(*19*). Professional opinions doubtless vary regarding the judgment used in selecting sites for extended investigation, and on other points. In general, the work has been carried on with a sense of scientific problem, though one wonders whether the possibilities of a full interdisciplinary approach were properly considered. As nonfederal participation increased and more cooperating agencies joined in the concentrated attack on the Middle Missouri, other investigators have found it possible to resume their researches elsewhere in the basin and not necessarily at sites directly or immediately threatened by the federal water-control program. This result has provided a healthy balance—not all our archeologic problems can be solved in the reservoir areas, though the work there may well suggest where researches are needed outside. Moreover, expanding agricultural and other activities are probably as destructive of sites in the long run (though not as spectacular) as is flooding.

In retrospect, the archeologic salvage program in the Missouri River Basin must be reckoned a major advance toward our understanding of human prehistory and its development in the region through some thousands of years. Any shortcomings in planning or performance are perhaps the price we must pay to learn that a tight administrative rein must be combined with a scholarly approach to scientific problems; that high standards of planning, staffing, field work, and reporting must be maintained at all times; that addition of any major field-research program to an existing agency's regular operations must be balanced by expansion in publications, with appropriate funding and expert supervision; and that proper public relations and free exchange of information and ideas must include the public and professional colleagues no less than immediate associates. When the vast quantities of raw data, artifacts, and records accumulated in the Missouri Basin salvage work have been critically studied and reported according to the highest possible standards of scholarship, American archeology will benefit from a store of new and important information much of which, but for the salvage effort, would have been lost forever.

References and Notes

1. F. Johnson, *Science* **152**, 1592 (1966).
2. W. D. Strong, in *Smithsonian Inst. Misc. Collections* **100**, 353 (1940); W. R. Wedel, *ibid.*, p. 291.
3. W. D. Strong, *ibid.* **93**, 224 (1935).
4. W. T. Mulloy, *Wyoming Univ. Publ.* **22** (1958).
5. H. L. Ickes, *Senate Doc. 191* (78th Congress, 2nd sess., 1944).
6. J. O. Brew, Ed., in *Amer. Antiquity* **12**, 209 (1947).
7. F. H. H. Roberts, Jr., in *78th Ann. Rept. Bur. Amer. Ethnol.* (1962), p. 21.
8. ———, Ed., *Bull. Bur. Amer. Ethnol. 176* (1960).
9. J. Deetz, *Illinois Studies in Anthropol. No. 4* (1965).
10. D. J. Lehmer, *Bull. Bur. Amer. Ethnol. 158* (1954).
11. W. R. Wedel, *ibid. 174* (1959), p. 557.
12. ———, *Amer. Anthropol.* **55**, 499 (1953).
13. M. F. Kivett, in *Bull. Bur. Amer. Ethnol. 154* (1953), p. 103; R. W. Neuman, *Amer. Antiquity* **26**, 78 (1960); G. W. Hewes, *ibid.* **14**, 322 (1949).
14. J. T. Hughes, *Amer. Antiquity* **14**, 266 (1949).
15. P. Holder and J. Wike, *ibid.*, p. 260 (1949); E. M. Davis, *Univ. Nebr. State Museum Spec. Publ. 3* (1962).
16. For earlier summaries see W. R. Wedel in *Bull. Bur. Amer. Ethnol. 154* (1953); P. L. Cooper, in *Smithsonian Inst. Misc. Collections* **126** (1955).
17. See River Basin Surveys statements in *Ann. Rept. Bur. Amer. Ethnol. 64th* (1946–47) (and since); F. H. H. Roberts, Jr., *Smithsonian Inst. Ann. Rept. 1960* (1961), p. 523.
18. R. L. Stephenson, *Amer. Antiquity* **28**, 277 (1963).
19. See R. F. Heizer, *The Masterkey* **40**, 54 (1966).

20. I thank T. D. Stewart, J. C. Ewers, and M. M. Wedel for criticizing the manuscript; and W. W. Caldwell, River Basin Surveys, for Figs. 3–9. All photographs by the Smithsonian Institution's River Basin Surveys.

Regulation of the Lac Operon

Recent studies on the regulation of lactose metabolism in *Escherichia coli* support the operon model.

Jonathan R. Beckwith

The mechanism of regulation of gene expression is today one of the most actively studied problems in molecular biology, in good part as a result of the pioneering work of Jacob and Monod on the control of the genes involved in lactose metabolism in the bacterium *Escherichia coli*. Since 1961, when Jacob and Monod first proposed the operon model for gene regulation (*1*), a number of alternative suggestions have been published for the ways in which genes are controlled (*2–5*), some of them radically different from the Jacob-Monod model.

The lactose (*lac*) system is still the best-studied example of gene regulation. In the years since 1961, there has been a considerable amount of new information, both genetic and biochemical, on the *lac* operon. On the basis of the most recent information, which will be discussed in this article, it appears that the original formulation of the operon, in most of its aspects, is still the simplest model fitting all the known facts about the *lac* system (*6*).

The author is an assistant professor in the Department of Bacteriology and Immunology, Harvard Medical School, Boston, Massachusetts.

Lactose Metabolism in Escherichia coli

The initial steps in the metabolism of lactose by *E. coli* involve two protein components: (i) a membrane-bound protein [M-protein (*7*) or permease (*8*)] which is probably responsible for both the transport of lactose into the bacterial cell and for its concentration therein; and (ii) the enzyme β-galactosidase which catalyzes the hydrolysis of lactose within the cell to glucose and galactose. The structure of these two proteins is determined by two chromosomal genes, *y* for the permease and *z* for β-galactosidase. In wild-type strains of *E. coli* grown on almost any carbon source but lactose, the activities of these genes are repressed, their products being found in only very small amounts. However, growth on lactose as sole carbon source, or addition to the growth medium of various compounds structurally related to lactose, results in the induction of gene expression, with an increase in

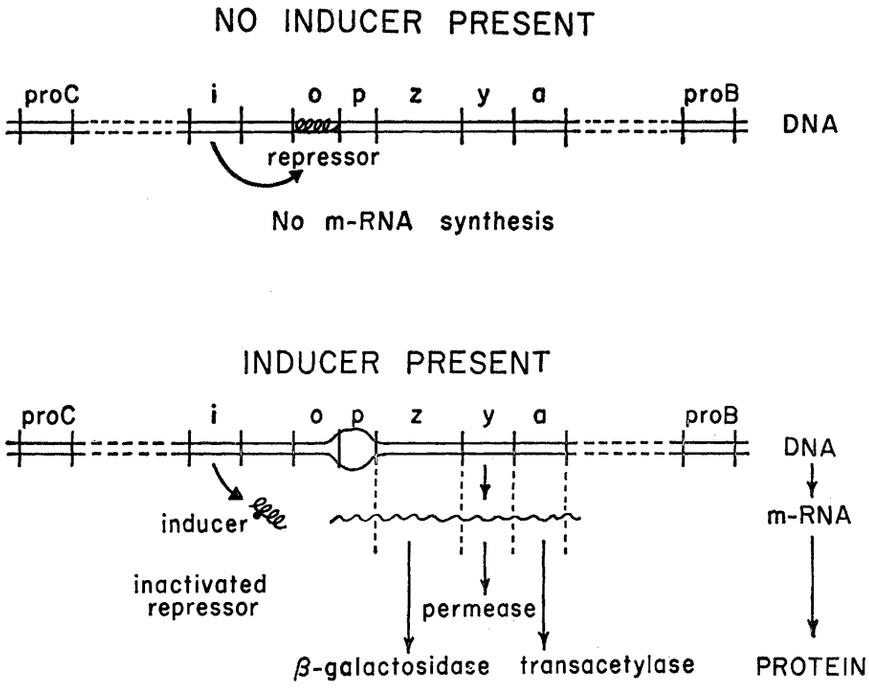

Fig. 1. The control of expression of the *lac* operon. The *lac* region and two of the markers surrounding it (genes involved in proline biosynthesis) on the chromosome are represented as regions of double-stranded DNA. The picture of the operator and promoter in this figure is only one of many ways of visualizing their interaction. The promoter is shown serving as an initiation site for mRNA synthesis only when its two DNA strands are not held together by hydrogen bonding in the "closed" configuration. The repressor binds to and closes the operator, resulting in a closing of the promoter. An inducer alters the repressor so that it can no longer interact with the operator in this way. Pro C and pro B indicate two genes involved in the biosynthesis of proline.

the amounts of these proteins of as much as 1000 times. Under these optimum conditions, β-galactosidase represents approximately 3 percent of the total protein of the cell. Induction also results in the appearance of another enzymic activity, thiogalactoside transacetylase (*9*), corresponding to a third structural gene, *a*. Although the *a* gene is regulated in parallel with the *z* and *y* genes, the enzyme plays no essential role in lactose metabolism (*10*). These three structural genes of the lactose system lie next to one another on the chromosome, mapping in the order *z-ya* (Fig. 1).

The Model

The Jacob-Monod operon model of control, with some additions and modifications resulting from recent work on the *lac* system (Fig. 1) may be described as follows. In the absence of any regulation, the expression of the three *lac* structural genes involves two steps. First, the information from these genes is transcribed into a single RNA (messenger RNA) molecule. The information from each gene-copy within the RNA is then translated by the protein-synthesizing machinery into the structure of the three protein products. The synthesis of this mRNA molecule is initiated at the promoter site, *p*, which is adjacent to, or part of the *z* gene. The structure of the promoter, or of a site in the same region, determines the rate at which the message is transcribed from these genes under conditions of maximal expression. The control of the expression of the *lac* structural genes is effected by a repressor molecule, the protein product of the closely linked *i* gene. The repressor acts on the DNA to inhibit the transcription of these genes. The combination of the repressor with the operator site, *o*, which is adjacent to *p*, in some way inhibits the initiation process. Compounds that cause induction of the expression of the *lac* genes act by either destroying or altering the repressor-operator complex, thus allowing initiation of mRNA synthesis at the promoter.

A group of genes whose activity is coordinated by an operator is termed an operon. According to this model, the operator is not essential for operon activity, but rather serves as a controlling site superimposed on a functioning unit.

Mapping of the Elements of the Lac Operon

Jacob and Monod have presented evidence from three-factor crosses demonstrating the gene order *iozy* (*11*). In addition, strains which carry deletions extending into the *lac* genes from outside either end of the *lac* operon have been isolated, providing further confirmation for the order *iozya* (see Fig. 4) (*10, 21*). Class-I deletions remove the *i* gene and leave the operator and structural genes intact, while class-IVa deletions remove the *a* gene but leave all other sites in the operon intact.

Translation of the Lac operon

The translation process. To discuss certain aspects of *lac* operon control, I must first describe what is known about the mode of translation of mRNA information into protein. The mRNA copies of genes are thought to be translated by the following process (*13*). Ribosomes attach to one end of the message (corresponding to the amino-terminal end of the protein), and, in conjunction with the other components of the protein-synthesizing machinery, begin to move along the message as the peptide bonds are formed. New ribosomes continually attach to this end of the message and proceed in this way, so that at any one time the mRNA will carry many ribosomes along its length. This complex of ribosomes and mRNA is called a polysome. At the end of the gene-copy in the mRNA, the translation machinery meets a codon that signals termination and release of the completed polypeptide chain.

As a result of mutation, chain-terminating codons may be introduced into the gene at various points preceding the normal site of chain termination. The two well-studied chain-terminating codons arising by mutation are the amber and ochre codons, UAG and UAA (for uracil, adenine, and guanine) respectively (*14, 15*). Chain termination by an amber mutation (and probably by an ochre mutation also) results in a quantitative release of an

NH_2-terminal fragment of the protein coded for by a particular gene. The length of the protein fragment depends upon the distance of the mutation from the beginning of the gene (corresponding to the NH_2-terminal end of the protein).

Another type of mutation that interferes with translation is the frame-shift mutation (*16*). These mutations, through the addition or removal of one or more base pairs from a gene, cause the reading of an incorrect sequence of codons in the mRNA gene-copy.

The operon messenger RNA. There is still no direct evidence that all the information from the structural genes of the *lac* operon is contained in a single piece of mRNA. However, studies on polarity (described later) do suggest that this is so. Also, Kiho and Rich (*17*) have presented evidence that an amber mutation prematurely terminating translation in the *y* gene affects the size of the polysome on which β-galactosidase is made. If the *y* and *z* gene-copies were on different mRNA molecules, no effect should have been observed. We shall assume that there is a single operon mRNA.

The amino-terminal end of the z *gene.* A knowledge of the direction of translation of the structural genes of the operon is important in interpreting the various experiments on operon expression. The direction of translation indicates the direction of transcription of the operon, since it has been shown by biochemical and genetic experiments that both these processes begin at the 5′ end of the RNA molecules (*18*). Studies by Fowler and Zabin (*19*) on three chain-terminating mutants of the *z* gene (*20*) indicate that translation is initiated at the proximal (*21*) end of the *z* gene-copy in the mRNA. These mutants, one of which lies very close to the *y* gene, map at the distal end of the *z* gene.

Each mutant makes a large amount of a protein that is immunologically similar to β-galactosidase. If this distal end of the *z* gene corresponded to the NH_2-terminal end of the protein, translation should have begun at this end, and, in these mutants, terminated within a short distance, releasing a small polypeptide fragment. The finding of large amounts of a large protein molecule is thus very strong evidence that translation begins at the proximal end of the *z* gene, and that in these mutants, most of the *z* polypeptide chain is made before the chain-terminating triplet is read and the protein released. As mentioned earlier, this conclusion also leads to the further conclusion that operon mRNA synthesis is initiated at this end of the operon. Even stronger evidence for the operator-proximal end of a gene corresponding to the NH_2-terminal end of the protein comes from studies on the tryptophan operon by Yanofsky and co-workers (*22*).

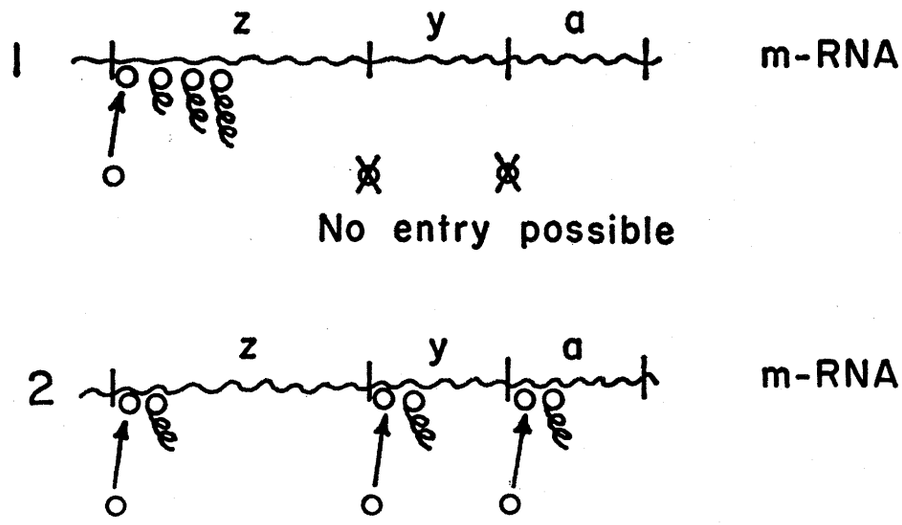

Fig. 2. Two possible models for translation of the *lac* operon mRNA. The circles represent ribosomes, and the squiggly lines attached to them represent the growing polypeptide chain.

Polar Mutants

Two different models for the translation of the gene-copies in an operon mRNA have been considered (Fig. 2). (i) A ribosome can enter onto the mRNA at only one end of the molecule (*2*); in the *lac* operon, this end would correspond to the NH_2-terminal end of the *z* gene. Then, during translation, the ribosome proceeds down the message and must complete passage through the *z* gene in order to start translation of *y* and *a*. (ii) Ribosomes can enter independently at the starting points for all three gene-copies in the mRNA, *z*, *y*, and *a*, without any requirement for having translated a previous gene-copy (*23*).

There is no conclusive evidence to distinguish between the two models. However, any model for the translation of the *lac* operon mRNA must take into account the class of mutants known as polar mutants. Polar mutations in the *z* gene of the *lac* operon are usually point mutations that not only abolish *z* gene activity, but, in addition, reduce or abolish the expression of the *y* and *a* genes (*24, 25*). Such mutations in the *y* gene affect the activity of the *a* gene but not of the *z* gene. Although many polar mutants are not well characterized, a large percentage are amber and ochre mutations (*23*). Since these chain-terminating mutations exert their effects on translation, the polarity must be due primarily to an interference with the translation of the operon mRNA.

The polarity of a chain-terminating mutation in the *z* gene depends upon its position within the gene (Fig. 3) (*23*). Amber and ochre mutants mapping toward the distal end of the *z* gene are not very strongly polar, whereas those chain-terminating mutants mapping in a proximal segment of the *z* gene do not make any detectable amounts of the *y* and *a* gene products. The extent of polarity of a chain-terminating mutation in *z* is determined by the distance of the mutation from the boundary between *z* and *y*, and not by its proximity to the operator. Thus, when an extremely polar mutation, mapping at the proximal end of the *z* gene, is combined with a succeeding deletion in the gene, effectively moving the mutation closer to the *y* gene, the polarity effects are markedly reduced (*23, 26, 27*). The distance from the operator end of the *z* gene to this mutation has not changed. This finding shows that, even if translation of the operon is terminated by a chain-terminating mutation very shortly after it begins, the rest of the operon can still be expressed if

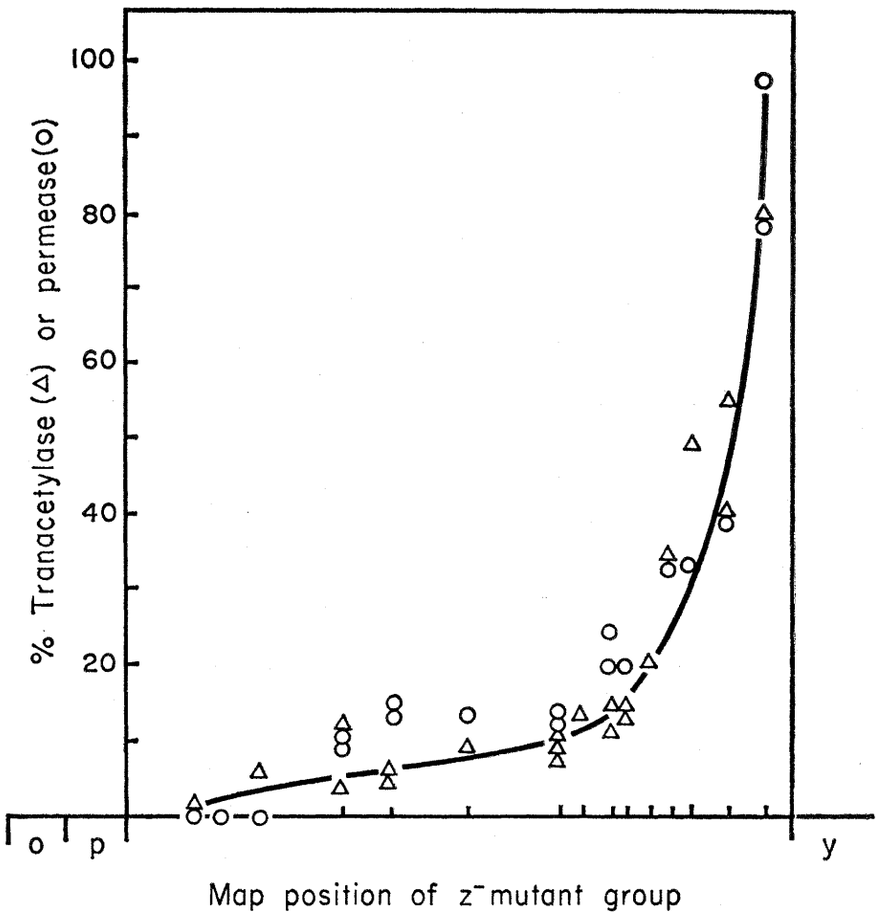

Fig. 3. The gradient of polarity of chain-terminating mutants in the *z* gene. This figure is copied from Newton *et al.* (*23*). The permease activity (○) and transacetylase activity (Δ) of many z^- amber and ochre mutants are plotted against the position within the *z* gene. These positions are only rough approximations of true location on the genetic map location. However, the order of the mutants is unambiguous.

this mutation is very close to the *y* gene.

Polar mutations in the *z* gene not only reduce the amounts of the proteins, permease and transacetylase, but also reduce the amount of *lac* mRNA present in the cell (*28*). Again, the reduction of the amount of *lac* mRNA is correlated with the position of the mutation in the *z* gene. Extracts of induced cells carrying one of the extremely polar mutations at the proximal end of the *z* gene contain no detectable amounts of *lac* mRNA; cells harboring z^- polar mutations closer to the *y* gene contain detectable, but still reduced, levels of *lac* mRNA. There is some indication that in these mutants there are normal amounts of the fragment of *lac* mRNA corresponding to the proximal segment of the *z* gene preceding the mutation, but that complete operon mRNA molecules are present only in reduced amounts. Thus, as a result of the interference by polar mutations with the translation process, a polarity effect on *lac* mRNA levels also results.

The results of studies on polar mutants fit both of the models for operon mRNA translation considered above. If the ribosomes can only enter at the beginning of *z* (model 1), then, after encountering an amber or ochre mutation in the *z* gene-copy, ribosomes would continue to pass down the message. However, since these ribosomes are no longer engaged in translation, there is an increased probability of their falling off the mRNA. The longer the distance to be traversed after the site of chain-termination, the lower the probability that the ribosome will reach the *y* gene-copy and there initiate translation. Recent studies of Malamy (*29*) on the strong polarity effects of what appear to be frame-shift mutations in the distal segment of the *z* gene suggest that ribosomes after encountering chain-terminating mutations do continue moving along the mRNA. He finds that double mutants, in which the "frame-shift" is preceded by an ochre or amber mutation in *z*, are still just as polar as the original frame-shift mutant. If ribosomes fell off the mRNA at chain-terminating codons, they should never encounter the frame-shift, and the polarity effects should be reversed. However, it should be noted that apparent frame-shift mutations in one gene of the histidine operon in *Salmonella typhimurium* do not have the strong polarity effects seen in similar mutants in the *lac* operon (*30*). Confirmation of the proposal that the extreme polarity effects of these mutants in the *lac* operon are due to their frame-shift character would be strong supporting evidence that ribosomes continue their movement after chain-terminating codons.

If ribosomes cannot continue along an mRNA after a chain-terminating codon, then, in order to explain why most amber and ochre mutants allow some expression of *y* and *a*, we must admit that ribosomes can enter the *y* and *a* gene-copies independently of the completion of passage of the *z* gene-copy (model 2). The polarity effects of chain-terminating mutants can be accounted for by one of the following additional hypotheses.

1) Initially, the operon mRNA sticks to the DNA. As ribosomes move down the mRNA during translation and get closer to the beginning of the *y* gene-copy, there is an increasing chance that they will cause release of this part of the mRNA, thus freeing it for ribosome entry (*23*). Premature chain-termination in the *z* gene and falling off of ribosomes would make it more likely that the mRNA will stick to the DNA.

2) The introduction of a chain-terminating mutation in the *z* gene, leaving an untranslated portion of the mRNA, results in a more rapid destruction of the distal segment of the mRNA (*15, 31*). The amount of destruction could depend upon the segment of mRNA which is not engaged in translation.

3) The secondary structure of the RNA surrounding the ribosome entry points for *y* and *a* is such that ribosomes cannot enter unless there has been enough progress of ribosomes on the *z* gene-copy to disrupt this structure (*32*). Premature chain-termination in the *z* gene and falling off of ribosomes would reduce the probability of exposing the entry site on *y*.

Although all of these hypotheses must

include some explanation for the decreased amounts of *lac* mRNA in polar mutants, it is not a very important consideration in formulating the models. If this decrease in mRNA is not directly predicted by the model for polarity, then the suggestion can always be added that, under the particular conditions, the mRNA is destroyed very rapidly. If, of course, it is proved that the mRNA is not rapidly destroyed in these mutants, then certain models of polarity must be discarded.

Certain important findings do come out of the studies on polarity. First, the conclusion that the amount of *lac* mRNA in a cell somehow depends upon the translation of the operon has led to new speculations on mechanisms of repression of operon activity involving effects on translation. Second, it can be concluded that translation of an early part of the *z* message is not necessary for the release of the operon mRNA from the DNA, except in so far as this region is distant from the beginning of the *y* part of the mRNA. Stent (*3*), Yanofsky (*33*), and others have suggested that there is a critical region at the beginning of the first gene-copy in an operon mRNA, and this region must be translated in order to release the entire mRNA from the DNA. In the *lac* operon, this model cannot account for the extreme polar mutants.

Sequential Expression of the Lac Genes

When the kinetics of expression of the *lac* operon are followed after the addition of inducer to a culture of *E. coli*, β-galactosidase activity begins to increase a minute or so before transacetylase activity rises (*34, 35*). Since the appearance of an enzymic activity requires transcription, translation, and assembly of a protein into the correct configuration, any or all of these processes could be responsible for this lag. Experiments designed to determine the basis of this lag do show that there is a measurable time lapse between the transcription of the *z* gene and of the *a* gene (*35, 36*). However, it is not clear yet whether this delay is responsible for the lag in appearance of transacetylase activity. In any case, the results showing the sequence of transcription of the *lac* genes are another indication that transcription begins at the operator end of the operon.

Unbalanced Translation of Lac Genes

Zabin has shown that the molar amounts of transacetylase synthesized are five or more times lower than the molar amounts of β-galactosidase (*37*). One explanation for this finding derives from the studies on polar mutants which show that the distance between a chain-termination event in one gene and the initiation of translation in the next determines how much of the next protein is made. It is possible that, after the chain-terminating triplet in either *z* or *y*, there is a "dead space" of untranslated nucleotide sequence, the length of which determines the amount of transacetylase made.

The Operator

I have described evidence that the transcription of the *lac* operon starts at the operator end. But is transcription actually initiated at the operator itself? In the original operon hypothesis, it was proposed that the operator has three functions. (i) It forms the first part of the *z* gene; (ii) it is the site of repressor action; and (iii) it is the site of initiation of transcription. The second function is confirmed by the existence of operator-constitutive mutants (O^c) (*1*), which have partly lost sensitivity to repressor. As a result, these mutants make large amounts of the *lac* enzymes in the absence of inducer, but can still be induced to make more. The nature of the operator is indicated by the finding that O^c mutations cause constitutivity only for the *lac* genes which are linked to that operator. Thus, in a partial diploid strain carrying two copies of the *lac* region, one of which is O^c, the genes attached to the O^+ operator are still normally repressible, while those linked to the O^c operator are still partially constitutive.

There is reason to believe that all O^c's are deletions (*38*). First, certain of the O^c's can be shown to be deletions, since they remove the *i* gene also (as discussed later) (*38*). Second, the frequency of O^c's is not increased by mutagens which only cause base substitutions, but it is increased by treatment with x-rays, which does cause deletions (*38*). Third, no suppressible O^c's have yet been found (*39*). With another operon there is also mutagenic evidence that O^c's are always deletions (*40*).

The third function for the operator was based on the existence of a second class of mutants ("O^0") mapping at the beginning of the *z* gene, which permanently shut off the *z*, *y*, and *a* genes (*1*). However, it was subsequently shown that complete activity is restored to the *y* and *a* genes by deletion of the "O^0" mutant site (*41*). Thus, this region of the operon is not necessary for operon expression and cannot be the site of initiation of mRNA transcription for the operon. In addition, it was shown that "O^0" mutations do not, in fact, lie in the operator region, but are only extreme examples of the polar mutants (*15, 23, 31, 41*). Therefore, the term "O^0" is a misnomer, since these mutants have nothing to do with operator control. As a result of these studies and further work indicating that O^c mutations do not affect the properties of β-galactosidase (*38, 41, 42*), it has been concluded that the operator lies outside the first structural gene of the operon.

The Promoter

The following discussion summarizes evidence which suggests that transcription begins not at the operator, but at the adjacent promoter site. It is an elaboration of the argument for the promoter presented by Jacob, Ullman, and Monod (*38*). This concept of operon function is based on the properties of mutants in which various components of the *lac* operon are deleted. Starting with strains in which the *lac* operon has been inactivated by certain mutations, it is possible to select O^c mutations that render the operon partially or fully constitutive (*38*). In addition, by selecting for O^c mutations under conditions where only *y* gene function is required, O^c's are found which are z^-y^+ (*38, 41*). In all cases examined, O^c mutations of this latter type (Fig. 4, class III) are a result of deletions which cover a proximal segment of the *z* gene, the operator, and the *i* gene, and presumably fuse the intact *y* and *a* genes to another operon or gene. In most cases, the gene or genes to which *lac* has become fused have not been identified. However, two sets of deletions of this type have been isolated in which *lac* has been fused to known operons. In one case, the *y* and *a* genes have come under the control of an operator

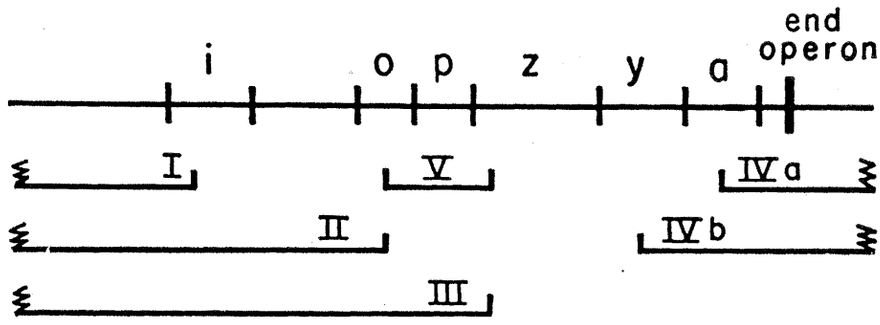

Fig. 4. Deletions of the *lac* operon.

for two genes involved in purine biosynthesis (*43*), and, in the other, the *lac* genes have come under the control of the operon determining tryptophan biosynthesis (Fig. 5) (*12*). In these cases, derepression of the tryptophan (*try*) or purine (*pur*) operons leads to a corresponding derepression of the remaining *lac* genes. In addition, the presence of purines or of tryptophan in the growth medium of the two types of strains results in strong repression of the synthesis of permease and transacetylase.

These results show that when a deletion removes the *lac* operator and a proximal segment of the *z* gene and fuses the *y* and *a* genes to an operon in the repressed state, the *y* and *a* genes are not expressed. Thus, it is not possible to restore *y* and *a* function to an inactivated *lac* operon merely by deleting the proximal end of the operon without regard to where the other end of the deletion lies. On the contrary, it seems likely that, to allow expression of the *y* and *a* genes, the deletions must connect these genes to another functioning gene or operon. Further evidence that this is the case comes from the failure to find, among a large number of O^c mutations, deletions covering part of the *z* gene which do not extend past the *i* gene. If it were possible to connect the *y* and *a* genes to any region of the DNA and allow function, one would have expected to find among the O^c–z^- deletions some in which the deletion either ended in the operator or in whatever space might exist between *i* and *o* (Fig. 4, class V). These results suggest that the *lac* operon has an associated site which is essential for expression and that removal of this site, without the substitution of a new one, inactivates the operon. Apparently, deletions of class V inactivate the operon, since they are not found. Therefore, this essential site, in the case of the *lac* operon, must lie between *o* and *z*. On the basis of this argument, Jacob, Ullman, and Monod (*38*) have proposed this location for the promoter site.

Second Function for the Promoter

The studies on deletions of the *lac* operon also provide information on the site or sites which determine the rate of expression of the operon under conditions of full induction. All deletions of class III (Fig. 4) result in a marked reduction in the amount of expression of the remaining *y* and *a* genes. In most cases, the activities of permease and transacetylase in these strains are only 10 percent or less of the activities seen in fully induced wild-type strains (*38, 41, 43*). Several explanations can account for these findings.

1) The *lac* operon may ordinarily be transcribed at a very fast rate compared to that of most other genes and operons, either because of repression of the other genes acting on transcription, or because of lower intrinsic potentials for transcription of the other genes. According to this proposal, the *y* and *a* genes are always likely to function at a lower rate when connected to a different gene or operon.

2) The *lac* operon mRNA is translated at a very fast rate compared to that of other genes or operons. This argument concerning the rate of translation is analogous to the first argument.

3) The deletions fusing *z* to some other gene create a sequence of nucleotides which results in polarity effects, thus reducing the expression of *y* and *a*. For example, the deletion may create a frame-shift in the operon or a nonsense mutation at the site of fusion. This explanation could account for low *y* and *a* in some cases, but it seems very unlikely that every deletion creates just the correct conditions for such an effect.

Another set of deletions was isolated which removed only part of the operator but again deleted the *i* gene (class II) (*38*), thus connecting the *lac* operon, through its operator, to some nearby segment of the chromosome. Deletions of class II are far more frequent than those of class III. None of these class II deletions completely abolishes operator function, as measured by its sensitivity to repressor. In the presence of the wild-type allele of the *i* gene, these O^c mutants are still partly repressible. But, the striking property of these deletions, in contrast to those of class III, is that they all produce maximally exactly the same amounts (100 percent) of *lac* gene products as the wild-type strain. Thus, since extensive deletion of the operator does not reduce the maximal rate of operon expression, the operator (*44*) cannot be the site which sets the maximal potential for expression of the *lac* operon.

How can we explain the difference between deletions of classes II and III? Since we have concluded that a site *p* lying between *o* and *z* is essential for expression of the operon, we may now suggest that this site (or a site in the same region) is also involved in determining the maximum rate of this expression (*45*). Deletions of class III remove this site and connect the *lac* operon to some other gene or operon with a promoter which functions at a lower rate. Deletions of class II, in contrast, do not delete the promoter, and so the site determining

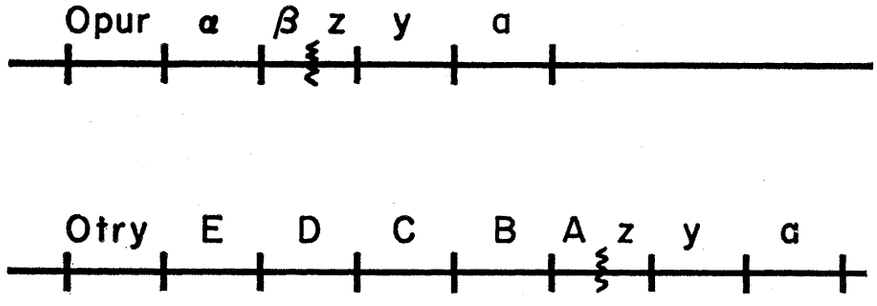

Fig. 5. Fused *lac* operons. The *site* of fusion of the *lac* and *try* operons was determined by E. R. Signer (*53*).

the maximum rate of expression of the operon is left fully functional. In addition, there are no sites at the distal end of the operon determining the maximum level of operon expression, since class IVa and IVb deletions have no effect on the expression of the intact proximal genes (*12*).

Regulation of Operon Transcription

The information obtained from study of these deletions is critical in analyzing the various mechanisms proposed for regulation and expression of the operon. These arguments can be reduced to the question of whether or not the operator (or some site previous to it) is the starting point of mRNA synthesis. I now consider these two possibilities for operon regulation and their implications.

1) The operator, in addition to carrying the information determining the sensitivity to repressor, is also the site of initiation of mRNA synthesis for the operon. Since we have concluded that the operator does not set the maximum level of operon expression, it follows from this model that the rate of initiation of mRNA synthesis is not the rate-limiting factor in the setting of this maximum. In other words, all genes or segments of the DNA to which the operon has become attached by deletions of class III are transcribed at the same rate as the *lac* operon is normally (when fully induced). The variation in maximal levels of gene expression is always a result of different rates of translation. An implication of this model is that *E. coli* has not evolved a mechanism for setting different maximal rates of RNA synthesis for different genes and operons. Although there is no strong evidence against this possibility, it is rather unattractive.

2) The only function of the operator is that of being the repressor-sensitive site. The promoter site is the initiation point for mRNA synthesis, and its structure determines how much transcription takes place in the absence of repression. According to this model, in contrast to the first possibility, different rates of mRNA synthesis can be set by different sequences of nucleotides (promoters) on the DNA.

In view of the implications of model 1, I favor this latter picture of transcription of the *lac* operon. Since the promoter would be the site of initiation of mRNA synthesis, according to this model, the operator cannot be transcribed as part of the operon mRNA. Therefore, the operon mRNA does not contain a repressor-recognition site. Then, in order to affect the expression of the *lac* structural genes, the repressor must recognize the operator on the chromosome. If repression does take place on the chromosome, the repressor could either directly repress the initiation of synthesis of operon mRNA or interact with the mRNA-DNA complex (or even a ribosome-RNA-DNA complex) to prevent further synthesis of operon mRNA.

It is very likely that a promoter region plays some part in determining the activity of the *lac* operon. It should be possible, therefore, to isolate mutants of the promoter which alter this maximal level. Mutants have been found which reduce the maximum expression of the *lac* operon and which have other properties expected of promoter mutants (*45*). These are mutants which, when fully induced, make only 5 to 10 percent of the normal fully induced levels of all three gene products. The mutants are not connected with the *i-o* control system, do not appear to be polar mutants, and map in the same region as the promoter.

The Repressor

The existence of two classes of mutations affecting *lac* regulation and defining the *i* locus, led to the concept of the *i* gene product as a repressor molecule which interacts both with the operator and the inducer (*1*). Strains carrying i^- mutations which abolish activity of the *i* gene result in a maximal synthesis of the *lac* proteins, even in the absence of inducer. Moreover, in partial diploid strains carrying both an i^- mutation and the wild-type i^+ allele of the *i* gene, the operon becomes repressible (*1, 46*). Thus, the *i* gene appears to be the structural gene for a diffusible product which is responsible for the repression of the operon. The i^s (super-repressed) mutations, in contrast, result in noninducibility of the *lac* operon (*24, 47*), in spite of the presence of unaltered operator, promoter, and structural genes. These mutants are thought to produce a repressor which has no affinity for inducer, so that the operon is permanently shut off. However, as pointed out by Brenner (*5*), on the basis of only these two classes of mutations one can devise somewhat more complex pictures of *i* gene action; for example, the *i* gene product is not itself the repressor but is an enzyme catalyzing the synthesis of a small molecule which is part of the repressor. However, if there were another molecule involved in repression, one would have expected to find mutations in another locus resulting in constitutivity. No such mutants have been found. Thus, the original Jacob-Monod idea of *i* gene action is still the simplest.

There is now both genetic and chemical evidence that the *i* gene ultimately directs the synthesis of a protein molecule. The genetic evidence is that there exist suppressible amber mutants (i^-) of the *i* gene (*39, 48*). Since it is known that these mutants affect translation, the *i* gene messenger must be translated into protein. Recently, Gilbert and Muller-Hill (*49*) have isolated the *i* gene product and have shown that it is at least partially composed of protein. The isolation was achieved by purification of a fraction of *E. coli* protein which binds the inducer IPTG (isopropyl-β-D-thiogalactoside) with the expected affinity constant. The identification of this protein as the *i* gene product was established first by the demonstration that this protein could not be detected by this technique in i^-, i^s and *i*-deletion strains; more importantly, the affinity constant of this protein for IPTG was altered in a strain carrying an *i* mutation leading to an increased efficiency of IPTG as an inducer.

Old Theories and Future Experiments

I have discussed the genetic work on the *lac* operon and its implications in terms of the Jacob-Monod operon hypothesis. As mentioned earlier, several alternative models of control have been proposed. In addition to the Jacob-Monod model, in which the operator is not transcribed into the operon mRNA, the following possibilities have been suggested. The first possibility is a model in which the operator is transcribed into the operon mRNA, but is not translated. For example, (i) the repressor binds to the mRNA copy of the operator, inhibiting the initiation of protein synthesis (at the promoter?), which, in turn, is necessary for continuing synthesis and release of the mRNA from the DNA. In this model, repression during translation inhibits the transcription process (*2, 3*). (ii) The operon mRNA is synthesized

in constant amounts, but the repressor inhibits initiation of protein synthesis and thus causes a very rapid destruction of the mRNA (*15, 31*). (iii) The operon mRNA is made in constant amounts, but the repressor is a ribonuclease which destroys the mRNA (*5, 31*).

The second possibility includes models in which the operator is transcribed into the operon mRNA and then must be translated in order to allow translation of the structural genes of the operon (*2–4*). Again, the failure to translate may result either in the mRNA's sticking to the DNA or in rapid destruction of the mRNA.

All of these models take into account the experiments of Attardi *et al.* (*28*) which show that repression results in a disappearance of *lac* mRNA from the cell. This finding, like the similar finding with the strong polar mutants, must be considered in formulating a model for *lac* operon regulation, but it does not severely limit the number of possibilities.

The thesis of this article has been that the current knowledge of the *lac* operon suggests that the operator is not transcribed into the operon mRNA and that, therefore, the Jacob-Monod suggestion for operon control is most likely to be correct for this system. Some of the other models listed are difficult to reconcile with some of the evidence discussed. One of the most striking of the recent findings is that most and probably all O^c's are deletions but still retain some sensitivity to repressor. Although this finding makes any picture of the repressor-operator interaction somewhat difficult to visualize, it makes particularly unattractive models in which the operator is translated into protein or is the initiation site for operon protein synthesis. One would expect deletions of the operator to have drastic effects on operon functioning in such models, and this does not appear to be so.

Final proof of one model or another will probably have to come from biochemical experiments on operon functioning. It is possible to set up a system in vitro in which *lac* mRNA is made, with RNA polymerase and DNA preparations that contain a high proportion of *lac* genes (*50*). In such experiments, other mRNA species are made also, but these can be eliminated by annealing with DNA preparations from appropriate strains. Then, the amount of *lac* mRNA made can be estimated. Using this system, Gilbert and Muller-Hill (*51*) are attempting to ascertain whether the *lac* repressor protein will inhibit synthesis of *lac* mRNA. In the same system, we are attempting to see whether the potential promoter mutants affect *lac* mRNA synthesis in vitro.

Other Regulatory Systems

Of the regulatory systems that have been studied in detail, the *lac* operon appears to be one of the simplest. Although the operon model can account for all the information concerning *lac*, in some other systems the control is clearly more complex, and even the basic control mechanism may be entirely different (*52*). However, in none of these cases is there yet any strong evidence against the operon model.

References and Notes

1. F. Jacob and J. Monod, *J. Mol. Biol.* **3**, 318 (1961).
2. B. N. Ames and P. E. Hartman, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 349 (1963).
3. G. Stent, *Science* **144**, 816 (1964).
4. W. Maas and E. McFall, *Ann. Rev. Microbiol.* **18**, 95 (1964).
5. S. Brenner, *Brit. Med. Bull.* **21**, 244 (1965).
6. Both Jacob [*Science* **152**, 1470 (1966)] and Monod [*ibid.* **154**, 475 (1966)] have in their Nobel laureate lectures discussed many historical and recent aspects of the *lac* operon and its regulation.
7. C. F. Fox and E. P. Kennedy, *Proc. Nat. Acad. Sci. U.S.* **54**, 891 (1965).
8. H. V. Rickenberg, G. N. Cohen, G. Buttin, J. Monod, *Ann. Inst. Pasteur* **91**, 829 (1956).
9. I. Zabin, A. Kepes, J. Monod, *J. Biol. Chem.* **237**, 253 (1962).
10. C. F. Fox, J. R. Beckwith, W. Epstein, E. R. Signer, *J. Mol. Biol.* **19**, 576 (1966).
11. F. Jacob and J. Monod, *Biochem. Biophys. Res. Commun.* **18**, 693 (1965).
12. J. R. Beckwith, E. R. Signer, W. Epstein, *Cold Spring Harbor Symp. Quant. Biol.* **31**, 393 (1967).
13. W. Gilbert, *J. Mol. Biol.* **6**, 374 (1963); A. Gierer, *ibid.*, p. 148; J. R. Warner, P. Knopf, A. Rich, *Proc. Nat. Acad. Sci. U.S.* **49**, 122 (1963); F. O. Wettstein, T. Staehelin, H. Noll, *Nature* **197**, 430 (1963); H. M. Dintzis, *Proc. Nat. Acad. Sci. U.S.* **48**, 247 (1961).
14. S. Benzer and S. P. Champe, *Proc. Nat. Acad. Sci. U.S.* **48**, 1114 (1962); A. Garen and O. Siddiqi, *ibid.*, p. 1121; A. S. Sarabhai, A. O. W. Stretton, S. Brenner, A. Bolle, *Nature* **201**, 13 (1964); S. Brenner, A. O. W. Stretton, S. Kaplan, *ibid.* **206**, 994 (1965); M. G. Weigert and A. Garen, *ibid.*, p. 992.
15. S. Brenner and J. R. Beckwith, *J. Mol. Biol.* **13**, 629 (1965).
16. F. H. C. Crick, L. Barnett, S. Brenner, R. J. Watts-Tobin, *Nature* **192**, 1227 (1961).
17. Y. Kiho and A. Rich, *Proc. Nat. Acad. Sci. U.S.* **54**, 1751 (1965).
18. R. E. Thach, M. A. Cecere, T. A. Sundararajan, P. Doty, *ibid.*, p. 1167; M. Salas, M. A. Smith, W. M. Stanley, A. J. Wahba, S. Ochoa, *J. Biol. Chem.* **240**, 3988 (1965); G. Streisinger, Y. Okada, E. Terzaghi, J. Emrich, A. Tsugita, M. Inouye, *Cold Spring Harbor Symp. Quant. Biol.*, in press; A. Goldstein, J. Kirschbaum, A. Roman, *Proc. Nat. Acad. Sci. U.S.* **54**, 1669 (1965); V. Maitra and J. Hurwitz, *ibid.*, p. 815; H. Bremer, M. W. Konrad, K. Gaines, G. S. Stent, *J. Mol. Biol.* **13**, 540 (1965); J. R. Guest and C. Yanofsky, *Nature* **210**, 799 (1966).
19. A. V. Fowler and I. Zabin, *Science* **154**, 1027 (1966).
20. Studies on complementation with z^- mutants suggest the possibility that the *z* gene is composed of three or more genes specifying distinct polypeptide chains [A. Ullman, D. Perrin, F. Jacob, J. Monod, *J. Mol. Biol.* **12**, 918 (1965)]. Although this question has not been conclusively resolved, most of the evidence indicates that the *z* gene directs the synthesis of only *one* polypeptide chain [J. L. Brown, S. Koorajian, J. Katze, I. Zabin, *J. Biol. Chem.* **241**, 2826 (1966); G. Craven, personal communication; D. Fan, personal communication].
21. In this paper, the proximal end of a gene or of the operon refers to that end which is closest to the operator. A distal end is that farthest from the operator.
22. C. Yanofsky, B. C. Carlton, J. R. Guest, D. R. Helinski, U. Henning, *Proc. Nat. Acad. Sci. U.S.* **51**, 266 (1964).
23. W. A. Newton, J. R. Beckwith, D. Zipser, S. Brenner, *J. Mol. Biol.* **14**, 290 (1965).
24. F. Jacob and J. Monod, *Cold Spring Harbor Symp. Quant. Biol.* **26**, 193 (1961).
25. N. C. Franklin and S. E. Luria, *Virology* **15**, 299 (1961).
26. W. A. Newton, *Cold Spring Harbor Symp. Quant. Biol.* **31**, 181 (1967).
27. D. Zipser and W. A. Newton, *J. Mol. Biol.*, in press.
28. G. S. Attardi, S. Naono, J. Rouviere, F. Jacob, F. Gros, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 363 (1963); G. Contesse, S. Naono, F. Gros, *Compt. Rend.* **263**, 1007 (1966).
29. M. Malamy, *Cold Spring Harbor Symp. Quant. Biol.* **31**, 89 (1967).
30. R. G. Martin, D. F. Silbert, D. W. E. Smith, H. J. Whitfield, *J. Mol. Biol.* **21**, 357 (1966).
31. J. R. Beckwith, *Abhandl. Deutsch. Akad. Wiss. Berlin Kl. Med.* **4**, 119 (1964).
32. M. R. Capecchi, *J. Mol. Biol.* **21**, 173 (1966); G. N. Gussin, *ibid.*, p. 435.
33. C. Yanofsky and J. Ito, *ibid.*, p. 313.
34. D. H. Alpers and G. M. Tomkins, *Proc. Nat. Acad. Sci. U.S.* **53**, 797 (1965).
35. A. Kepes, in preparation.
36. D. H. Alpers and G. M. Tomkins, *J. Biol. Chem.* **241**, 4434 (1966).
37. I. Zabin, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 431 (1963); W. Epstein, personal communication; D. Perrin, personal communication.
38. F. Jacob, A. Ullman, J. Monod, *Compt. Rend.* **258**, 3125 (1964).
39. S. Bourgeois, M. Cohn, L. E. Orgel, *J. Mol. Biol.* **14**, 300 (1965); S. Bourgeois, personal communication.
40. T. Ramakrishnan and E. A. Adelberg, *J. Bacteriol.* **87**, 566 (1964).
41. J. R. Beckwith, *J. Mol. Biol.* **8**, 427 (1964).
42. E. Steers, G. R. Craven, C. B. Anfinsen, *Proc. Nat. Acad. Sci. U.S.* **54**, 1174 (1965).
43. F. Jacob, A. Ullman, J. Monod, *J. Mol. Biol.* **13**, 704 (1965).
44. One might consider that since no 100-percent O^c's have been found, a right-hand extremity of the operator is involved in determining the rate of operon expression. This argument is not very different from the one discussed in that rather than one site with two distinguishable regions there are two sites.
45. J. Scaife and J. R. Beckwith, *Cold Spring Harbor Symp. Quant. Biol.* **31**, 403 (1967).
46. A. B. Pardee, F. Jacob, J. Monod, *J. Mol. Biol.* **1**, 165 (1959).
47. C. Willson, D. Perrin, M. Cohn, F. Jacob, J. Monod, *ibid.* **8**, 582 (1964).
48. B. Muller-Hill, *ibid.* **15**, 374 (1966).
49. W. Gilbert and B. Muller-Hill, *Proc. Nat. Acad. Sci. U.S.* **56**, 1891 (1966).
50. High-frequency transducing lysates for the *lac* genes have been isolated with the restricted transducing phage $\phi 80$ [J. R. Beckwith and E. R. Signer, *J. Mol. Biol.* **19**, 254 (1966)]. Isolation of DNA from phage lysates results in a DNA preparation which contains genetic material, the *lac* region, and a small portion of the chromosome surrounding *lac*.
51. W. Gilbert and B. Muller-Hill, personal communication.
52. E. Engelsberg, J. Irr, J. Power, N. Lee, *J. Bacteriol.* **90**, 946 (1965) (arabinose metabolism); H. Echols, A. Garen, S. Garen, A. Torriani, *J. Mol. Biol.* **3**, 425 (1961); A. Garen and N. Otsuji, *ibid.* **8**, 841 (1964) (alkaline phosphatase); M. Schwartz, personal communication (maltose metabolism).
53. E. R. Signer, unpublished data.

54. I thank Dr. B. Davis, Dr. W. Epstein, Dr. J. Scaife and Dr. E. Signer for discussions and criticisms of the manuscript. Supported by PHS research career program award GM-9027, PHS research grant GM-13017, and by NSF grant GB-5763.