M^{-1}. Such stronger sites must exhibit repressor affinities in excess of about 10^{11} M^{-1} if they are to have a significant effect on the basal level. A small number of sites with K_{RD} as large as 10^{13} M^{-1} would have relatively little effect on $[O]/[O_T]$.¶

¶ Gilbert and Reznikoff (personal communication cited in ref. 10) have both reported the existence of at least one site with $K_{RD} \simeq 10^{13}$ M^{-1} in the *z*-gene region of the *lac* operon.

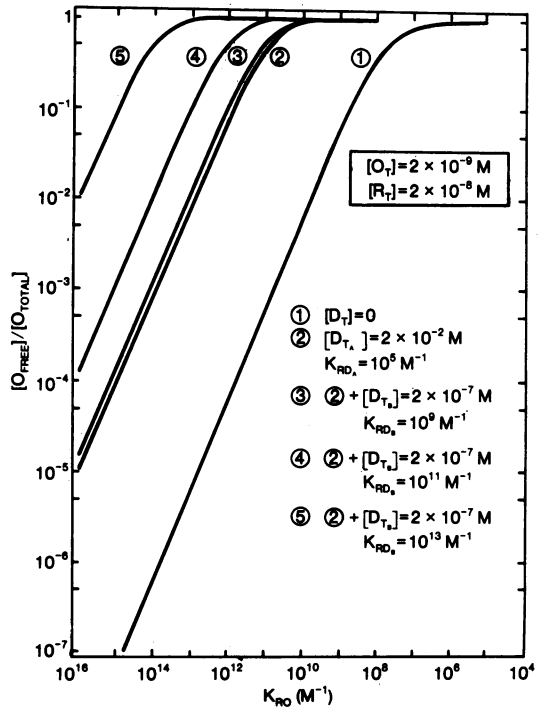

FIG. 3. Effect of non-specific binding sites on the fraction of free operator for various values of the repressor–operator association constant (K_{RO}). The effects of a second class of non-specific sites, with various association constants, are also illustrated.

Operator-Constitutive (O^c) *Mutations.* Fig. 3 also shows that decreases of one to two orders of magnitude in K_{RO} (O^c mutations) would still not bring $[O]/[O_T]$ to the observed wild-type basal level in the absence of competitive *R* binding to non-specific sites (curve 1), but would indeed show just the effects observed in the basal level with O^c mutants (fact *f*, above) when non-specific binding is included (curve 2).

Effects of Inducers. Fig. 4 shows the effects of inducer binding on the calculated fraction of free *lac* operator. From the lower (solid) curve it can be seen that, in the absence of non-specific binding of repressor and repressor–inducer complexes, derepression of the *lac* operon, even at super-saturating concentrations of *I*, cannot take place. This follows because the (*I*-saturated) RI_n complex binds to *O* with only about three orders of magnitude less affinity than does *R* itself (fact *g*, above); thus in the absence of the non-specific sites as "sink" for the RI_n complex, the latter itself binds to *O* and maintains repression at calculated values *below* the basal level.

The upper curves in Fig. 4 show the effects of adding inducer when the binding of *R* and RI_n to non-specific sites is included; calculations have been made for three different values of K_{RD}. We note that only the curve for $K_{RD} = 10^5$ M^{-1} starts at the observed basal level at low $[I_T]$, and reaches approximately the constitutive level (within a factor of two) at saturating inducer concentrations. This observation provides an independent confirmation that the value of K_{RD} chosen for the "base-set" of parameters is indeed a reasonable representation of the *in vivo* situation or, more accurately, that the *ratio* of K_{RO} to K_{RD} is approximately correct. Fig. 4 also shows that derepression of the *lac* operon occurs at the inducer concentrations observed *in vivo* (fact *e*, above) when non-specific binding of *R* and RI_n is taken into account, but not otherwise.

Finally, we have performed a series of calculations (not shown) in which the ratio of K_{RIO}/K_{RO} to K_{RID}/K_{RD} was varied. It can easily be shown by such procedures that if $K_{RIO}/K_{RO} > K_{RID}/K_{RD}$, the effect of increasing inducer concentrations is to further *repress* the *lac* operon. Thus, a mutational alteration in repressor structure producing such changes in the binding constants could provide at least a partial explanation of the behavior of the i^r mutant phenotype (11, 13).

II. Application to RNA polymerase–sigma factor–promoter–non-specific DNA interactions

Chamberlin and co-workers (14, 15) have made an extensive study of the binding of RNA polymerase [in both the "core-" and the "holo-" (core plus σ-factor) enzyme forms] to promoters and to non-specific DNA. The possible relevance of non-specific DNA binding to the control of polymerase function was certainly, at least qualitatively, a consideration in motivating their studies. For our purposes their most important result is that the holoenzyme appears to bind about 10^3 times less tightly to nonpromoter-containing DNA sequences than does the core enzyme. This suggested to us (in partial analogy to the induction of the *lac* system), the simplistic notion that sigma factor might serve as a general control element in RNA polymerase function by "inducing" the polymerase off non-specific DNA binding loci, thus increasing its availability to promoter sites. In this way, by controlling the concentration of active sigma factor, the cell could exert general transcriptional control over the intracellular level of protein synthesis.

We have applied the computational approach and model outlined in Eqs. 1 and 2 and Fig. 1 to examine the RNA polymerase system, substituting σ-factor for inducer, core enzyme for repressor, and holoenzyme for the repressor–inducer complex. Thus, in this Section, C represents core (RNA) polymerase, σ is sigma subunit, $C\sigma$ is holoenzyme, P represents promoter sites, and D (as before) indicates non-specific DNA sites. Mass action equations and conservation constraints corresponding to Eqs. 1 and 2 were used, assuming the following "base-set" parameters for the *E. coli* system *in vivo*: $K_{C\sigma P} = 10^{14}\ M^{-1}$ (for the final, melted-in, complex); $K_{C\sigma D} = 2 \times 10^8\ M^{-1}$; $K_{CD} = K_{CP} = 2 \times 10^{11}\ M^{-1}$; $K_{C\sigma} = 10^{10}\ M^{-1}$; $[P_T] = 2 \times 10^{-7}$ M (about 100 promoters per cell); $[C_T] = 1 \times 10^{-5}$ M (about 5000 core enzymes per cell); $[D_T] = 2 \times 10^{-2}$ M; and $[\sigma_T] = 1 \times 10^{-5}$ M (about 5000 σ-subunits per cell)‖.

Computations were made using both the "base-set" parameters and some variations of these quantities to indicate the nature of the resulting changes. The significant finding is that when polymerase is present predominantly in the "core" form (due, perhaps, to a low concentration of active σ-factor in the cell) it would be largely unavailable to promoter sites (P sites

‖ Several of these assumed "base-set" parameters are quite arbitrary, since the measurements available are not as complete as those for the *lac* repressor system. The "base-set" values for K_{CD}, $K_{C\sigma D}$, and $K_{C\sigma P}$ are a selfconsistent equilibrium set measured by Hinkle and Chamberlin (14) for conditions comparable to those under which the *lac* repressor "base-set" was determined. It is assumed that $K_{CP} = K_{CD}$. No measurements for $K_{C\sigma}$ or $[\sigma_T]$ in the *E. coli* cell are available, so they are treated as adjustable parameters to attain a reasonable degree of $C\sigma$ complex formation.

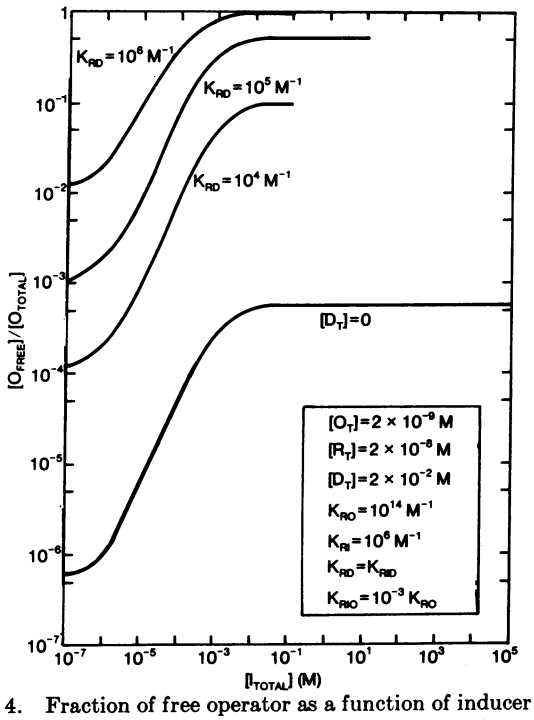

FIG. 4. Fraction of free operator as a function of inducer concentration, at several values of the association constant for repressor and repressor-inducer complex to non-specific DNA sites.

80 to > 99% free), because of the competition of the non-specific sites. This competitive binding to non-specific sites is much less for the holoenzyme, so in the presence of excess σ-factor the promoters are largely saturated.

III. Repressor–polymerase interaction at the *lac* operon

We now combine the above approaches to consider quantitatively the interaction of *lac* repressor and RNA polymerase in regulating the function of the *lac* operon. It is clear that a complete description of *lac* operon control will require appreciably more data for the entire RNA polymerase system; nevertheless it is instructive to consider some minimal models.

(*1*) *Non-competitive Binding.* In this model it is assumed that the nucleotide sequences which define the operator and promoter do not overlap, and that repressor and polymerase bind independently. Thus, the presence of bound R on O does not interfere with the formation of a stable, "melted-in" initiation complex, but totally prevents extension of the mRNA transcript through the operator region into the *lac* genes. (This model is implicitly assumed in the arguments of the previous sections.) For this model (as shown above) the basal level of *lac* enzyme synthesis should be simply the product of the probability that the operator is not covered by repressor, and the probability that the promoter is complexed with a functional polymerase. The latter probability enters equally into the constitutive rate, so for this model $(Z_{\text{basal}}/Z_{\text{const.}}) = [O]/[O_T]$. From Fig. 2 or 3 we see that for our "base-set" parameters for the repressor–operator system, the non-competitive model predicts that $(Z_{\text{basal}}/Z_{\text{const.}}) \simeq 10^{-3}$.

(*2*) *Competitive Binding.* The operator and promoter sequences are assumed to overlap, and thus repressor and polymerase compete for the same binding site. Assuming equilibrium, i.e., that the rate constant for product (mRNA) forma-

tion is small relative to rate constants for the formation and dissociation of the polymerase–promoter ("enzyme-binding site") and the repressor–operator ("inhibitor-binding site") complexes, we may apply the relations of competitive inhibition from enzyme kinetics to obtain:

$$(Z_{\text{basal}}/Z_{\text{const.}}) = \frac{1 + ([C\sigma]/K_{C\sigma P})}{1 + ([C\sigma]/K_{C\sigma P})\,(1 + K_{RO}\cdot[R])} \quad [3]$$

We assume that all the RNA polymerase is present as the holoenzyme. The concentrations of free holopolymerase ($[C\sigma]$) and free repressor ($[R]$) are controlled ("buffered") by the non-specific binding equilibria; these quantities have been evaluated using the procedures outlined in Sections I and II. For our "base-set" binding constants and constraints, we find, using Eq. 3, that $(Z_{\text{basal}}/Z_{\text{const.}}) \simeq 0.22$. This value is clearly too large, but the discrepancy should not be taken too seriously in view of the approximate nature of the polymerase–promoter "base-set". Calculations made with Eq. 3 are much more sensitive to the exact values used for binding constants, etc. than are the conclusions reached in Sections I and II. Thus, a decrease in the assumed values of $K_{C\sigma}$ or $[\sigma_T]$, or a lower value of $K_{C\sigma P}$ could all decrease $(Z_{\text{basal}}/Z_{\text{const.}})$ calculated for this model to a more reasonable value.

We emphasize that for this (competitive binding) model, repression would be virtually impossible in the absence of non-specific binding, since the free concentration of polymerase would then greatly exceed that of repressor. It is only the fact that polymerase (in both its forms) binds much more strongly than repressor to non-specific sites which reduces the ratio of free polymerase to repressor concentration to manageable levels.

(3) *"Interactive" Binding.* In this model the operator and promoter sequences do not overlap, but the binding of polymerase is modulated by the binding of repressor. We have found that repressor is a double-helix stabilizing protein, i.e., it increases the melting temperature of native DNA to which it is bound (1). Thus, the presence of repressor on the operator could inhibit the formation of a "melted-in" polymerase complex at an adjacent promoter site, even in the absence of sequence overlap. In the limit of complete inhibition, this model would resemble the competitive binding situation; partial inhibition would result in a situation intermediate between models (*1*) and (*2*). In either case, this model would be characterized by a distinctive temperature dependence.

These models are in principle distinguishable because they predict different responses to changes in parameters such as component concentrations, binding constants, temperature, etc. At present there are insufficient data available to exploit this possibility, but the quantitative formulation presented above can serve as a guide for experiments to discriminate between the models.

IV. Conclusions

In vitro measurements indicate that regulatory proteins will be bound nonspecifically to the *E. coli* chromosome *in vivo*. Our calculations, based on simple equilibrium considerations, imply that the observed levels of repression of the *lac* operon under a variety of physiological and genetic conditions can be rationalized when non-specific binding is taken into account, and that such binding must be considered as an integral component of any quantitative model in which the concentration of free protein is an essential parameter.

While lack of knowledge of the exact internal composition of the cell injects some uncertainty into these computations, the quantitative agreement demonstrated with *in vivo* results is not greatly affected by changes in the values of the "base-set" parameters. From Fig. 3, for instance, it is clear that if the value of K_{RO} is in fact 10^{13} M^{-1} (instead of 10^{14} M^{-1}, as we have used) the predicted basal level of repression in the absence of non-specific binding will still be far below that observed. Furthermore, we note that even if other regulatory proteins were to bind tightly to as much as 90% of the genome, thus making this portion of the DNA unavailable for non-specific repressor binding, the effect on the calculated basal level would be relatively small (see Fig. 2, solid curve; compare $[O]/[O_T]$ for $[D_T] = 10^{-3}$ M and 10^{-2} M). Similarly, other non-specifically binding components in the cell, e.g., polyamines, divalent cations, etc., should reduce K_{RO} and K_{RD} (and the equivalent polymerase binding constants) approximately in parallel, and thus have little effect on the *ratio* of these constants, which is the crucial parameter in this model for regulation.

It is tempting to extend these ideas to a consideration of possible roles that non-specific binding might play in the control of the function of regulatory proteins in eukaryotic cells. For example, these results suggest that competition between histones and regulatory proteins for non-specific, as well as for specific, DNA sites in such cells must be considered as a component of the complete regulatory system.

We are grateful to colleagues in our laboratory and in the Institute of Molecular Biology for stimulating discussions of these matters; we thank Dr. James McGhee also for help with the initial stages of the computer programming. The work was supported in part by USPHS Research Grant GM-15792 and GM-15423, a John Simon Guggenheim Memorial Research Fellowship (to P.H.v.H.), and USPHS Post-Doctoral Fellowships GM-55928 (to A.R.) and GM-43987 (to C.A.G.).

1. von Hippel, P. H., Revzin, A., Gross, C. A. & Wang, A. C. (1974) in *Symposium on Protein-Ligand Interaction* (W. de Gruyter, Berlin), in press.
2. Jacob, F. (1966) *Science* **152**, 1470–1478.
3. Monod, J. (1966) *Science* **154**, 475–483.
4. Gilbert, W. & Müller-Hill, B. (1970) in *The Lactose Operon*, eds. Beckwith, J. R. & Zipser, D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), pp. 93–109.
5. Riggs, A. D., Bourgeois, S., Newby, R. F. & Cohn, M. (1968) *J. Mol. Biol.* **34**, 365–368.
6. Jacob, F. & Monod, J. (1961) *Cold Spring Harbor Symp. Quant. Biol.* **26**, 193–211.
7. Sadler, J. R. & Novick, A. (1965) *J. Mol. Biol.* **12**, 305–327.
8. Bourgeois, S. & Jobe, A. (1970) in *The Lactose Operon*, eds. Beckwith, J. R. & Zipser, D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), pp. 325–341.
9. Smith, T. F. & Sadler, J. R. (1971) *J. Mol. Biol.* **59**, 273–305.
10. Jobe, A., Sadler, J. R. & Bourgeois, S. (1974) *J. Mol. Biol.* **85**, 231–248.
11. Jobe, A. & Bourgeois, S. (1972) *J. Mol. Biol.* **72**, 139–152.
12. Riggs, A. D., Suzuki, H. & Bourgeois, S. (1970) *J. Mol. Biol.* **48**, 67–83.
13. Myers, G. L. & Sadler, J. R. (1971) *J. Mol. Biol.* **58**, 1–28.
14. Hinkle, D. C. & Chamberlin, M. J. (1972) *J. Mol. Biol.* **70**, 157–185.
15. Chamberlin, M. J. (1974) *Annu. Rev. Biochem.* **43**, 721–775.