*J. Mol. Biol.* (1961) **3**, 318–356

REVIEW ARTICLE

# Genetic Regulatory Mechanisms in the Synthesis of Proteins †

FRANÇOIS JACOB AND JACQUES MONOD

*Services de Génétique Microbienne et de Biochimie Cellulaire, Institut Pasteur, Paris*

(*Received 28 December 1960*)

The synthesis of enzymes in bacteria follows a double genetic control. The so-called structural genes determine the molecular organization of the proteins. Other, functionally specialized, genetic determinants, called regulator and operator genes, control the rate of protein synthesis through the intermediacy of cytoplasmic components or repressors. The repressors can be either inactivated (induction) or activated (repression) by certain specific metabolites. This system of regulation appears to operate directly at the level of the synthesis by the gene of a short-lived intermediate, or messenger, which becomes associated with the ribosomes where protein synthesis takes place.

## 1. Introduction

According to its most widely accepted modern connotation, the word "gene" designates a DNA molecule whose specific self-replicating structure can, through mechanisms unknown, become translated into the specific structure of a polypeptide chain.

This concept of the "structural gene" accounts for the multiplicity, specificity and genetic stability of protein structures, and it implies that such structures are not controlled by environmental conditions or agents. It has been known for a long time, however, that the synthesis of individual proteins may be provoked or suppressed within a cell, under the influence of specific external agents, and more generally that the relative rates at which different proteins are synthesized may be profoundly altered, depending on external conditions. Moreover, it is evident from the study of many such effects that their operation is absolutely essential to the survival of the cell.

It has been suggested in the past that these effects might result from, and testify to, complementary contributions of genes on the one hand, and some chemical factors on the other in determining the final structure of proteins. This view, which contradicts at least partially the " structural gene" hypothesis, has found as yet no experimental support, and in the present paper we shall have occasion to consider briefly some of this negative evidence. Taking, at least provisionally, the structural gene hypothesis in its strictest form, let us assume that the DNA message contained within a gene is both necessary and sufficient to define the structure of a protein. The elective effects of agents other than the structural gene itself in promoting or suppressing the synthesis of a protein must then be described as operations which control the rate of transfer of structural information from gene to protein. Since it seems to be established

† This work has been aided by grants from the National Science Foundation, the Jane Coffin Childs Memorial Fund for Medical Research, and the Commissariat à l'Energie Atomique.

that proteins are synthesized in the cytoplasm, rather than directly at the genetic level, this transfer of structural information must involve a chemical intermediate synthesized by the genes. This hypothetical intermediate we shall call the structural messenger. The rate of information transfer, i.e. of protein synthesis, may then depend either upon the activity of the gene in synthesizing the messenger, or upon the activity of the messenger in synthesizing the protein. This simple picture helps to state the two problems with which we shall be concerned in the present paper. If a given agent specifically alters, positively or negatively, the rate of synthesis of a protein, we must ask:

(a) Whether the agent acts at the cytoplasmic level, by controlling the activity of the messenger, or at the genetic level, by controlling the synthesis of the messenger.

(b) Whether the specificity of the effect depends upon some feature of the information transferred from structural gene to protein, or upon some specialized controlling element, not represented in the structure of the protein, gene or messenger.

The first question is easy to state, if difficult to answer. The second may not appear so straightforward. It may be stated in a more general way, by asking whether the genome is composed exclusively of structural genes, or whether it also involves determinants which may control the rates of synthesis of proteins according to a given set of conditions, without determining the structure of any individual protein. Again it may not be evident that these two statements are equivalent. We hope to make their meaning clear and to show that they are indeed equivalent, when we consider experimental examples.

The best defined systems wherein the synthesis of a protein is seen to be controlled by specific agents are examples of enzymatic adaptation, this term being taken here to cover both enzyme induction, i.e. the formation of enzyme electively provoked by a substrate, and enzyme repression, i.e. the specific inhibition of enzyme formation brought about by a metabolite. Only a few inducible and repressible systems have been identified both biochemically and genetically to an extent which allows discussion of the questions in which we are interested here. In attempting to generalize, we will have to extrapolate from these few systems. Such generalization is greatly encouraged, however, by the fact that lysogenic systems, where phage protein synthesis might be presumed to obey entirely different rules, turn out to be analysable in closely similar terms. We shall therefore consider in succession certain inducible and repressible enzyme systems and lysogenic systems.

It might be best to state at the outset some of the main conclusions which we shall arrive at. These are:

(a) That the mechanisms of control in all these systems are negative, in the sense that they operate by inhibition rather than activation of protein synthesis.

(b) That in addition to the classical structural genes, these systems involve two other types of genetic determinants (regulator and operator) fulfilling specific functions in the control mechanisms.

(c) That the control mechanisms operate at the genetic level, i.e. by regulating the activity of structural genes.

## 2. Inducible and Repressible Enzyme Systems

### (a) *The phenomenon of enzyme induction. General remarks*

It has been known for over 60 years (Duclaux, 1899; Dienert, 1900; Went, 1901) that certain enzymes of micro-organisms are formed only in the presence of their

specific substrate. This effect, later named "enzymatic adaptation" by Karstrom (1938), has been the subject of a great deal of experimentation and speculation. For a long time, "enzymatic adaptation" was not clearly distinguished from the selection of spontaneous variants in growing populations, or it was suggested that enzymatic adaptation and selection represented *alternative* mechanisms for the acquisition of a "new" enzymatic property. Not until 1946 were adaptive enzyme systems shown to be controlled in bacteria by discrete, specific, stable, i.e. genetic, determinants (Monod & Audureau, 1946). A large number of inducible systems has been discovered and studied in bacteria. In fact, enzymes which attack exogeneous substrates are, as a general rule, inducible in these organisms. The phenomenon is far more difficult to study in tissues or cells of higher organisms, but its existence has been established quite clearly in many instances. Very often, if not again as a rule, the presence of a substrate induces the formation not of a single but of several enzymes, sequentially involved in its metabolism (Stanier, 1951).

Most of the fundamental characteristics of the induction effect have been established in the study of the "lactose" system of *Escherichia coli* (Monod & Cohn, 1952; Cohn, 1957; Monod, 1959) and may be summarized in a brief discussion of this system from the biochemical and physiological point of view. We shall return later to the genetic analysis of this system.

### (b) *The lactose system of* Escherichia coli

Lactose and other β-galactosides are metabolized in *E. coli* (and certain other enteric bacteria) by the hydrolytic transglucosylase β-galactosidase. This enzyme was isolated from *E. coli* and later crystallized. Its specificity, activation by ions and transglucosylase *vs* hydrolase activity have been studied in great detail (*cf.* Cohn, 1957). We need only mention the properties that are significant for the present discussion. The enzyme is active exclusively on β-galactosides unsubstituted on the galactose ring. Activity and affinity are influenced by the nature of the aglycone moiety both being maximum when this radical is a relatively large, hydrophobic group. Substitution of sulfur for oxygen in the galactosidic linkage of the substrate abolishes hydrolytic activity completely, but the thiogalactosides retain about the same affinity for the enzyme site as the homologous oxygen compounds.

As isolated by present methods, β-galactosidase appears to form various polymers (mostly hexamers) of a fundamental unit with a molecular weight of 135,000. There is one end group (threonine) and also one enzyme site (as determined by equilibrium dialysis against thiogalactosides) per unit. It is uncertain whether the monomer is active as such, or exists *in vivo*. The hexameric molecule has a turnover number of 240,000 $\text{mol} \times \text{min}^{-1}$ at 28°C, pH 7·0 with *o*-nitrophenyl-β-D-galactoside as substrate and $Na^+$ (0·01 M) as activator.

There seems to exist only a single homogeneous β-galactosidase in *E. coli*, and this organism apparently cannot form any other enzyme capable of metabolizing lactose, as indicated by the fact that mutants that have lost β-galactosidase activity cannot grow on lactose as sole carbon source.

However, the possession of β-galactosidase activity is not sufficient to allow utilization of lactose by *intact E. coli* cells. Another component, distinct from β-galactosidase, is required to allow penetration of the substrate into the cell (Monod, 1956; Rickenberg, Cohen, Buttin & Monod, 1956; Cohen & Monod, 1957; Pardee, 1957; Képès, 1960). The presence and activity of this component is determined by measuring the rate of

entry and/or the level of accumulation of radioactive thiogalactosides into intact cells. Analysis of this active permeation process shows that it obeys classical enzyme kinetics allowing determination of $K_m$ and $V_{max}$. The specificity is high since the system is active only with galactosides ($\beta$ or $\alpha$), or thiogalactosides. The spectrum of apparent affinities ($1/K_m$) is very different from that of $\beta$-galactosidase. Since the permeation system, like $\beta$-galactosidase, is inducible (see below) its formation can be studied *in vivo*, and shown to be invariably associated with protein synthesis. By these criteria, there appears to be little doubt that this specific permeation system involves a specific protein (or proteins), formed upon induction, which has been called galactoside-permease. That this protein is distinct from and independent of $\beta$-galactosidase is shown by the fact that mutants that have lost $\beta$-galactosidase retain the capacity to concentrate galactosides, while mutants that have lost this capacity retain the power to synthesize galactosidase. The latter mutants (called cryptic) cannot however use lactose, since the intracellular galactosidase is apparently accessible exclusively *via* the specific permeation system.

Until quite recently, it had not proved possible to identify *in vitro* the inducible protein (or proteins) presumably responsible for galactoside-permease activity. During the past year, a protein characterized by the ability to carry out the reaction:

Ac. Coenzyme A + Thiogalactoside → 6-Acetylthiogalactoside + Coenzyme A

has been identified, and extensively purified from extracts of *E. coli* grown in presence of galactosides (Zabin, Képès & Monod, 1959). The function of this enzyme in the system is far from clear, since formation of a free covalent acetyl-compound is almost certainly not involved in the permeation process *in vivo*. On the other hand:

(a) mutants that have lost $\beta$-galactosidase and retained galactoside-permease, retain galactoside-acetylase;

(b) most mutants that have lost permease cannot form acetylase;

(c) permeaseless acetylaseless mutants which revert to the permease-positive condition simultaneously regain the ability to form acetylase.

These correlations strongly suggest that galactoside-acetylase is somehow involved in the permeation process, although its function *in vivo* is obscure, and it seems almost certain that other proteins (specific or not for this system) are involved. In any case, we are interested here not in the mechanisms of permeation, but in the control mechanisms which operate with $\beta$-galactosidase, galactoside-permease and galactoside-acetylase. The important point therefore is that, as we shall see, galactoside-acetylase invariably obeys the same controls as galactosidase.†

### (c) *Enzyme induction and protein synthesis*

Wild type *E. coli* cells grown in the absence of a galactoside contain about 1 to 10 units of galactosidase per mg dry weight, that is, an average of 0·5 to 5 active molecules

† For reasons which will become apparent later it is important to consider whether there is any justification for the assumption that galactosidase and acetylase activities might be associated with the same fundamental protein unit. We should therefore point to the following observation:

(a) There are mutants which form gacatosidase and no acetylase, and *vice versa*.

(b) Purified acetylase is devoid of any detectable galactosidase activity.

(c) The specificity of the two enzymes is very different.

(d) The two enzymes are easily and completely separated by fractional precipitation.

(e) Acetylase is highly heat-resistant, under conditions where galactosidase is very labile.

(f) Anti-galactosidase serum does not precipitate acetylase; nor does anti-acetylase serum precipitate galactosidase.

There is therefore no ground for the contention that galactosidase and acetylase activities are associated with the same protein.

per cell or 0·15 to 1·5 molecules per nucleus. Bacteria grown in the presence of a suitable inducer contain an average of 10,000 units per mg dry weight. This is the induction effect.

A primary problem, to which much experimental work has been devoted, is whether this considerable increase in specific activity corresponds to the synthesis of entirely "new" enzyme molecules, or to the activation or conversion of pre-existing protein precursors. It has been established by a combination of immunological and isotopic methods that the enzyme formed upon induction:

(a) is distinct, as an antigen, from all the proteins present in uninduced cells (Cohn & Torriani, 1952);

(b) does not derive any significant fraction of its sulfur (Monod & Cohn, 1953; Hogness, Cohn & Monod, 1955) or carbon (Rotman & Spiegelman, 1954) from pre-existing proteins.

The inducer, therefore, brings about the complete *de novo* synthesis of enzyme molecules which are new by their specific structure as well as by the origin of their elements. The study of several other induced systems has fully confirmed this conclusion, which may by now be considered as part of the *definition* of the effect. We will use the term "induction" here as meaning "activation by inducer of enzyme-protein synthesis."

### (d) *Kinetics of induction*

Accepting (still provisionally) the structural gene hypothesis, we may therefore consider that the inducer somehow accelerates the rate of information transfer from gene to protein. This it could do either by provoking the synthesis of the messenger or by activating the messenger. If the messenger were a *stable* structure, functioning as a catalytic template in protein synthesis, one would expect different kinetics of induction, depending on whether the inducer acted at the genetic or at the cytoplasmic level.

The kinetics of galactosidase induction turn out to be remarkably simple when determined under proper experimental conditions (Monod, Pappenheimer & Cohen-Bazire, 1952; Herzenberg, 1959). Upon addition of a suitable inducer to a growing culture, enzyme activity increases at a rate proportional to the increase in total protein within the culture; i.e. a linear relation is obtained (Fig. 1) when total enzyme activity is plotted against mass of the culture. The slope of this line:

$$P = \frac{\Delta z}{\Delta M}$$

is the "differential rate of synthesis," which is taken by definition as the measure of the effect. Extrapolation to the origin indicates that enzyme formation begins about three minutes (at 37 °C) after addition of inducer (Pardee & Prestidge, 1961). Removal of the inducer (or addition of a specific anti-inducer, see below) results in cessation of enzyme synthesis within the same short time. The differential rate of synthesis varies with the concentration of inducer reaching a different saturation value for different inducers. The inducer therefore acts in a manner which is (kinetically) similar to that of a dissociable activator in an enzyme system: activation and inactivation follow very rapidly upon addition or removal of the activator.

The conclusion which can be drawn from these kinetics is a negative one: the inducer does not appear to activate the synthesis of a stable intermediate able to accumulate in the cell (Monod, 1956).

Similar kinetics of induction have been observed with most or all other systems which have been adequately studied (Halvorson, 1960) with the exception of penicillinase of *Bacillus cereus*. The well-known work of Pollock has shown that the synthesis of this enzyme continues for a long time, at a decreasing rate, after removal of inducer (penicillin) from the medium. This effect is apparently related to the fact that minute amounts of penicillin are retained irreversibly by the cells after transient exposure to the drug (Pollock, 1950). The unique behavior of this system therefore does not contradict the rule that induced synthesis stops when the inducer is removed from the cells. Using this system, Pollock & Perret (1951) were able to show that the inducer acts catalytically, in the sense that a cell may synthesize many more enzyme molecules than it has retained inducer molecules.

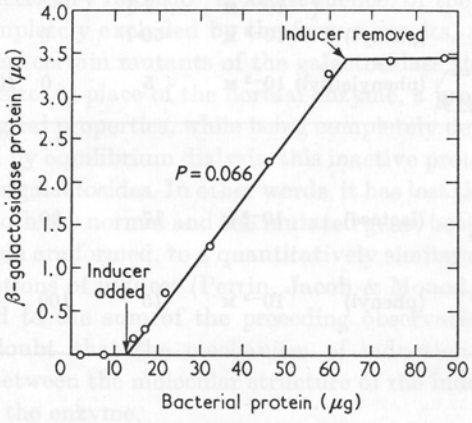



FIG. 1. Kinetics of induced enzyme synthesis. Differential plot expressing accumulation of β-galactosidase as a function of increase of mass of cells in a growing culture of *E. coli*. Since abscissa and ordinates are expressed in the same units (micrograms of protein) the slope of the straight line gives galactosidase as the fraction ($P$) of total protein synthesized in the presence of inducer. (After Cohn, 1957.)

### (e) *Specificity of induction*

One of the most conspicuous features of the induction effect is its extreme specificity. As a general rule, only the substrate of an enzyme, or substances very closely allied to the normal substrate, are endowed with inducer activity towards this enzyme. This evidently suggests that a correlation between the molecular structure of the inducer and the structure of the catalytic center on the enzymes is *inherently* involved in the mechanism of induction. Two main types of hypotheses have been proposed to account for this correlation, and thereby for the mechanism of action of the inducer:

(a) The inducer serves as "partial template" in enzyme synthesis, molding as it were the catalytic center.

(b) The inducer acts by combining specifically with preformed enzyme (or "pre-enzyme"), thereby somehow accelerating the synthesis of further enzyme molecules.

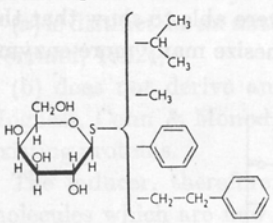
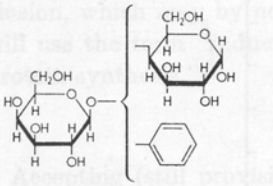
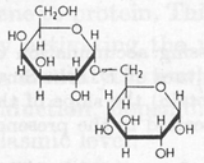
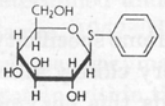
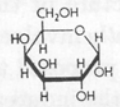
It is not necessary to discuss these "classical" hypotheses in detail, because it seems to be established now that the correlation in question is in fact *not* inherent to the mechanism of induction.

Table 1 lists a number of compounds tested as inducers of galactosidase, and as substrates (or specific inhibitors) of the enzyme. It will be noted that:

(a) no compound that does not possess an intact unsubstituted galactosidic residue induces;

TABLE 1

*Induction of galactosidase and galactoside-transacetylase by various galactosides*

| Compound | Concentrations | β-galactosidase | | | Galactoside-transacetylase | |
|---|---|---|---|---|---|---|
| | | Induction value | $V$ | $1/K_m$ | Induction value | $V/K_m$ |
| **β-D-thiogalactosides** | | | | | | |
|  (*iso*propyl) | $10^{-4}$ M | 100 | 0 | 140 | 100 | 80 |
| (methyl) | $10^{-4}$ M | 78 | 0 | 7 | 74 | 30 |
| | $10^{-5}$ M | 7·5 | — | — | 10 | — |
| (phenyl) | $10^{-3}$ M | <0·1 | 0 | 100 | <1 | 100 |
| (phenylethyl) | $10^{-3}$ M | 5 | 0 | 10,000 | 3 | — |
| **β-D-galactosides** | | | | | | |
|  (lactose) | $10^{-3}$ M | 17 | 30 | 14 | 12 | 35 |
| (phenyl) | $10^{-3}$ M | 15 | 100 | 100 | 11 | — |
| **α-D-galactoside** | | | | | | |
|  (melibiose) | $10^{-3}$ M | 35 | 0 | <0·1 | 37 | <1 |
| **β-D-glucoside** | | | | | | |
|  (phenyl) | $10^{-3}$ M | <0·1 | 0 | 0 | <1 | 50 |
|  (galactose) | $10^{-3}$ M | <0·1 | — | 4 | <1 | <1 |
| Methyl-β-D-thiogalactoside ($10^{-4}$ M) + phenyl-β-D-thiogalactoside ($10^{-3}$ M) | | 52 | — | — | 63 | — |

Columns "induction value" refer to specific activities developed by cultures of wild type *E. coli* K12 grown on glycerol as carbon source with each galactoside added at molar concentration stated. Values are given in percent of values obtained with *iso*propyl-thiogalactoside at $10^{-4}$ M (for which actual units were about 7,500 units of β-galactosidase and 300 units of galactoside-transacetylase per mg of bacteria). Column $V$ refers to maximal substrate activity of each compound with respect to galactosidase. Values are given in percent of activity obtained with phenyl-galactoside. Column $1/K_m$ expresses affinity of each compound with respect to galactosidase. Values are given in percent of that observed with phenylgalactoside. In case of galactoside-transacetylase, only the relative values $V/K_m$ are given since low affinity of this enzyme prevents independent determination of the constants. (Computed from Monod & Cohn, 1952; Monod *et al.*, 1952; Buttin, 1956; Zabin *et al.*, 1959; Képès *et al.*, unpublished results.)

(b) many compounds which are not substrates (such as the thiogalactosides) are excellent inducers (for instance *iso*propyl thiogalactoside);

(c) there is no correlation between affinity for the enzyme and capacity to induce (*cf.* thiophenylgalactoside and melibiose).

The possibility that the enzyme formed in response to different inducers may have somewhat different specific properties should also be considered, and has been rather thoroughly tested, with entirely negative results (Monod & Cohn, 1952).

There is therefore no quantitative correlation whatever between inducing capacity and the substrate activity or affinity parameters of the various galactosides tested. The fact remains, however, that only galactosides will induce galactosidase, whose binding site is complementary for the galactose ring-structure. The possibility that this correlation is a necessary requisite, or consequence, of the induction mechanism was therefore not completely excluded by the former results.

As we shall see later, certain mutants of the galactosidase structural gene (*z*) have been found to synthesize, in place of the normal enzyme, a protein which is identical to it by its immunological properties, while being completely devoid of any enzymatic activity. When tested by equilibrium dialysis, this inactive protein proved to have no measurable affinity for galactosides. In other words, it has lost the specific binding site. In diploids carrying both the normal and the mutated gene, both normal galactosidase and the inactive protein are formed, to a quantitatively similar extent, in the presence of different concentrations of inducer (Perrin, Jacob & Monod, 1960).

This finding, added to the sum of the preceding observations, appears to prove beyond reasonable doubt that the mechanism of induction does not imply any inherent correlation between the molecular structure of the inducer and the structure of the binding site of the enzyme.

On the other hand, there is complete correlation in the induction of galactosidase and acetylase. This is illustrated by Table 1 which shows not only that the same compounds are active or inactive as inducers of either enzyme, but that the relative amounts of galactosidase and acetylase synthesized in the presence of different inducers or at different concentrations of the same inducer are constant, even though the absolute amounts vary greatly. The remarkable qualitative and quantitative correlation in the induction of these two widely different enzyme proteins strongly suggests that the synthesis of both is directly governed by a common controlling element with which the inducer interacts. This interaction must, at some point, involve stereospecific binding of the inducer, since induction is sterically specific, and since certain galactosides which are devoid of any inducing activity act as competitive inhibitors of induction in the presence of active inducers (Monod, 1956; Herzenberg, 1959). This suggests that an enzyme, or some other protein, distinct from either galactosidase or acetylase, acts as "receptor" of the inducer. We shall return later to the difficult problem raised by the identification of this "induction receptor."

### (f) *Enzyme repression*

While positive enzymatic adaptation, i.e. induction, has been known for over sixty years, negative adaptation, i.e. specific inhibition of enzyme synthesis, was discovered only in 1953, when it was found that the formation of the enzyme tryptophan-synthetase was inhibited selectively by tryptophan and certain tryptophan analogs (Monod & Cohen-Bazire, 1953). Soon afterwards, other examples of this effect were observed (Cohn, Cohen & Monod, 1953; Adelberg & Umbarger, 1953; Wijesundera &

Woods, 1953), and several systems were studied in detail in subsequent years (Gorini & Maas, 1957; Vogel, 1957*a*,*b*; Yates & Pardee, 1957; Magasanik, Magasanik & Neidhardt, 1959). These studies have revealed that the "repression" effect, as it was later named by Vogel (1957*a*,*b*), is very closely analogous, albeit symmetrically opposed, to the induction effect.

Enzyme repression, like induction, generally involves not a single but a sequence of enzymes active in successive metabolic steps. While inducibility is the rule for catabolic enzyme sequences responsible for the degradation of exogeneous substances, repressibility is the rule for anabolic enzymes, involved in the synthesis of essential metabolites such as amino acids or nucleotides.† Repression, like induction, is highly specific, but while inducers generally are substrates (or analogs of substrates) of the sequence, the repressing metabolites generally are the product (or analogs of the product) of the sequence.

That the effect involves inhibition of enzyme *synthesis*, and not inhibition (directly or indirectly) of enzyme *activity* was apparent already in the first example studied (Monod & Cohen-Bazire, 1953), and has been proved conclusively by isotope incorporation experiments (Yates & Pardee, 1957). It is important to emphasize this point, because enzyme repression must not be confused with another effect variously called "feedback inhibition" or "retro-inhibition" which is equally frequent, and may occur in the same systems. This last effect, discovered by Novick & Szilard (in Novick, 1955), involves the inhibition of activity of an early enzyme in an anabolic sequence, by the ultimate product of the sequence (Yates & Pardee, 1956; Umbarger, 1956). We shall use "repression" exclusively to designate specific inhibition of enzyme *synthesis*.‡

### (g) *Kinetics and specificity of repression*

The kinetics of enzyme synthesis provoked by "de-repression" are identical to the kinetics of induction (see Fig. 2). When wild type *E. coli* is grown in the presence of arginine, only traces of ornithine-carbamyltransferase are formed. As soon as arginine is removed from the growth medium, the differential rate of enzyme synthesis increases about 1,000 times and remains constant, until arginine is added again, when it immediately falls back to the repressed level. The repressing metabolite here acts (kinetically) as would a dissociable inhibitor in an enzyme system.

The specificity of repression poses some particularly significant problems. As a rule, the repressing metabolite of an anabolic sequence is the ultimate product of this sequence. For instance, L-arginine, to the exclusion of any other amino acid, represses the enzymes of the sequence involved in the biosynthesis of arginine. Arginine shows no specific affinity for the early enzymes in the sequence, such as, in particular, ornithine-carbamyltransferase. In this sense, arginine is a "gratuitous" repressing metabolite for this protein, just as galactosides are "gratuitous inducers" for the mutated (inactive) galactosidase. The possibility must be considered however that arginine may be converted back, through the sequence itself, to an intermediate product

† Certain enzymes which attack exogeneous substrates are controlled by repression. Alkaline phosphatase (*E. coli*) is not induced by phosphate esters, but it is repressed by orthophosphate. Urease (*Pseudomonas*) is repressed by ammonia.

‡ We should perhaps recall the well-known fact that glucose and other carbohydrates inhibit the synthesis of many *inducible* enzymes, attacking a variety of substrates (Dienert, 1900; Gale, 1943; Monod, 1942; Cohn & Horibata, 1959). It is probable that this non-specific "glucose effect" bears some relation to the repressive effect of specific metabolites, but the relationship is not clear (Neidhardt & Magasanik, 1956*a*,*b*). We shall not discuss the glucose effect in this paper.

or substrate of the enzyme. This has been excluded by Gorini & Maas (1957) who showed that, in mutants lacking one of the enzymes involved in later steps of the sequence, ornithine transcarbamylase is repressed by arginine to the same extent as in the wild type. Moreover, neither ornithine nor any other intermediate of the sequence is endowed with repressing activity in mutants which cannot convert the intermediate into arginine. It is quite clear therefore that the specificity of action of the repressing metabolite does not depend upon the specific configuration of the enzyme site.

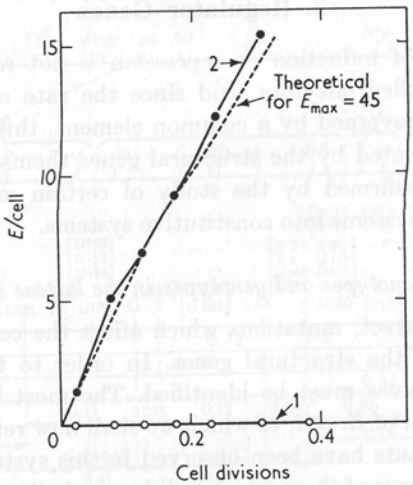



FIG. 2. Repression of ornithine-transcarbamylase by arginine. *E. coli* requiring both histidine and arginine were grown in a chemostat with 1 $\mu$g/ml. histidine + 6 $\mu$g/ml. arginine (curve 1) or with 10 $\mu$g/ml. histidine + 5 $\mu$g/ml. arginine (curve 2). Cultures are inoculated with washed cells taken from cultures growing exponentially in excess of arginine. The theoretical curve was calculated from the constant enzyme/cell value reached after 4 cell divisions. (After Gorini & Maas, 1958.)

The same conclusion is applicable to the enzymes of the histidine synthesizing pathway which are repressed in the presence of histidine, both in the wild type and in different mutants lacking one of the enzymes. The work of Ames & Garry (1959) has shown that the rates of synthesis of different enzymes in this sequence vary in *quantitatively* constant ratios under any set of medium conditions, and that the ratios are the same in various mutants lacking one of the enzymes and in the wild type. Here again, as in the case of the lactose system, the synthesis of widely different, albeit functionally related, enzymes appears to be controlled by a single common mechanism, with which the repressing metabolite specifically interacts.

In summary, repression and induction appear as closely similar effects, even if opposed in their results. Both control the rate of synthesis of enzyme proteins. Both are highly specific, but in neither case is the specificity related to the specificity of action (or binding) of the controlled enzyme. The kinetics of induction and repression are the same. Different functionally related enzymes are frequently co-induced or co-repressed, quantitatively to the same extent, by a single substrate or metabolite.

The remarkable similarity of induction and repression suggests that the two effects represent different manifestations of fundamentally similar mechanisms (Cohn & Monod, 1953; Monod, 1955; Vogel, 1957*a*, *b*; Pardee, Jacob & Monod, 1959; Szilard,

1960). This would imply either that in inducible systems the inducer acts as an antagonist of an internal repressor or that in repressible systems the repressing metabolite acts as an antagonist of an internal inducer. This is not an esoteric dilemma since it poses a very pertinent question, namely what would happen in an adaptive system of either type, when *both* the inducer and the repressor were eliminated? This, in fact, is the main question which we shall try to answer in the next section.

## 3. Regulator Genes

Since the specificity of induction or repression is not related to the structural specificity of the controlled enzymes, and since the rate of synthesis of different enzymes appears to be governed by a common element, this element is presumably not controlled or represented by the structural genes themselves. This inference, as we shall now see, is confirmed by the study of certain mutations which convert inducible or repressible systems into constitutive systems.

### (a) *Phenotypes and genotypes in the lactose systems*

If this inference is correct, mutations which affect the controlling system should not behave as alleles of the structural genes. In order to test this prediction, the structural genes themselves must be identified. The most thoroughly investigated case is the lactose system of *E. coli*, to which we shall now return. Six phenotypically different classes of mutants have been observed in this system. For the time being, we shall consider only three of them which will be symbolized and defined as follows:

(1) Galactosidase mutations: $z^+ \rightleftharpoons z^-$ expressed as the loss of the capacity to synthesize active galactosidase (with or without induction).

(2) Permease mutations: $y^+ \rightleftharpoons y^-$ expressed as the loss of the capacity to form galactoside-permease. Most, but not all, mutants of this class simultaneously lose the capacity to synthesize active acetylase. We shall confine our discussion to the acetylaseless subclass.

(3) Constitutive mutations: $i^+ \rightleftharpoons i^-$ expressed as the ability to synthesize large amounts of galactosidase *and* acetylase in the absence of inducer (Monod, 1956; Rickenberg *et al.*, 1956; Pardee *et al.*, 1959).

The first two classes are specific for either galactosidase or acetylase: the galactosidaseless mutants form normal amounts of acetylase; conversely the acetylaseless mutants form normal amounts of galactosidase. In contrast, the constitutive mutations, of which over one hundred recurrences have been observed, invariably affect both the galactosidase and the permease (acetylase).† There are eight possible combinations of these phenotypes, and they have all been observed both in *E. coli* ML and K12.

The loci corresponding to a number of recurrences of each of the three mutant types have been mapped by recombination in *E. coli* K12. The map (Fig. 3) also

† The significance of this finding could be questioned since, in order to isolate constitutive mutants, one must of course use selective media, and this procedure might be supposed to favour double mutants, where the constitutivity of galactosidase and permease had arisen independently. It is possible, however, to select for $i^+ \rightarrow i^-$ mutants in organisms of type $i^+z^+y^-$, i.e. permeaseless. Fifty such mutants were isolated, giving rise to "constitutive cryptic" types $i^-z^+y^-$ from which, by reversion of $y^-$, fifty clones of constitutive $i^-z^+y^+$ were obtained. It was verified that in each of these fifty clones the permease was constitutive.

indicates the location of certain other mutations (*o* mutations) which will be discussed later. As may be seen, all these loci are confined to a very small segment of the chromosome, the *Lac* region. The extreme proximity of all these mutations raises the question whether they belong to a single or to several independent functional units. Such functional analysis requires that the biochemical expression of the various genetic structures be studied in heterozygous diploids. Until quite recently, only transient diploids were available in *E. coli*; the recent discovery of a new type of gene transfer in these bacteria (sexduction) has opened the possibility of obtaining stable clones which are diploid (or polyploid) for different small segments of the chromosome.

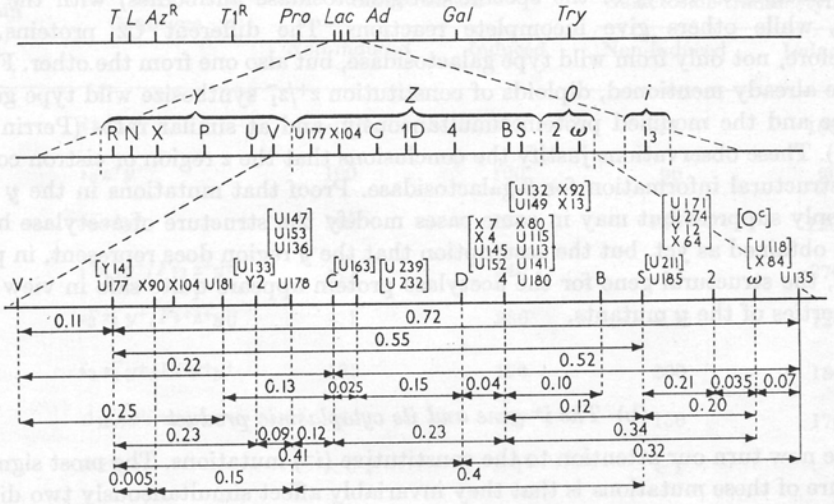



FIG. 3. Diagrammatic map of the lactose region of *E. coli* K12. The upper line represents the position of the *Lac* region with respect to other known markers. The middle line represents an enlargement of the *Lac* region with the four loci *y*, *z*, *o* and *i*. The lower line represents an enlargement of the *z* and *o* loci. Recombination frequencies (given at the bottom) are obtained in two factor crosses of the type $Hfr\ Lac\bar{A}ad^+S^s \times F^-Lac\bar{B}ad^-S^r$, from the ratios "recombinants $Lac^+ad^+S^r$/recombinants $ad^+S^r$." The total length of the *z* gene may be estimated to be 0·7 map units, i.e. about 3,500 nucleotide pairs for about 1,000 amino acids in the monomer of β-galactosidase.

In this process, small fragments of the bacterial chromosome are incorporated into the sex factor, *F*. This new unit of replication is transmissible by conjugation, and is then added to the normal genome of the recipient bacterium which becomes diploid for the small chromosomal fragment. Among the units thus isolated, one carries the whole *Lac* region (Jacob & Adelberg, 1959; Jacob, Perrin, Sanchez & Monod, 1960). To symbolize the genetic structure of these diploids, the chromosomal alleles are written in the usual manner, while the alleles attached to the sex factor are preceded by the letter *F*.

Turning our attention to the behaviour of *z* and *y* mutant types, we may first note that diploids of structure $z^+y^-/Fz^-y^+$ or $z^-y^+/Fz^+y^-$ are wild type, being able to ferment lactose, and forming normal amounts of both galactosidase and acetylase. This complete complementation between $z^-$ and $y^-$ mutants indicates that they belong to independent cistrons. Conversely, no complementation is observed between different $y^-$ mutants, indicating that they all belong to a single cistron. No complementation is observed between most $z^-$ mutants. Certain diploids of structure $z_a^-z_b^+/Fz_a^+z_b^-$ synthesize galactosidase in reduced amounts, but pairs of mutually non-complementing mutants overlap mutually complementing mutants, suggesting again

that a single cistron is involved, as one might expect, since the monomer of galactosidase has a single N-terminal group. It should be recalled that intracistronic partial complementation has been observed in several cases (Giles, 1958), and has (tentatively) been explained as related to a polymeric state of the protein.

Mutations in the $z$ gene affect the structure of galactosidase. This is shown by the fact that most of the $z^-$ mutants synthesize, in place of active enzyme, a protein which is able to displace authentic (wild type) galactosidase from its combination with specific antibody (Perrin, Bussard & Monod, 1959). Among proteins synthesized by different $z^-$ mutants (symbolized $Cz_1$, $Cz_2$, etc.) some give complete cross reactions (i.e. precipitate 100% of the specific antigalactosidase antibodies) with the serum used, while others give incomplete reactions. The different Cz proteins differ therefore, not only from wild type galactosidase, but also one from the other. Finally, as we already mentioned, diploids of constitution $z^+/z_1^-$ synthesize wild type galactosidase and the modified protein simultaneously, and at similar rates (Perrin *et al.*, 1960). These observations justify the conclusions that the $z$ region or cistron contains the structural information for $\beta$-galactosidase. Proof that mutations in the $y$ region not only suppress but may in some cases modify the structure of acetylase has not been obtained as yet, but the assumption that the $y$ region does represent, in part at least, the structural gene for the acetylase protein appears quite safe in view of the properties of the $y$ mutants.

### (b) *The* $i^+$ *gene and its cytoplasmic product*

We now turn our attention to the constitutive ($i^-$) mutations. The most significant feature of these mutations is that they invariably affect simultaneously two different enzyme-proteins, each independently determined, as we have just seen, by different structural genes. In fact, most $i^-$ mutants synthesize more galactosidase and acetylase than induced wild type cells, but it is quite remarkable that the *ratio* of galactosidase to acetylase is the same in the constitutive cells as in the induced wild type, strongly suggesting that the mechanism controlled by the $i$ gene is the same as that with which the inducer interacts.

The study of double heterozygotes of structures: $i^+z^-/Fi^-z^+$ or $i^-y^+/Fi^+y^-$ shows (Table 2, lines 4 and 5) that the inducible $i^+$ allele is dominant over the constitutive and that it is active in the *trans* position, with respect to both $y^+$ and $z^+$.

Therefore the $i$ mutations belong to an independent cistron, governing the expression of $y$ and $z$ *via* a cytoplasmic component. The dominance of the inducible over the constitutive allele means that the former corresponds to the active form of the $i$ gene. This is confirmed by the fact that strains carrying a *deletion* of the *izy* region behave like $i^-$ in diploids (Table 2, line 7). However, two different interpretations of the function of the $i^+$ gene must be considered.

(a) The $i^+$ gene determines the synthesis of a repressor, inactive or absent in the $i^-$ alleles.

(b) The $i^+$ gene determines the synthesis of an enzyme which destroys an inducer, produced by an independent pathway.

The first interpretation is the most straightforward, and it presents the great interest of implying that the fundamental mechanisms of control may be the same in inducible and repressible systems. Several lines of evidence indicate that it is the correct interpretation.

First, we may mention the fact that constitutive synthesis of β-galactosidase by $i^-z^+y^+$ types is not inhibited by thiophenyl-galactoside which has been shown (Cohn & Monod, 1953) to be a competitive inhibitor of induction by exogenous galactosides (see p. 325).

TABLE 2

*Synthesis of galactosidase and galactoside-transacetylase by haploids and heterozygous diploids of regulator mutants*

| Strain No. | Genotype | Galactosidase | | Galactoside-transacetylase | |
|---|---|---|---|---|---|
| | | Non-induced | Induced | Non-induced | Induced |
| 1 | $i^+z^+y^+$ | <0·1 | 100 | <1 | 100 |
| 2 | $i_6^-z^+y^+$ | 100 | 100 | 90 | 90 |
| 3 | $i_3^-z^+y^+$ | 140 | 130 | 130 | 120 |
| 4 | $i^+z_1^-y^+/Fi_3^-z^+y^+$ | <1 | 240 | 1 | 270 |
| 5 | $i_3^-z_1^-y^+/Fi^+z^+y_U^-$ | <1 | 280 | <1 | 120 |
| 6 | $i_3^-z_1^-y^+/Fi^-z^+y^+$ | 195 | 190 | 200 | 180 |
| 7 | $\Delta_{izy}/Fi^-z^+y^+$ | 130 | 150 | 150 | 170 |
| 8 | $i^sz^+y^+$ | <0·1 | <1 | <1 | <1 |
| 9 | $i^sz^+y^+/Fi^+z^+y^+$ | <0·1 | 2 | <1 | 3 |

Bacteria are grown in glycerol as carbon source and induced, when stated, by *iso*propyl-thio-galactoside, $10^{-4}$ M. Values are given as a percentage of those observed with induced wild type (for absolute values, see legend of Table 1). $\Delta_{izy}$ refers to a deletion of the whole *Lac* region. It will be noted that organisms carrying the wild allele of one of the structural genes (*z* or *y*) on the *F* factor form more of the corresponding enzyme than the haploid. This is presumably due to the fact that several copies of the *F-Lac* unit are present per chromosome. In $i^+/i^-$ heterozygotes, values observed with uninduced cells are sometimes higher than in the haploid control. This is due to the presence of a significant fraction of $i^-/i^-$ homozygous recombinants in the population.

A direct and specific argument comes from the study of one particular mutant of the lactose system. This mutant ($i^s$) has lost the capacity to synthesize *both* galactosidase and permease. It is not a deletion because it recombines, giving $Lac^+$ types, with all the $z^-$ and $y^-$ mutants. In crosses with $z^-i^-$ organisms the progeny is *exclusively* $i^-$ while in crosses with $z^-i^+$ it is *exclusively* $i^+$, indicating exceedingly close linkage of this mutation with the *i* region. Finally, in diploids of constitution $i^s/i^+$, $i^s$ turns out to be *dominant*: the diploids cannot synthesize either galactosidase or acetylase (see Table 2, lines 8 and 9).

These unique properties appear exceedingly difficult to account for, except by the admittedly very specific hypothesis that mutant $i^s$ is an allele of *i* where the *structure* of the repressor is such that it cannot be antagonized by the inducer any more. If this hypothesis is correct, one would expect that the $i^s$ mutant could regain the ability to metabolize lactose, not only by reversion to wild type ($i^s \rightarrow i^+$) but also, and probably more frequently, by inactivation of the *i* gene, that is to say by achieving the

constitutive condition ($i^s \rightarrow i^-$). Actually, *Lac*$^+$ "revertants" are very frequent in populations of mutant $i^s$, and 50% of these "revertants" are indeed constitutives of the $i^-$ (recessive) type. (The other revertants are also constitutives, but of the $o^c$ class which we shall mention later.) The properties of this remarkable mutant could evidently not be understood under the assumption that the *i* gene governs the synthesis of an inducer-destroying enzyme (Willson, Perrin, Jacob & Monod, 1961).

Accepting tentatively the conclusion that the $i^+$ gene governs the synthesis of an intracellular repressor, we may now consider the question of the presence of this substance in the cytoplasm, and of its chemical nature.

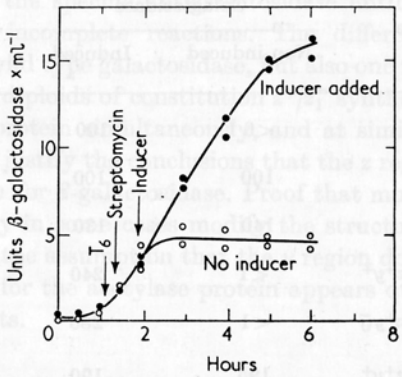



FIG. 4. Synthesis of β-galactosidase by merozygotes formed by conjugation between inducible, galactosidase-positive males and constitutive, galactosidase-negative females. Male (*Hfr* $i^+z^+T6^sS^s$) and female ($F^-$ $i^-z^-T6^rS^r$) bacteria grown in a synthetic medium containing glycerol as carbon source are mixed in the same medium (time O) in the absence of inducer. In such a cross, the first zygotes which receive the *Lac* region from the males are formed from the 20th min. The rate of enzyme synthesis is determined from enzyme activity measurement on the whole population, to which streptomycin and phage T6 are added at times indicated by arrows to block further formation of recombinants and induction of the male parents. It may be seen that in the absence of inducer enzyme synthesis stops about 60 to 80 min after penetration of the first $z^+i^+$ segment but is resumed by addition of inducer (From Pardee *et al*, 1959).

Important indications on this question have been obtained by studying the kinetics and conditions of expression of the $i^+$ and $z^+$ genes when they are introduced into the cytoplasm of cells bearing the inactive ($z^-$ and $i^-$) alleles. The sexual transfer of the *Lac* segment from male to female cells provides an adequate experimental system for such studies. It should be recalled that conjugation in *E. coli* involves essentially the transfer of a male chromosome (or chromosome segment) to the female cell. This transfer is oriented, always beginning at one extremity of the chromosome, and it is progressive, each chromosome segment entering into the recipient cell at a fairly precise time following inception of conjugation in a given mating pair (Wollman & Jacob, 1959). The conjugation does not appear to involve any significant cytoplasmic mixing, so that the zygotes inherit virtually all their cytoplasm from the female cell, receiving only a chromosome or chromosome segment from the male. In order to study galactosidase synthesis by the zygotes, conditions must be set up such that the unmated parents cannot form the enzymes. This is the case when mating between inducible galactosidase-positive, streptomycin-sensitive males ($\sigma^{\!\!\!

become streptomycin sensitive (because the $Sm^s$ gene is transferred only to a small proportion of them, and at a very late time). The results of such an experiment, performed in the absence of inducer, are shown in Fig. 4. It is seen that galactosidase synthesis starts almost immediately following actual entry of the $z^+$ gene. We shall return later to a more precise analysis of the expression of the $z^+$ gene. The important point to be stressed here is that during this initial period the zygotes behave like *constitutive* cells, synthesizing enzyme in the *absence* of inducer. Approximately sixty minutes later, however, the rate of galactosidase synthesis falls off to zero. If at that time inducer is added, the maximum rate of enzyme synthesis is resumed. We are, in other words, witnessing the conversion of the originally $i^-$ phenotype of the zygote cell, into an $i^+$ phenotype. And this experiment clearly shows that the "inducible" state is associated with the presence, at a sufficient level, of a *cytoplasmic* substance synthesized under the control of the $i^+$ gene. (It may be pointed out that the use of a female strain carrying a *deletion* of the *Lac* region instead of the $i^-z^-$ alleles gives the same results (Pardee *et al.*, 1959).)

If now 5-methyltryptophan is added to the mated cells a few minutes before entry of the $z^+$ gene, no galactosidase is formed because, as is well known, this compound inhibits tryptophan synthesis by retro-inhibition, and therefore blocks protein synthesis. If the repressor is a protein, or if it is formed by a specific enzyme, the synthesis of which is governed by the $i^+$ gene, its accumulation should also be blocked. If on the other hand the repressor is not a protein, and if its synthesis does not require the preliminary synthesis of a specific enzyme controlled by the $i^+$ gene, it may accumulate in presence of 5-methyltryptophan which is known (Gros, unpublished results) *not* to inhibit energy transfer or the synthesis of nucleic acids.

The results of Pardee & Prestidge (1959) show that the repressor *does* accumulate under these conditions, since the addition of tryptophan 60 min after 5-methyltryptophan allows immediate and complete resumption of enzyme synthesis, *but only in the presence of inducer*; in other words, the cytoplasm of the zygote cells has been converted from the constitutive to the inducible state during the time that protein synthesis was blocked. This result has also been obtained using chloramphenicol as the agent for blocking protein synthesis, and it has been repeated using another system of gene transfer (Luria *et al.*, unpublished results).

This experiment leads to the conclusion that the repressor is not a protein, and this again excludes the hypothesis that the $i^+$ gene controls an inducer-destroying enzyme. We should like to stress the point that this conclusion does not imply that no enzyme is involved in the synthesis of the repressor, but that the enzymes which may be involved are *not* controlled by the $i^+$ gene. The experiments are negative, as far as the chemical nature of the repressor itself is concerned, since they only eliminate protein as a candidate. They do, however, invite the speculation that the repressor may be the primary product of the $i^+$ gene, and the further speculation that such a primary product may be a polyribonucleotide.

Before concluding this section, it should be pointed out that constitutive mutations have been found in several inducible systems; in fact wherever they have been searched for by adequate selective techniques (amylomaltase of *E. coli* (Cohen-Bazire & Jolit, 1953), penicillinase of *B. cereus* (Kogut, Pollock & Tridgell, 1956), glucuronidase of *E. coli* (F. Stoeber, unpublished results), galactokinase and galactose-transferase (Buttin, unpublished results)). That *any* inducible system should be potentially capable of giving rise to constitutive mutants, strongly indicates that such mutations occur, or at least can always occur, by a loss of function. In the case of the "galactose"

system of *E. coli*, it has been found that the constitutive mutation is pleiotropic, affecting a sequence of three different enzymes (galactokinase, galactose-transferase, UDP-galactose epimerase), and occurs at a locus distinct from that of the corresponding structural genes (Buttin, unpublished results).

The main conclusions from the observations reviewed in this section may be summarized as defining a new type of gene, which we shall call a "regulator gene" (Jacob & Monod, 1959). A regulator gene does not contribute structural information to the proteins which it controls. The specific product of a regulator gene is a cytoplasmic substance, which inhibits information transfer from a structural gene (or genes) to protein. In contrast to the classical structural gene, a regulator gene may control the synthesis of several different proteins: the one-gene one-protein rule does not apply to it.

We have already pointed out the profound similarities between induction and repression which suggest that the two effects represent different manifestations of the same fundamental mechanism. If this is true, and if the above conclusions are valid, one expects to find that the genetic control of repressible systems also involves regulator genes.

### (c) *Regulator genes in repressible systems*

The identification of constitutive or "de-repressed" mutants of several repressible systems has fulfilled this expectation. For the selection of such mutants, certain analogs of the normal repressing metabolite may be used as specific selective agents, because they cannot substitute for the metabolite, except as repressing metabolites. For instance, 5-methyltryptophan does not substitute for tryptophan in protein synthesis (Munier, unpublished results), but it represses the enzymes of the tryptophan-synthesizing sequence (Monod & Cohen-Bazire, 1953). Normal wild type *E. coli* does not grow in the presence of 5-methyltryptophan. Fully resistant stable mutants arise, however, a large fraction of which turn out to be constitutive for the tryptophan system.† The properties of these organisms indicate that they arise by mutation of a regulator gene $R_T$ (Cohen & Jacob, 1959). In these mutants tryptophan-synthetase as well as at least two of the enzymes involved in previous steps in the sequence are formed at the same rate irrespective of the presence of tryptophan, while in the wild type all these enzymes are strongly repressed. Actually the mutants form more of the enzymes in the presence of tryptophan, than does the wild type in its absence (just as $i^-z^+$ mutants form more galactosidase in the absence of inducer than the wild type does at saturating concentration of inducer). The capacity of the mutants to concentrate tryptophan from the medium is not impaired, nor is their tryptophanase activity increased. The loss of sensitivity to tryptophan as repressing metabolite cannot therefore be attributed to its destruction by, or exclusion from, the cells, and can only reflect the breakdown of the control system itself. Several recurrences of the $R_T$ mutation have been mapped. They are all located in the same small section of the chromosome, at a large distance from the cluster of genes which was shown by Yanofsky & Lennox (1959) to synthesize the different enzymes of the sequence. One of these genes (comprising two cistrons) has been very clearly identified by the work of Yanofsky (1960) as the structural gene for tryptophan synthetase, and it is a safe assumption that the other genes in this cluster determine the structure of the preceding

† Resistance to 5-methyltryptophan may also arise by other mechanisms in which we are not interested here.

enzymes in the sequence. The $R_T$ gene therefore controls the rate of synthesis of several different proteins without, however, determining their structure. It can only do so *via* a cytoplasmic intermediate, since it is located quite far from the structural genes. To complete its characterization as a regulator gene, it should be verified that the constitutive ($R_T^-$) allele corresponds to the inactive state of the gene (or gene product), i.e. is recessive. Stable heterozygotes have not been available in this case, but the transient (sexual) heterozygotes of a cross $\sigma R_T^- \times \varphi R_T^+$ are sensitive to 5-methyltryptophan, indicating that the repressible allele is dominant (Cohen & Jacob, 1959).

In the arginine-synthesizing sequence there are some seven enzymes, simultaneously repressible by arginine (Vogel, 1957*a*,*b*; Gorini & Maas, 1958). The specific (i.e. probably structural) genes which control these enzymes are dispersed at various loci on the chromosome. Mutants resistant to canavanine have been obtained, in which several (perhaps all) of these enzymes are simultaneously de-repressed. These mutations occur at a locus (near $Sm^r$) which is widely separated from the loci corresponding (probably) to the structural genes. The dominance relationships have not been analysed (Gorini, unpublished results; Maas, Lavallé, Wiame & Jacob, unpublished results).

The case of alkaline phosphatase is particularly interesting because the structural gene corresponding to this protein is well identified by the demonstration that various mutations at this locus result in the synthesis of altered phosphatase (Levinthal, 1959). The synthesis of this enzyme is repressed by orthophosphate (Torriani, 1960). Constitutive mutants which synthesize large amounts of enzyme in the presence of orthophosphate have been isolated. They occur at two loci, neither of which is allelic to the structural gene, and the constitutive enzyme is identical, by all tests, to the wild type (repressible) enzyme. The constitutive alleles for both of the two loci have been shown to be recessive with respect to wild type. Conversely, mutations in the structural ($P$) gene do not affect the regulatory mechanism, since the altered (inactive) enzyme formed by mutants of the $P$ gene is repressed in the presence of orthophosphate to the same extent as the wild type enzyme (Echols, Garen, Garen & Torriani, 1961).

### (d) *The interaction of repressors, inducers and co-repressors*

The sum of these observations leaves little doubt that repression, like induction, is controlled by specialized regulator genes, which operate by a basically similar mechanism in both types of systems, namely by governing the synthesis of an intracellular substance which inhibits information transfer from structural genes to protein.

It is evident therefore that the metabolites (such as tryptophan, arginine, orthophosphate) which inhibit enzyme synthesis in repressible systems are not active by themselves, but only by virtue of an interaction with a repressor synthesized under the control of a regulator gene. Their action is best described as an activation of the genetically controlled repression system. In order to avoid confusion of words, we shall speak of repressing metabolites as "co-repressors" reserving the name "repressors" (or apo-repressors) for the cytoplasmic products of the regulator genes.

The nature of the interaction between repressor and co-repressor (in repressible systems) or inducer (in inducible systems) poses a particularly difficult problem. As a purely formal description, one may think of inducers as antagonists, and of co-repressors as activators, of the repressor. A variety of chemical models can be imagined

to account for such antagonistic or activating interactions. We shall not go into these speculations since there is at present no evidence to support or eliminate any particular model. But it must be pointed out that, in any model, the structural specificity of inducers or co-repressors must be accounted for, and can be accounted for, only by the assumption that a stereospecific receptor is involved in the interaction. The fact that the repressor is apparently not a protein then raises a serious difficulty since the capacity to form stereospecific complexes with small molecules appears to be a privilege of proteins. If a protein, perhaps an enzyme, is responsible for the specificity, the structure of this protein is presumably determined by a structual gene and mutation in this gene would result in loss of the capacity to be induced (or repressed). Such mutants, which would have precisely predictable properties (they would be pleiotropic, recessive, and they would be complemented by mutants of the other structural genes) have not been encountered in the lactose system, while the possibility that the controlled enzymes themselves (galactosidase or acetylase) play the role of "induction enzyme" is excluded.

It is conceivable that, in the repressible systems which synthesize amino acids, this role is played by enzymes simultaneously responsible for essential functions (e.g. the activating enzymes) whose loss would be lethal, but this seems hardly conceivable in the case of most inducible systems. One possibility which is not excluded by these observations is that the repressor itself synthesizes the "induction protein" and remains thereafter associated with it. Genetic inactivation of the induction enzyme would then be associated with structural alterations of the repressor itself and would generally be expressed as constitutive mutations of the regulator gene.† This possibility is mentioned here only as an illustration of the dilemma which we have briefly analysed, and whose solution will depend upon the chemical identification of the repressor.

### (e) *Regulator genes and immunity in temperate phage systems*

One of the most conspicuous examples of the fact that certain genes may be either allowed to express their potentialities, or specifically prohibited from doing so, is the phenomenon of immunity in temperate phage systems (*cf.* Lwoff, 1953; Jacob, 1954; Jacob & Wollman, 1957; Bertani, 1958; Jacob, 1960).

The genetic material of the so-called temperate phages can exist in one of two states within the host cell:

(1) In the *vegetative state*, the phage genome multiplies autonomously. This process, during which all the phage components are synthesized, culminates in the production of infectious phage particles which are released by lysis of the host cell.

(2) In the *prophage state*, the genetic material of the phage is attached to a specific site of the bacterial chromosome in such a way that both genetic elements replicate as a single unit. The host cell is said to be "lysogenic." As long as the phage genome remains in the prophage state, phage particles are not produced. For lysogenic bacteria to produce phage, the genetic material of the phage must undergo a transition from the prophage to the vegetative state. During normal growth of lysogenic bacteria, this event is exceedingly rare. With certain types of prophages, however, the transition can be induced in the whole population by exposure of the culture to u.v. light,

† Such a model could account for the properties of the $i^s$ (dominant) mutant of the regulator gene in the lactose system, by the assumption that in this mutant the repressor remains active, while having lost the capacity to form its associated induction protein.

X-rays or various compounds known to alter DNA metabolism (Lwoff, Siminovitch & Kjeldgaard, 1950; Lwoff, 1953; Jacob, 1954).

The study of "defective" phage genomes, in which a mutation has altered one of the steps required for the production of phage particles, indicates the existence of at least two distinct groups of viral functions, both of which are related to the capacity of synthesizing specific proteins (Jacob, Fuerst & Wollman, 1957). Some "early" functions appear as a pre-requisite for the vegetative multiplication of the phage genome and, at least in virulent phages of the T-even series, it is now known that they correspond to the synthesis of a series of new enzymes (Flaks & Cohen, 1959; Kornberg, Zimmerman, Kornberg & Josse, 1959). A group of "late" functions correspond to the synthesis of the structural proteins which constitute the phage coat. The expression of these different viral functions appears to be in some way co-ordinated by a sequential process, since defective mutations affecting some of the early functions may also result in the loss of the capacity to perform several later steps of phage multiplication (Jacob *et al.*, 1957).

In contrast, the viral functions are not expressed in the prophage state and the protein constituents of the phage coat cannot be detected within lysogenic bacteria. In addition, lysogenic bacteria exhibit the remarkable property of being specifically *immune* to the very type of phage particles whose genome is already present in the cell as prophage. When lysogenic cells are infected with homologous phage particles, these particles absorb onto the cells and inject their genetic material, but the cell survives. The injected genetic material does not express its viral functions: it is unable to initiate the synthesis of the protein components of the coat and to multiply vegetatively. It remains inert and is diluted out in the course of bacterial multiplication (Bertani, 1953; Jacob, 1954).

The inhibition of phage-gene functions in lysogenic bacteria therefore applies not only to the prophage, but also to additional homologous phage genomes. It depends only upon the presence of the prophage (and not upon a permanent alteration, provoked by the prophage, of bacterial genes) since loss of the prophage is both necessary and sufficient to make the bacteria sensitive again.

Two kinds of interpretation may be considered to account for these "immunity" relationships:

(a) The prophage occupies and blocks a *chromosomal* site of the host, specifically required in some way for the vegetative multiplication of the homologous phage.

(b) The prophage produces a *cytoplasmic* inhibitor preventing the completion of some reactions (presumably the synthesis of a particular protein) necessary for the initiation of vegetative multiplication.

A decision between these alternative hypotheses may be reached through the study of persistent diploids, heterozygous for the character lysogeny. A sex factor has been isolated which has incorporated a segment of the bacterial chromosome carrying the genes which control galactose fermentation, *Gal*, and the site of attachment of prophage, $\lambda$. Diploid heterozygotes with the structure $Gal^- \lambda^-/F\ Gal^+ \lambda^+$ or $Gal^- \lambda^+/F\ Gal^+ \lambda^-$ are immune against superinfection with phage $\lambda$, a result which shows that "immunity" is dominant over "non-immunity" and has a cytoplasmic expression (Jacob, Schaeffer & Wollman, 1960).

The study of transient zygotes formed during conjugation between lysogenic ($\lambda^+$) and non-lysogenic ($\lambda^-$) cells leads to the same conclusion. In crosses $\male\lambda^+ \times \female\lambda^-$, the transfer of the prophage carried by the male chromosome into the non-immune

recipient results in transition to the vegetative state: multiplication of the phage occurs in the zygotes, which are lysed and release phage particles. This phenomenon is known as "zygotic induction" (Jacob & Wollman, 1956). In the *reverse* cross $\male\lambda^- \times \female\lambda^+$, however, *no zygotic induction occurs*. The transfer of the "non-lysogenic" character carried by the male chromosome into the immune recipient does not bring about the development of the prophage and the zygotes are immune against superinfection with phage $\lambda$.

The opposite results obtained in reciprocal crosses of lysogenic by non-lysogenic male and female cells are entirely analogous to the observations made with the lactose system in reciprocal crosses of inducible by non-inducible cells. In both cases, it is evident that the decisive factor is the origin of the *cytoplasm* of the zygote, and the conclusion is inescapable, that the immunity of lysogenic bacteria is due to a cytoplasmic constituent, in the presence of which the viral genes cannot become expressed (Jacob, 1960).

The same two hypotheses which we have already considered for the interpretation of the product of the regulator gene in the lactose system, apply to the cytoplasmic inhibitor insuring immunity in lysogenic bacteria.

(a) The inhibitor is a specific repressor which prevents the synthesis of some early protein(s) required for the initiation of vegetative multiplication.

(b) The inhibitor is an enzyme which destroys a metabolite, normally synthesized by the non-lysogenic cell and specifically required for the vegetative multiplication of the phage.

Several lines of evidence argue against the second hypothesis (Jacob & Campbell, 1959; Jacob, 1960). First, for a given strain of bacteria, many temperate phages are known, each of which exhibits a different immunity pattern. According to the second hypothesis, each of these phages would specifically require for vegetative multiplication a different metabolite normally produced by the non-lysogenic cells, an assumption which appears extremely unlikely. The second argument stems from the fact that, like the repressor of the lactose system, the inhibitor responsible for immunity is synthesized in the presence of chloramphenicol, i.e. in the absence of protein synthesis: when crosses $\male\lambda^+ \times \female\lambda^-$ are performed in the presence of chloramphenicol, no zygotic induction occurs and the prophage is found to segregate normally among recombinants.

In order to explain immunity in lysogenic bacteria, we are led therefore to the same type of interpretation as in the case of adaptive enzyme systems. According to this interpretation, the prophage controls a cytoplasmic repressor, which inhibits specifically the synthesis of one (or several) protein(s) necessary for the initiation of vegetative multiplication. In this model, the introduction of the genetic material of the phage into a non-lysogenic cell, whether by infection or by conjugation, results in a "race" between the synthesis of the specific repressor and that of the early proteins required for vegetative multiplication. The fate of the host-cell, survival with lysogenization or lysis as a result of phage multiplication, depends upon whether the synthesis of the repressor or that of the protein is favoured. Changes in the cultural conditions favoring the synthesis of the repressor such as infection at low temperature, or in the presence of chloramphenicol, would favor lysogenization and *vice versa*. The phenomenon of induction by u.v. light could then be understood, for instance, in the following way: exposure of inducible lysogenic bacteria to u.v. light or X-rays would transiently disturb the regulation system, for example by preventing further synthesis of the repressor. If the repressor is unstable, its concentration inside the cell would

decrease and reach a level low enough to allow the synthesis of the early proteins. Thus the vegetative multiplication would be irreversibly initiated.

The similarity between lysogenic systems and adaptive systems is further strengthened by the genetic analysis of immunity. Schematically, the genome of phage λ appears to involve two parts (see Fig. 5): a small central segment, the *C* region, contains a few determinants which control various functions involved in lysogenization (Kaiser, 1957); the rest of the linkage group contains determinants which govern the "viral functions," i.e. presumably the structural genes corresponding to the different phage proteins. Certain strains of temperate phages which exhibit different immunity patterns are able nevertheless to undergo genetic recombination. The specific immunity pattern segregates in such crosses, proving to be controlled by a small segment "*im*" of the *C* region (Kaiser & Jacob, 1957). In other words, a prophage contains in its *C* region a small segment "*im*" which controls the synthesis of a specific repressor, active on the phage genome carrying a homologous "*im*" segment.

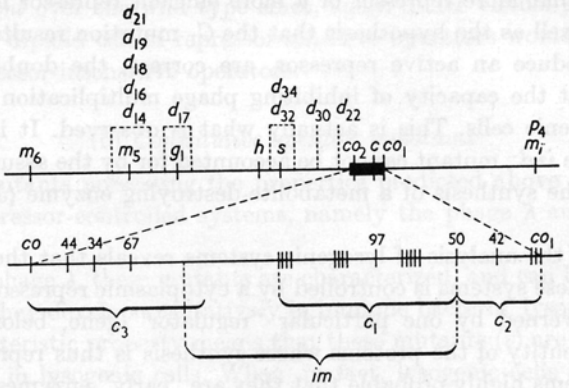



FIG. 5. Diagrammatic representation of the linkage group of the temperate bacteriophage λ. The upper diagram represents the linear arrangement of markers. Symbols refer to various plaque size, plaque type and host-range markers. Symbols *d* refer to various defective mutations. The *C* region represented by a thicker line is enlarged in the lower diagram. The figures correspond to various *C* mutations. The *C* region can be subdivided into three functional units, $C_1$, $C_2$ and $C_3$; the segment controlling immunity is designated *im*.

In the "*im*" region, two types of mutations arise, whose properties are extremely similar to those of the different mutations affecting the regulator genes of adaptive enzyme systems.

(1) Some mutations ($C_I^+ \rightarrow C_I$) result in the complete loss of the capacity for lysogenization in single infection. All the $C_I$ mutations are located in a cluster, in a small part of the "*im*" segment, and they behave as belonging to a single cistron in complementation tests.

In mixed infections with both $C_I$ and $C_I^+$ phages, double lysogenic clones carrying both $C_I$ and $C_I^+$ prophages can be recovered. In such clones, single lysogenic cells segregate, which carry the $C^+$ type alone but never the $C_I$ type alone. These findings indicate that the wild allele is dominant over the mutant $C_I$ alleles and is cytoplasmically expressed, repressing the mutant genome into the prophage state. The properties of the $C_I$ mutations are therefore similar to those of the recessive constitutive mutations of adaptive systems. The evidence suggests that the $C_I$ locus controls the synthesis of the repressor responsible for immunity, and that the $C_I$ mutations correspond to inactivation of this locus, or of its product.

(2) A mutation ($ind^+ \rightarrow ind^-$) has been found which results in the loss of the inducible property of the prophage, i.e. of its capacity to multiply vegetatively upon exposure of lysogenic bacteria to u.v. light, X-rays or chemical inducers. This mutation is located in the $C_I$ segment. The mutant allele $ind^-$ is dominant over the wild allele $ind^+$ since double lysogenic $\lambda ind^+/\lambda ind^-$ or diploid heterozygotes of structure $Gal^- \lambda ind^+/F\ Gal^+ \lambda ind^-$ or $Gal^- \lambda ind^-/F\ Gal^+ \lambda ind^+$ are all non-inducible. In addition, the mutant $\lambda ind^-$ exhibits a unique property. If lysogenic bacteria K12 ($\lambda^+$) carrying a wild type prophage are exposed to u.v. light, the whole population lyses and releases phage. Infection of such cells with $\lambda ind^-$ mutants, either before or immediately after irradiation, completely inhibits phage production and lysis.

The properties of the $ind^-$ mutant appear in every respect similar to those of the previously described mutant $i^s$ of the lactose system. The unique properties of the $ind^-$ mutants can be explained only by the same type of hypothesis, namely that the mutation $ind^-$ affects, quantitatively or qualitatively, the synthesis of the repressor in such a way that more repressor or a more efficient repressor is produced. If this assumption as well as the hypothesis that the $C_I$ mutation results in the loss of the capacity to produce an active repressor, are correct, the double mutants $C_I ind^-$ should have lost the capacity of inhibiting phage multiplication upon infection of wild type lysogenic cells. This is actually what is observed. It is evident that the properties of the $ind^-$ mutant cannot be accounted for by the assumption that the $C_I$ locus controls the synthesis of a metabolite-destroying enzyme (Jacob & Campbell, 1959).

In summary, the analysis of lysogenic systems reveals that the expression of the viral genes in these systems is controlled by a cytoplasmic repressor substance, whose synthesis is governed by one particular "regulator" gene, belonging to the viral genome. The identity of the proteins whose synthesis is thus repressed is not established, but it seems highly probable that they are "early" enzymes which initiate the whole process of vegetative multiplication. With the (important) limitation that they are sensitive to entirely different types of inducing conditions, the phage repression systems appear entirely comparable to the systems involved in enzymatic adaptation.

## 4. The Operator and the Operon

### (a) *The operator as site of action of the repressor*

In the preceding section we have discussed the evidence which shows that the transfer of information from structural genes to protein is controlled by specific repressors synthesized by specialized regulator genes. We must now consider the next problem, which is the site and mode of action of the repressor.

In regard to this problem, the most important property of the repressor is its characteristic pleiotropic specificity of action. In the lactose system of *E. coli*, the repressor is both *highly specific* since mutations of the *i* gene do not affect any other system, and *pleiotropic* since both galactosidase and acetylase are affected simultaneously and quantitatively to the same extent, by such mutations.

The specificity of operation of the repressor implies that it acts by forming a stereospecific combination with a constituent of the system possessing the proper (complementary) molecular configuration. Furthermore, it must be assumed that the flow of information from gene to protein is interrupted when this element is combined with

the repressor. This controlling element we shall call the "*operator*" (Jacob & Monod, 1959). We should perhaps call attention to the fact that, once the existence of a specific repressor is considered as established, the existence of an operator element defined as above follows necessarily. Our problem, therefore, is not whether an operator exists, but where (and how) it intervenes in the system of information transfer.

An important prediction follows immediately from the preceding considerations. Under any hypothesis concerning the nature of the operator, its specific complementary configuration must be genetically determined; therefore it could be affected by mutations which would alter or abolish its specific affinity for the repressor, without necessarily impairing its activity as initiator of information-transfer. Such mutations would result in *constitutive* synthesis of the protein or proteins. These mutations would define an "operator locus" which should be genetically distinct from the regulator gene (i.e. its mutations should not behave as alleles of the regulator); the most distinctive predictable property of such mutants would be that the constitutive allele should be *dominant* over the wild type since, again under virtually any hypothesis, the presence in a diploid cell of repressor-sensitive operators would not prevent the operation of repressor-insensitive operators.

### (b) *Constitutive operator mutations*

Constitutive mutants possessing the properties predicted above have so far been found in two repressor-controlled systems, namely the phage $\lambda$ and *Lac* system of *E. coli*.

In the case of phage $\lambda$, these mutants are characterized, and can be easily selected, by the fact that they develop vegetatively in immune bacteria, lysogenic for the wild type. This characteristic property means that these mutants ($v$) are *insensitive* to the repressor present in lysogenic cells. When, in fact, lysogenic cells are infected with these mutant particles, the development of the wild type prophage is induced, and the resulting phage population is a mixture of $v$ and $v^+$ particles. This is expected, since presumably the initiation of prophage development depends only on the formation of one or a few "early" enzyme-proteins, which are supplied by the virulent particle (Jacob & Wollman, 1953).

In the *Lac* system, dominant constitutive ($o^c$) mutants have been isolated by selecting for constitutivity in cells diploid for the *Lac* region, thus virtually eliminating the recessive ($i^-$) constitutive mutants (Jacob *et al.*, 1960*a*). By recombination, the $o^c$ mutations can be mapped in the *Lac* region, between the $i$ and the $z$ loci, the order being (*Pro*) *yzoi* (*Ad*) (see Fig. 3). Some of the properties of these mutants are summarized in Table 3. To begin with, let us consider only the effects of this mutation on galactosidase synthesis. It will be noted that in the absence of inducer, these organisms synthesize 10 to 20% of the amount of galactosidase synthesized by $i^-$ mutants, i.e. about 100 to 200 times more than uninduced wild type cells (Table 3, lines 3 and 7). In the presence of inducer, they synthesize maximal amounts of enzyme. They are therefore only partially constitutive (except however under conditions of starvation, when they form maximum amounts of galactosidase in the absence of inducer (Brown, unpublished results)). The essential point however is that the enzyme is synthesized constitutively by diploid cells of constitution $o^c/o^+$ (see Table 3). The $o^c$ allele therefore is "dominant."

If the constitutivity of the $o^c$ mutant results from a loss of sensitivity of the operator to the repressor, the $o^c$ organisms should also be insensitive to the presence of the

altered repressor synthesized by the $i^s$ (dominant) allele of the $i^+$ gene (see page 331). That this is indeed the case, as shown by the constitutive behavior of diploids with the constitution $i^s o^+/Fi^+o^c$ (see Table 3, line 12), is a very strong confirmation of the interpretation of the effects of *both* mutations ($i^s$ and $o^c$). In addition, and as one would expect according to this interpretation, $o^c$ mutants frequently arise as lactose positive "revertants" in populations of $i^s$ cells (see p. 332).

TABLE 3

*Synthesis of galactosidase, cross-reacting material (CRM), and galactoside-transacetylase by haploid and heterozygous diploid operator mutants*

| Strain No. | Genotype | Galactosidase | | Cross-reacting material | |
|---|---|---|---|---|---|
| | | Non-induced | Induced | Non-induced | Induced |
| 1 | $o^+z^+$ | <0·1 | 100 | — | — |
| 2 | $o^+z^+/Fo^+z_1^-$ | <0·1 | 105 | <1 | 310 |
| 3 | $o^cz^+$ | 15 | 90 | — | — |
| 4 | $o^+z^+/Fo^cz_1$ | <0·1 | 90 | 30 | 180 |
| 5 | $o^+z_1^-/Fo^cz^+$ | 90 | 250 | <1 | 85 |
| | | Galactosidase | | Galactoside-transacetylase | |
| | | Non-Induced | Induced | Non-induced | Induced |
| 6 | $o^+z^+y^+$ | <0·1 | 100 | <1 | 100 |
| 7 | $o^cz^+y^+$ | 25 | 95 | 15 | 110 |
| 8 | $o^+z^+y_U^-/Fo^cz^+y^+$ | 70 | 220 | 50 | 160 |
| 9 | $o^+z_1^-y^+/Fo^cz^+y_U^-$ | 180 | 440 | <1 | 220 |
| 10 | $i^+o_{84}^cz^+y^+$ | <0·1 | <0·1 | <1 | <1 |
| 11 | $i^+o_{84}^cz^+y^+/Fi^-o^+z^+y^+$ | 1 | 260 | 2 | 240 |
| 12 | $i^so^+z^+y^+/Fi^+o^cz^+y^+$ | 190 | 210 | 150 | 200 |

Bacteria are grown in glycerol as carbon source and induced when stated, with *iso*propyl-thiogalactoside, $10^{-4}$ M. Values of galactosidase and acetylase are given as a percentage of those observed with induced wild type. Values of CRM are expressed as antigenic equivalents of galactosidase. Note that the proteins corresponding to the alleles carried by the sex factor are often produced in greater amount than that observed with induced haploid wild type. This is presumably due to the existence of several copies of the *F-Lac* factor per chromosome. In $o^c$ mutants, haploid or diploid, the absolute values of enzymes produced, especially in the non-induced cultures varies greatly from day to day depending on the conditions of the cultures.

We therefore conclude that the $o^+ \rightarrow o^c$ mutations correspond to a modification of the specific, repressor-accepting, structure of the operator. This identifies the operator locus, i.e. the genetic segment responsible for the structure of the operator, but not the operator itself.

(c) *The operon*

Turning now to this problem, we note that the $o^c$ mutation (like the $i^-$ mutation) is pleiotropic: it affects simultaneously and quantitatively to the same extent, the synthesis of galactosidase and acetylase (see Table 3, lines 7 and 8). The structure of the operator, or operators, which controls the synthesis of the two proteins, therefore, is controlled by a single determinant.†

Two alternative interpretations of this situation must be considered:

(a) A single operator controls an *integral* property of the *z-y* genetic segment, or of its cytoplasmic product.

(b) The specific product of the operator locus is able to associate in the cytoplasm, with the products of the *z* and *y* cistrons, and thereby governs the expression of both structural genes.

The second interpretation implies that mutations of the operator locus should behave as belonging to a cistron *independent* of both the *z* and *y* cistrons. The first interpretation requires, on the contrary, that these mutations behave functionally as if they *belonged to both cistrons simultaneously*. These alternative interpretations can therefore be distinguished without reference to any particular physical model of operator action by testing for the *trans* effect of *o* alleles, that is to say for the constitutive *vs* inducible expression of the two structural genes in $o^+/o^c$ diploids, heterozygous for one or both of these structural genes.

The results obtained with diploids of various structures are shown in Table 3. We may first note that in diploids of constitution $o^+z^+/Fo^cz_1^-$ or $o^+z_1^-/Fo^cz^+$ (lines 4 and 5), both the normal galactosidase produced by the $z^+$ allele and the altered protein (CRM) produced by the $z_1^-$ allele are formed in the presence of inducer, while in the *absence* of inducer, *only the protein corresponding to the* z *allele in position cis to the* $o^c$ *is produced.* The $o^c$ therefore has no effect on the *z* allele in position *trans*. Or putting it otherwise: the expression of the *z* allele attached to an $o^+$ remains fully repressor-sensitive even in the presence of an $o^c$ in position *trans*. The *o* locus might be said to behave as belonging to the same cistron as the *z* markers. But as we know already, the $o^c$ mutation is equally effective towards the acetylase which belongs to a cistron independent of *z*, and not adjacent to the operator locus. The results shown in Table 3, lines 8 and 9, confirm that the $o \rightarrow y$ relationship is the same as the $o \rightarrow z$ relationship, that is, the effect of the $o^c$ allele extends *exclusively* to the *y* allele in the *cis* position. For instance, in the diploid $o^+z^-y^+/Fo^cz^+y_U^-$ the galactosidase is constitutive and the acetylase is inducible, while in the diploid $o^+z^+y_U^-/Fo^cz^+y^+$ both enzymes are constitutive.

These observations, predicted by the first interpretation, are incompatible with the second and lead to the conclusion that the operator governs an integral property of the genetic segment *ozy*, or of its cytoplasmic product (Jacob *et al.*, 1960*a*; Képès, Monod & Jacob, 1961).

This leads to another prediction. Certain mutations of the *o* segment could modify the operator in such a way as to inactivate the whole *ozy* segment resulting in the loss of the capacity to synthesize *both* galactosidase and permease.

These "$o^o$" mutants would be *recessive* to $o^+$ or $o^c$, and they would *not* be complemented either by $o^+z^+y^-$ or by $o^+z^-y^+$ mutants. Several point-mutants, possessing

† Let us recall again that no *non-pleiotropic* constitutive mutants of any type have been isolated in this system, in spite of systematic screening for such mutants.

precisely these properties, have been isolated (Jacob *et al.*, 1960*a*). They all map very closely to $o^c$, as expected (see Fig. 3). It is interesting to note that in these mutants the $i^+$ gene is functional (Table 3, line 11), which shows clearly, not only that the *i* and *o* mutants are not alleles, but that the *o* segment, while governing the expression of the *z* and *y* genes, does not affect the expression of the regulator gene.

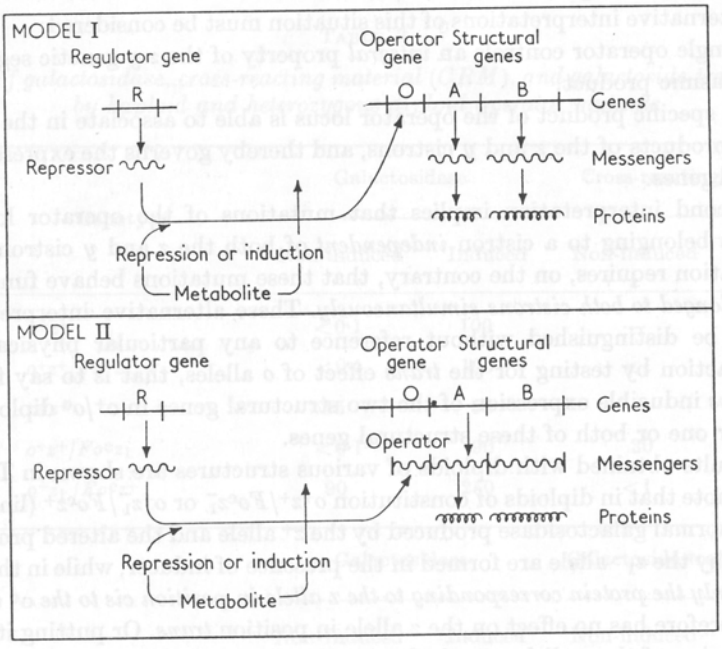



FIG. 6. Models of the regulation of protein synthesis.

In conclusion, the integral or *co-ordinate* expression of the *ozy* genetic segment signifies that the operator, which controls this expression, is and remains attached (see Fig. 6):

(a) either to the genes themselves (Fig. 6, I),

(b) or to the cytoplasmic messenger of the linked *z* and *y* genes which must then be assumed to form a single, integral, particle corresponding to the structure of the whole *ozy* segment, and functioning as a whole (Fig. 6, II).

In the former case, *the operator would in fact be identical with the* o *locus* and it would govern directly the activity of the genes, i.e. the synthesis of the structural messengers.

Both of these models are compatible with the observations which we have discussed so far. We shall return in the next section to the question whether the operator, i.e. the site of specific interaction with the repressor, is genetic or cytoplasmic. In either case, the *ozy* segment, although containing at least two independent structural genes, governing two independent proteins, behaves as a *unit* in the transfer of information. This *genetic unit of co-ordinate expression* we shall call the "*operon*" (Jacob *et al.*, 1960*a*).

The existence of such a unit of genetic expression is proved so far only in the case of the *Lac* segment. As we have already seen, the *v* mutants of phage λ, while illustrating the existence of an operator in this system, do not define an operon (because the

number and the functions of the structural genes controlled by this operator are unknown). However, many observations hitherto unexplained by or even conflicting with classical genetic theory, are immediately accounted for by the operon theory. It is well known that, in bacteria, the genes governing the synthesis of different enzymes sequentially involved in a metabolic pathway are often found to be extremely closely linked, forming a cluster (Demerec, 1956). Various not very convincing speculations have been advanced to account for this obvious correlation of genetic structure and biochemical function (see Pontecorvo, 1958). Since it is now established that simultaneous induction or repression also generally prevails in such metabolic sequences, it seems very likely that the gene clusters represent units of co-ordinate expression, i.e. operons.

We have already mentioned the fact that two inducible enzymes sequentially involved in the metabolism of galactose by *E. coli*, galactokinase and UDP-galactose-transferase, are simultaneously induced by galactose, or by the gratuitous inducer D-fucose (Buttin, 1961). The genes which control specifically the synthesis of these enzymes, i.e. presumably the structural genes, are closely linked, forming a cluster on the *E. coli* chromosome. (Kalckar, Kurahashi & Jordan, 1959; Lederberg, 1960; Yarmolinsky & Wiesmeyer, 1960; Adler, unpublished results.) Certain point-mutations which occur in this chromosome segment abolish the capacity to synthesize both enzymes. These pleiotropic loss mutations are not complemented by any one of the specific (structural) loss mutations, an observation which is in apparent direct conflict with the one-gene one-enzyme hypothesis. These relationships are explained and the conflict is resolved if it is assumed that the linked structural genes constitute an operon controlled by a single operator and that the pleiotropic mutations are mutations of the operator locus.

We have also already discussed the system of simultaneous repression which controls the synthesis of the enzymes involved in histidine synthesis in *Salmonella*. This system involves eight or nine reaction steps. The enzymes which catalyse five of these reactions have been identified. The genes which individually determine these enzymes form a closely linked cluster on the *Salmonella* chromosome. Mutations in each of these genes result in a loss of capacity to synthesize a single enzyme; however, certain mutations at one end of the cluster abolish the capacity to synthesize all the enzymes simultaneously, and these mutations are not complemented by any one of the specific mutations (Ames, Garry & Herzenberg, 1960; Hartman, Loper & Serman, 1960). It will be recalled that the relative rates of synthesis of different enzymes in this sequence are constant under any set of conditions (see p. 327). All these remarkable findings are explained if it is assumed that this cluster of genes constitutes an operon, controlled by an operator associated with the *g* cistron.

The rule that genes controlling metabolically sequential enzymes constitute genetic clusters does not apply, in general, to organisms other than bacteria (Pontecorvo, 1958). Nor does it apply to all bacterial systems, even where simultaneous repression is known to occur and to be controlled by a single regulator gene, as is apparently the case for the enzymes of arginine biosynthesis. In such cases, it must be supposed that several identical or similar operator loci are responsible for sensitivity to repressor of each of the independent information-transfer systems.

It is clear that when an operator controls the expression of only a single structural cistron, the concept of the operon does not apply, and in fact there are no conceivable genetic-biochemical tests which could identify the operator-controlling genetic

segment as distinct from the structural cistron itself.† One may therefore wonder whether it will be possible experimentally to extend this concept to dispersed (as opposed to clustered) genetic systems. It should be remarked at this point that many enzyme proteins are apparently made up of two (or more) different polypeptide chains. It is tempting to predict that such proteins will often be found to be controlled by two (or more) adjacent and co-ordinated structural cistrons, forming an operon.

## 5. The Kinetics of Expression of Structural Genes, and the Nature of the Structural Message

The problem we want to discuss in this section is whether the repressor-operator system functions at the genetic level by governing the *synthesis* of the structural message or at the cytoplasmic level, by controlling the protein-synthesizing *activity* of the messenger (see Fig. 6). These two conceivable models we shall designate respectively as the "genetic operator model" and the "cytoplasmic operator model."

The existence of units of co-ordinate expression involving several structural genes appears in fact difficult to reconcile with the cytoplasmic operator model, if only because of the size that the cytoplasmic unit would have to attain. If we assume that the message is a polyribonucleotide and take a coding ratio of 3, the "unit message" corresponding to an operon governing the synthesis of three proteins of average (monomeric) molecular weight 60,000 would have a molecular weight about $1{\cdot}8 \times 10^6$; we have seen that operons including up to 8 structural cistrons may in fact exist. On the other hand, RNA fractions of *E. coli* and other cells do not appear to include polyribonucleotide molecules of molecular weight exceeding $10^6$.

This difficulty is probably not insuperable; and this type of argument, given the present state of our knowledge, cannot be considered to eliminate the cytoplasmic operator model, even less to establish the validity of the genetic model. However, it seems more profitable tentatively to adopt the genetic model and to see whether some of the more specific predictions which it implies are experimentally verified.

The most immediate and also perhaps the most striking of these implications is that the structural message must be carried by a very short-lived intermediate both rapidly formed and rapidly destroyed during the process of information transfer. This is required by the kinetics of induction. As we have seen, the addition of inducer, or the removal of co-repressor, provokes the synthesis of enzyme at maximum rate within a matter of a few minutes, while the removal of inducer, or the addition of co-repressor interrupts the synthesis within an equally short time. Such kinetics are incompatible with the assumption that the repressor-operator interaction controls the rate of synthesis of *stable* enzyme-forming templates (Monod, 1956, 1958). Therefore, if the genetic operator model is valid, one should expect the kinetics of structural gene expression to be *essentially the same* as the kinetics of induction: injection of a "new" gene into an otherwise competent cell should result in virtually immediate synthesis of the corresponding protein at maximum rate; while removal of the gene should be attended by concomitant cessation of synthesis.

† It should be pointed out that the operational distinction between the operator locus and the structural cistron to which it is directly adjacent rests exclusively on the fact that the operator mutations affect the synthesis of several proteins governed by linked cistrons. This does not exclude the possibility that the operator locus is actually *part* of the structural cistron to which it is "adjacent." If it were so, one might expect certain constitutive operator mutations to involve an alteration of the structure of the protein governed by the "adjacent" cistron. The evidence available at present is insufficient to confirm or eliminate this assumption.

(a) *Kinetics of expression of the galactosidase structural gene*

Additions and removals of genes to and from cells are somewhat more difficult to perform than additions or removals of inducer. However, it can be done. Gene injection without cytoplasmic mixing occurs in the conjugation of *Hfr* male and $F^-$ female *E. coli*. In a mixed male and female population the individual pairs do not all mate at the same time, but the distribution of times of injection of a *given* gene can be rather accurately determined by proper genetic methods. The injection of the $z^+$ (galactosidase) gene from male cells into galactosidase-negative ($z^-$) female cells is rapidly followed by enzyme synthesis within zygotes (cf. p. 332). When the rate of enzyme synthesis in the population is expressed as a function of time, taking into

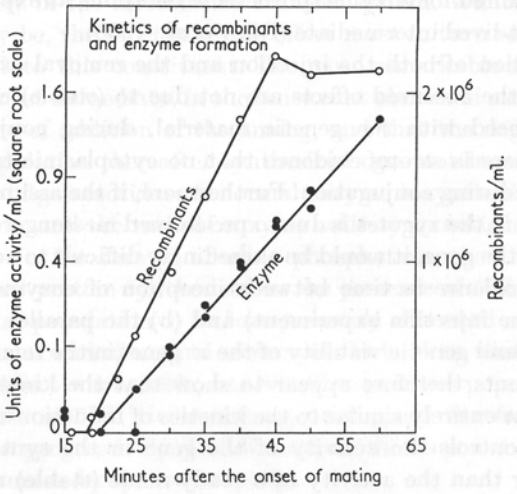



FIG. 7. Kinetics of enzyme production by merozygotes formed by conjugation between inducible galactosidase-positive males and constitutive galactosidase-negative females. Conditions are such that only the zygotes can form enzyme. Increase in the number of $z^+$ containing zygotes is determined by counting recombinants on adequate selective medium. Formation of enzyme is followed by enzyme activity measurements on the total population. It is seen that the enzyme increases linearly with the square of time. Since the zygote population increases linearly with time, it is apparent that the rate of enzyme synthesis per zygote is constant from the time of penetration of the $z^+$ gene. (From Riley *et al.*, 1960.)

account the increase with time of the number of $z^+$ containing zygotes, it is found (see Fig. 7):

(1) that enzyme synthesis begins within two minutes of the penetration of the $z^+$ gene;

(2) that the rate per zygote is constant and maximum over at least the first 40 min following penetration (Riley, Pardee, Jacob & Monod, 1960).

These observations indicate that the structural messenger is very rapidly formed by the $z^+$ gene, and does not accumulate. This could be interpretated in one of two ways:

(a) the structural messenger is a short-lived intermediate;

(b) the structural messenger is stable, but the gene rapidly forms a limited number of messenger molecules, and thereafter stops functioning.

If the second assumption is correct, removal of the gene after the inception of enzyme synthesis should not prevent the synthesis from continuing. This possibility is tested by the "removal" experiment, which is performed by loading the male

chromosome with $^{32}P$ before injection. Following injection (into unlabelled female cells), ample time (25 min) is allowed for expression of the $z^+$ gene, before the zygotes are frozen to allow $^{32}P$ decay for various lengths of time. The rate of galactosidase synthesis by the population is determined immediately after thawing. It is found to decrease sharply as a function of the fraction of $^{32}P$ atoms decayed. If a longer period of time (110 min) is allowed for expression before freezing, no decrease in either enzyme-forming capacity or in viability of the $z^+$ marker are observed. This is to be expected, since by that time most of the $z^+$ genes would have replicated, and this observation provides an internal control showing that no indirect effects of $^{32}P$ disintegrations are involved.

This experiment therefore indicates that even after the $z^+$ gene has become expressed its integrity is required for enzyme synthesis to continue, as expected if the messenger molecule is a short-lived intermediate (Riley *et al.*, 1960).

The interpretation of both the injection and the removal experiment rests on the assumption that the observed effects are not due to (stable) cytoplasmic messenger molecules introduced with the genetic material, during conjugation. As we have already noted, there is strong evidence that no cytoplasmic transfer, even of small molecules, occurs during conjugation. Furthermore, if the assumption were made that enzyme synthesis in the zygotes is due to pre-formed messenger molecules rather than to the activity of the gene, it would be exceedingly difficult to account for both (a) the very precise coincidence in time between inception of enzyme synthesis and entry of the gene (in the injection experiment) and (b) the parallel behaviour of enzyme-forming capacity and genetic viability of the $z^+$ gene (in the removal experiment).

These experiments therefore appear to show that the kinetics of expression of a structural gene are entirely similar to the kinetics of induction-repression, as expected if the operator controls the activity of the gene in the synthesis of a short-lived messenger, rather than the activity of a ready-made (stable) messenger molecule in synthesizing protein.

It is interesting at this point to recall the fact that infection of *E. coli* with virulent ($\phi$II, T2, T4) phage is attended within 2 to 4 minutes by inhibition of *bacterial* protein synthesis, including in particular $\beta$-galactosidase (Cohen, 1949; Monod & Wollman, 1947; Benzer, 1953). It is known on the other hand that phage-infection results in rapid visible lysis of bacterial nuclei, while no major destruction of pre-formed bacterial RNA appears to occur (Luria & Human, 1950). It seems very probable that the inhibition of specific bacterial protein synthesis by virulent phage is due essentially to the depolymerization of bacterial DNA, and this conclusion also implies that the integrity of bacterial genes is required for continued synthesis of bacterial protein. In confirmation of this interpretation, it may be noted that infection of *E. coli* by phage $\lambda$, which does not result in destruction of bacterial nuclei, allows $\beta$-galactosidase synthesis to continue almost to the time of lysis (Siminovitch & Jacob, 1952).

### (b) *Structural effects of base analogs*

An entirely different type of experiment also leads to the conclusion that the structural messenger is a short-lived intermediate and suggests, furthermore, that this intermediate is a ribonucleotide. It is known that certain purine and pyrimidine analogs are incorporated by bacterial cells into ribo- and deoxyribonucleotides, and it has been found that the synthesis of protein, or of some proteins, may be inhibited in the presence of certain of these analogs. One of the mechanisms by which these effects

could be explained may be that certain analogs are incorporated into the structural messenger. If so, one might hope to observe that the molecular structure of specific proteins formed in the presence of an analog is modified. It has in fact been found that the molecular properties of β-galactosidase and of alkaline phosphatase synthesized by *E. coli* in the presence of 5-fluorouracil (5FU) are strikingly altered. In the case of β-galactosidase, the ratio of enzyme activity to antigenic valency is decreased by 80%. In the case of alkaline phosphatase, the rate of thermal inactivation (of this normally highly heat-resistant protein) is greatly increased (Naono & Gros, 1960*a*,*b*; Bussard, Naono, Gros & Monod, 1960).

It can safely be assumed that such an effect cannot result from the mere presence of 5FU in the cells, and must reflect incorporation of the analog into a constituent involved in some way in the information transfer system. Whatever the identity of this constituent may be, the kinetics of the effect must in turn reflect the kinetics of 5FU incorporation into this constituent. The most remarkable feature of the 5FU effect is that it is almost immediate, in the sense that abnormal enzyme is synthesized almost from the time of addition of the analog, and that the degree of abnormality of the molecular population thereafter synthesized does not increase with time. For instance, in the case of galactosidase abnormal enzyme is synthesized within 5 min of addition of the analog, and the ratio of enzyme activity to antigenic valency remains constant thereafter. In the case of alkaline phosphatase, the thermal inactivation curve of the abnormal protein synthesized in the presence of 5FU is monomolecular, showing the molecular population to be *homogeneously* abnormal rather than made up of a mixture of normal and abnormal molecules. It is clear that if the constituent responsible for this effect were stable, one would expect the population of molecules made in the presence of 5FU to be heterogeneous, and the fraction of abnorma molecules to increase progressively. It follows that the responsible constituent must be formed, and also must decay, very rapidly.

Now it should be noted that, besides the structural gene-synthesized messenger, the information transfer system probably involves other constituents responsible for the correct translation of the message, such as for instance the RNA fractions involved in amino acid transfer. The 5FU effect could be due to incorporation into one of these fractions rather than to incorporation into the messenger itself. However, the convergence of the results of the different experiments discussed above strongly suggests that the 5FU effect does reflect a high rate of turnover of the messenger itself.

### (c) *Messenger RNA*

Accepting tentatively these conclusions, let us then consider what properties would be required of a cellular constituent, to allow its identification with the structural messenger. These qualifications based on general assumptions, and on the results discussed above, would be as follows:

(1) The "candidate" should be a polynucleotide.

(2) The fraction would presumably be very heterogeneous with respect to molecular weight. However, assuming a coding ratio of 3, the average molecular weight would not be lower than $5 \times 10^5$.

(3) It should have a base composition reflecting the base composition of DNA.

(4) It should, at least temporarily or under certain conditions, be found associated with ribosomes, since there are good reasons to believe that ribosomes are the seat of protein synthesis.

(5) It should have a very high rate of turnover and in particular it should saturate with 5FU within less than about 3 min.

It is immediately evident that none of the more classically recognized cellular RNA fractions meets these very restrictive qualifications. Ribosomal RNA, frequently assumed to represent the "template" in protein synthesis, is remarkably homogeneous in molecular weight. Its base composition is similar in different species, and does not reflect the variations in base ratios found in DNA. Moreover it appears to be entirely stable in growing cells (Davern & Meselson, 1960). It incorporates 5FU only in proportion to net increase.

Transfer RNA, or (sRNA) does not reflect DNA in base composition. Its average molecular weight is much lower than the $5 \times 10^5$ required for the messenger. Except perhaps for the terminal adenine and cytidine, its rate of incorporation of bases, including in particular 5FU, is not higher than that of ribosomal RNA.

However, a small fraction of RNA, first observed by Volkin & Astrachan (1957) in phage infected *E. coli*, and recently found to exist also in normal yeasts (Yčas & Vincent, 1960) and coli (Gros, *et al.*, 1961), does seem to meet all the qualifications listed above.

This fraction (which we shall designate "messenger RNA" or M-RNA) amounts to only about 3% of the total RNA; it can be separated from other RNA fractions by column fractionation or sedimentation (Fig. 8). Its average sedimentation velocity coefficient is 13, corresponding to a minimum molecular weight of $3 \times 10^5$, but since the molecules are presumably far from spherical, the molecular weight is probably much higher. The rate of incorporation of $^{32}P$, uracil or 5FU into this fraction is extremely rapid: half saturation is observed in less than 30 sec, indicating a rate of synthesis several hundred times faster than any other RNA fraction. Its half life is also very short, as shown by the disappearance of radioactivity from this fraction in pre-labelled cells. At high concentrations of $Mg^{2+}$ (0·005 M) the fraction tends to associate with the 70s ribosomal particles, while at lower $Mg^{2+}$ concentrations it sediments independently of the ribosomal particles (Gros *et al.*, 1961).

The striking fact, discovered by Volkin & Astrachan, that the base-composition of this fraction in T2-infected cells reflects the base composition of *phage* (rather than bacterial) DNA, had led to the suggestion that it served as a precursor of phage DNA The agreement between the properties of this fraction and the properties of a short-lived structural messenger suggests that, in phage infected cells as well as in normal cells, this fraction served in fact in the transfer of genetic information from phage DNA to the protein synthesizing centers. This assumption implies that the same protein-forming centers which, in uninfected cells, synthesize bacterial protein, also serve in infected cells to synthesize phage protein according to the new structural information provided by phage DNA, *via* M-RNA. This interpretation is strongly supported by recent observations made with T4 infected *E. coli*. (Brenner, Jacob & Meselson, 1961).

Uninfected cells of *E. coli* were grown in the presence of $^{15}N$. They were then infected and resuspended in $^{14}N$ medium. Following infection, they were exposed to short pulses of $^{32}P$ or $^{35}S$, and the ribosomes were analysed in density gradients. It was found:

(1) that no detectable amounts of ribosomal RNA were synthesized after infection;

(2) labelled M-RNA formed *after* infection became associated with unlabelled ribosomal particles formed *before* infection;

(3) newly formed (i.e. phage-determined) protein, identified by its $^{35}S$ content, was found associated with the 70s particles before it appeared in the soluble protein fraction.

These observations strongly suggest that phage protein is synthesized by *bacterial* ribosomes formed before infection and associated with *phage-determined* M-RNA. Since the structural information for phage protein could not reside in the bacterial ribosomes, it must be provided by the M-RNA fraction.

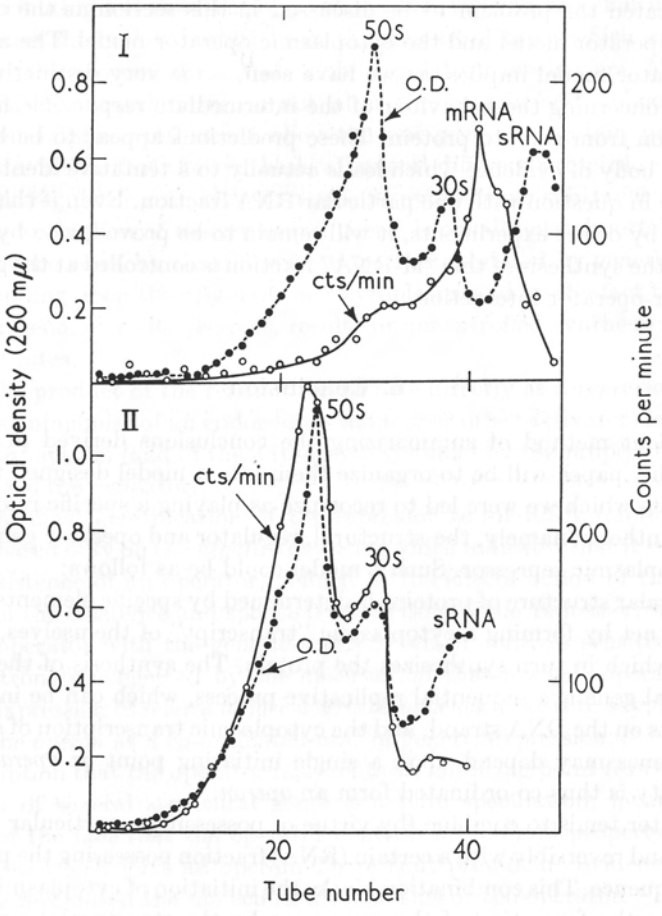



FIG. 8. Incorporation and turnover of uracil in messenger RNA. *E. coli* growing exponentially in broth were incubated for 5 sec with [$^{14}$C]-uracil. The bacteria were centrifuged, washed and resuspended in the original volume of the same medium containing 100-fold excess of [$^{12}$C]-uracil. Half the bacteria were then harvested and frozen (I) and the remainder were incubated for 15 min at 37°C (II) prior to harvesting and freezing. The frozen samples were ground with alumina and extracted with tris buffer (2-amino-2 hydroxymethylpropane-1:3-diol) containing $10^{-4}$M-Mg, treated with DNase and applied to a sucrose gradient. After 3 hr, sequential samples were taken for determination of radioactivity and absorption at 260 mμ. It may be seen (part I) that after 5 sec, M-RNA is the only labelled fraction, and that subsequently (part II) uracil incorporated into M-RNA is entirely renewed. (From Gros *et al.*, 1961.)

Finally, the recent experiments of Lamfrom (1961) independently repeated by Kruh, Rosa, Dreyfus & Schapira (1961) have shown directly that species specificity in the synthesis of haemoglobin is determined by a "soluble" RNA-containing fraction rather than by the ribosomal fraction. Lamfrom used reconstructed systems, containing ribosomes from one species (rabbit) and soluble fractions from another (sheep) and found that the haemoglobin formed *in vitro* by these systems belonged in part to the

type characteristic of the species used to prepare the *soluble* fraction. It is not, of course, positively proved that *inter-specific* differences in haemoglobin structure are gene-determined rather than cytoplasmic, but the assumption seems safe enough. In any case, Lamfrom's experiment proves beyond doubt that the ribosomes cannot be considered to determine entirely (if at all) the specific structure of proteins.

We had stated the problem to be discussed in this section as the choice between the genetic operator model and the cytoplasmic operator model. The adoption of the genetic operator model implies, as we have seen, some very distinctive and specific predictions concerning the behaviour of the intermediate responsible for the transfer of information from gene to protein. These predictions appear to be borne out by a considerable body of evidence which leads actually to a tentative identification of the intermediate in question with one particular RNA fraction. Even if this identification is confirmed by direct experiments, it will remain to be proved, also by direct experiments, that the synthesis of this "M-RNA" fraction is controlled at the genetic level by the repressor-operator interaction.

## 6. Conclusion

A convenient method of summarizing the conclusions derived in the preceding sections of this paper will be to organize them into a model designed to embody the main elements which we were led to recognize as playing a specific role in the control of protein synthesis; namely, the structural, regulator and operator genes, the operon, and the cytoplasmic repressor. Such a model could be as follows:

The molecular structure of proteins is determined by specific elements, the *structural genes*. These act by forming a cytoplasmic "transcript" of themselves, the structural messenger, which in turn synthesizes the protein. The synthesis of the messenger by the structural gene is a sequential replicative process, which can be initiated only at certain points on the DNA strand, and the cytoplasmic transcription of several, linked, structural genes may depend upon a single initiating point or *operator*. The genes whose activity is thus co-ordinated form an *operon*.

The operator tends to combine (by virtue of possessing a particular base sequence) specifically and reversibly with a certain (RNA) fraction possessing the proper (complementary) sequence. This combination blocks the initiation of cytoplasmic transcription and therefore the formation of the messenger by the structural genes in the whole operon. The specific "repressor" (RNA?), acting with a given operator, is synthesized by a *regulator gene*.

The repressor in certain systems (inducible enzyme systems) tends to combine specifically with certain specific small molecules. The combined repressor has no affinity for the operator, and the combination therefore results in *activation of the operon*.

In other systems (repressible enzyme systems) the repressor by itself is inactive (i.e. it has no affinity for the operator) and is activated only by combining with certain specific small molecules. The combination therefore leads to *inhibition of the operon*.

The structural messenger is an unstable molecule, which is destroyed in the process of information transfer. The rate of messenger synthesis, therefore, in turn controls the rate of protein synthesis.

This model was meant to summarize and express conveniently the properties of the different factors which play a specific role in the control of protein synthesis. In

order concretely to represent the functions of these different factors, we have had to introduce some purely speculative assumptions. Let us clearly discriminate the experimentally established conclusions from the speculations:

(1) The most firmly grounded of these conclusions is the existence of *regulator* genes, which control the rate of information-transfer from *structural* genes to proteins, without contributing any information to the proteins themselves. Let us briefly recall the evidence on this point: mutations in the structural gene, which are reflected as alterations of the protein, do not alter the regulatory mechanism. Mutations that alter the regulatory mechanism do not alter the protein and do not map in the structural genes. Structural genes obey the one-gene one-protein principle, while regulator genes may affect the synthesis of several different proteins.

(2) That the regulator gene acts *via* a specific cytoplasmic substance whose effect is to *inhibit* the expression of the structural genes, is equally clearly established by the *trans* effect of the gene, by the different properties exhibited by genetically identical zygotes depending upon the origin of their cytoplasm, and by the fact that absence of the regulator gene, or of its product, results in uncontrolled synthesis of the protein at maximum rates.

(3) That the product of the regulator gene acts directly as a *repressor* (rather than indirectly, as antagonist of an endogenous inducer or other activator) is proved in the case of the *Lac* system (and of the $\lambda$ lysogenic systems) by the properties of the dominant mutants of the regulator.

(4) The chemical identification of the repressor as an RNA fraction is a logical assumption based only on the *negative* evidence which indicates that it is not a protein.

(5) The existence of an operator, defined as the site of action of the repressor, is deduced from the existence and specificity of action of the repressor. The identification of the operator with the genetic segment which controls sensitivity to the repressor, is strongly suggested by the observation that a *single* operator gene may control the expression of *several adjacent structural genes*, that is to say, by the demonstration of the *operon* as a co-ordinated unit of genetic expression.

The assumption that the operator represents an initiating point for the cytoplasmic transcription of several structural genes is a pure speculation, meant only as an illustration of the fact that the operator controls an integral property of the group of linked genes which form an operon. There is at present no evidence on which to base any assumption on the molecular mechanisms of the operator.

(6) The assumptions made regarding the interaction of the repressor with inducers or co-repressors are among the weakest and vaguest in the model. The idea that specific coupling of inducers to the repressor could result in inactivation of the repressor appears reasonable enough, but it raises a difficulty which we have already pointed out. Since this reaction between repressor and inducer must be stereospecific (for both) it should presumably require a specific enzyme; yet no evidence, genetic or biochemical, has been found for such an enzyme.

(7) The property attributed to the structural messenger of being an unstable intermediate is one of the most specific and novel implications of this scheme; it is required, let us recall, by the kinetics of induction, once the assumption is made that the control systems operate at the genetic level. This leads to a new concept of the mechanism of information transfer, where the protein synthesizing centers (ribosomes) play the role of non-specific constituents which can synthesize different proteins, according to specific instructions which they receive from the genes through M-RNA. The already fairly impressive body of evidence, kinetic and analytical, which supports

this new interpretation of information transfer, is of great interest in itself, even if some of the other assumptions included in the scheme turn out to be incorrect.

These conclusions apply strictly to the bacterial systems from which they were derived; but the fact that adaptive enzyme systems of both types (inducible and repressible) and phage systems appear to obey the same fundamental mechanisms of control, involving the same essential elements, argues strongly for the generality of what may be called "repressive genetic regulation" of protein synthesis.

One is led to wonder whether all or most structural genes (i.e. the synthesis of most proteins) are submitted to repressive regulation. In bacteria, virtually all the enzyme systems which have been adequately studied have proved sensitive to inductive or repressive effects. The old idea that such effects are characteristic only of "non-essential" enzymes is certainly incorrect (although, of course, these effects can be detected only under conditions, natural or artificial, such that the system under study is at least partially non-essential (gratuitous). The results of mutations which abolish the control (such as constitutive mutations) illustrate its physiological importance. Constitutive mutants of the lactose system synthesize 6 to 7% of all their proteins as $\beta$-galactosidase. In constitutive mutants of the phosphatase system, 5 to 6% of the total protein is phosphatase. Similar figures have been obtained with other constitutive mutants. It is clear that the cells could not survive the breakdown of more than two or three of the control systems which keep in pace the synthesis of enzyme proteins.

The occurrence of inductive and repressive effects in tissues of higher organisms has been observed in many instances, although it has not proved possible so far to analyse any of these systems in detail (the main difficulty being the creation of controlled conditions of gratuity). It has repeatedly been pointed out that enzymatic adaptation, as studied in micro-organisms, offers a valuable model for the interpretation of biochemical co-ordination within tissues and between organs in higher organisms. The demonstration that adaptive effects in micro-organisms are primarily negative (repressive), that they are controlled by functionally specialized genes and operate at the genetic level, would seem greatly to widen the possibilities of interpretation. The fundamental problem of chemical physiology and of embryology is to understand why tissue cells do not all express, all the time, all the potentialities inherent in their genome. The survival of the organism requires that many, and, in some tissues most, of these potentialities be unexpressed, that is to say *repressed*. Malignancy is adequately described as a breakdown of one or several growth controlling systems, and the genetic origin of this breakdown can hardly be doubted.

According to the strictly structural concept, the genome is considered as a mosaic of independent molecular blue-prints for the building of individual cellular constituents. In the execution of these plans, however, co-ordination is evidently of absolute survival value. The discovery of regulator and operator genes, and of repressive regulation of the activity of structural genes, reveals that the genome contains not only a series of blue-prints, but a co-ordinated program of protein synthesis and the means of controlling its execution.

## REFERENCES

Adelberg, E. A. & Umbarger, H. E. (1953). *J. Biol. Chem.* **205**, 475.
Ames, B. N. & Garry, B. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1453.
Ames, B. N., Garry, B. & Herzenberg, L. A. (1960). *J. Gen. Microbiol.* **22**, 369.
Benzer, S. (1953). *Biochim. biophys. Acta*, **11**, 383.
Bertani, G. (1953). *Cold. Spr. Harb. Symp. Quant. Biol.* **18**, 65.

Bertani, G. (1958). *Advanc. Virus Res.* **5**, 151.
Brenner, S., Jacob, F. & Meselson, M. (1961). *Nature*, **190**, 576.
Bussard, A., Naono, S., Gros, F. & Monod, J. (1960). *C. R. Acad. Sci., Paris*, **250**, 4049.
Buttin, G. (1956). Diplôme Et. Sup., Paris.
Buttin, G. (1961). *C. R. Acad. Sci., Paris*, in the press.
Cohen, G. N. & Jacob, F. (1959). *C. R. Acad. Sci., Paris*, **248**, 3490.
Cohen, G. N. & Monod, J. (1957). *Bact. Rev.* **21**, 169.
Cohen, S. S. (1949). *Bact. Rev.* **13**, 1.
Cohen-Bazire, G. & Jolit, M. (1953). *Ann. Inst. Pasteur*, **84**, 1.
Cohn, M. (1957). *Bact. Rev.* **21**, 140.
Cohn, M., Cohen, G. N. & Monod, J. (1953). *C. R. Acad. Sci., Paris*, **236**, 746.
Cohn, M. & Horibata, K. (1959). *J. Bact.* **78**, 624.
Cohn, M. & Monod, J. (1953). In *Adaptation in Micro-organisms*, p. 132. Cambridge University Press.
Cohn, M. & Torriani, A. M. (1952). *J. Immunol.* **69**, 471.
Davern, C. I. & Meselson, M. (1960). *J. Mol. Biol.* **2**, 153.
Demerec, M. (1956). *Cold Spr. Harb. Symp. Quant. Biol.* **21**, 113.
Dienert, F. (1900). *Ann. Inst. Pasteur*, **14**, 139.
Duclaux, E. (1899). *Traité de Microbiologie*. Paris: Masson et Cie.
Echols, H., Garen, A., Garen, S. & Torriani, A. M. (1961). *J. Mol. Biol.*, in the press.
Flaks, J. G. & Cohen, S. S. (1959). *J. Biol. Chem.* **234**, 1501.
Gale, E. F. (1943). *Bact. Rev.* **7**, 139.
Giles, N. H. (1958). *Proc. Xth Intern. Cong. Genetics*, Montreal, **1**, 261.
Gorini, L. & Maas, W. K. (1957). *Biochim. biophys. Acta*, **25**, 208.
Gorini, L. & Maas, W. K. (1958). In *The Chemical Basis of Development*, p. 469. Baltimore. Johns Hopkins Press.
Gros, F., Hiatt, H., Gilbert, W., Kurland, C. G., Risebrough, R. W. & Watson, J. D. (1961). *Nature*, **190**, 581.
Halvorson, H. O. (1960). *Advanc. Enzymol.* in the press.
Hartman, P. E., Loper, J. C. & Serman, D. (1960). *J. Gen. Microbiol.* **22**, 323.
Herzenberg, L. (1959). *Biochim. biophys. Acta*, **31**, 525.
Hogness, D. S., Cohn, M. & Monod, J. (1955). *Biochim. biophys. Acta*, **16**, 99.
Jacob, F. (1954). *Les Bactéries Lysogènes et la Notion de Provirus*. Paris: Masson et Cie.
Jacob, F. (1960). *Harvey Lectures*, 1958–1959, series **54**, 1.
Jacob, F. & Adelberg, E. A. (1959). *C.R. Acad. Sci., Paris*, **249**, 189.
Jacob, F. & Campbell, A. (1959). *C.R. Acad. Sci., Paris*, **248**, 3219.
Jacob, F., Fuerst, C. R. & Wollman, E. L. (1957). *Ann. Inst. Pasteur*, **93**, 724.
Jacob, F. & Monod, J. (1959). *C.R. Acad. Sci., Paris*, **249**, 1282.
Jacob, F., Perrin, D., Sanchez, C. & Monod, J. (1960*a*). *C.R. Acad. Sci., Paris*, **250**, 1727.
Jacob, F., Schaeffer, P. & Wollman, E. L. (1960*b*). In *Microbial Genetics*, Xth Symposium of the Society for General Microbiology, p. 67.
Jacob, F. & Wollman, E. L. (1953). *Cold Spr. Harb. Symp. Quant. Biol.* **18**, 101.
Jacob, F. & Wollman, E. L. (1956). *Ann. Inst. Pasteur*, **91**, 486.
Jacob, F. & Wollman, E. L. (1957). In *The Chemical Basis of Heredity*, p. 468. Baltimore: Johns Hopkins Press.
Kaiser, A. D. (1957). *Virology*, **3**, 42.
Kaiser, A. D. & Jacob, F. (1957). *Virology*, **4**, 509.
Kalckar, H. M., Kurahashi, K. & Jordan, E. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1776.
Karstrom, H. (1938). *Ergebn. Enzymforsch.* **7**, 350.
Képès, A. (1960). *Biochim. biophys. Acta*, **40**, 70.
Képès, A., Monod, J. & Jacob, F. (1961). In preparation.
Kogut, M., Pollock, M. & Tridgell, E. J. (1956). *Biochem. J.* **62**, 391.
Kornberg, A., Zimmerman, S. B., Kornberg, S. R. & Josse, J. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 772.
Kruh, J., Rosa, J., Dreyfus, J.-C. & Schapira, G. (1961). *Biochim. biophys. Acta*, in the press.
Lamfrom, H. (1961). *J. Mol. Biol.* **3**, 241.
Lederberg, E. (1960). In *Microbial Genetics*, The Xth Symposium of the Society of General Microbiology, p. 115.

Levinthal, C. (1959). In *Structure and Function of Genetic Elements*, Brookhaven Symposia in Biology, p. 76.
Luria, S. E. & Human, M. L. (1950). *J. Bact.* **59**, 551.
Lwoff, A. (1953). *Bact. Rev.* **17**, 269.
Lwoff, A., Siminovitch, L. & Kjeldgaard, N. (1950). *Ann. Inst. Pasteur*, **79**, 815.
Magasanik, B., Magasanik, A. K. & Neidhardt, F. C. (1959). In *A Ciba Symposium on the Regulation of Cell Metabolism*, p. 334. London: Churchill.
Monod, J. (1942). *Recherches sur la Croissance des Cultures Bactériennes*. Paris: Hermann.
Monod, J. (1955). *Exp. Ann. Biochim. Méd.* série XVII, p. 195. Paris: Masson et Cie.
Monod, J. (1956). In *Units of Biological Structure and Function*, p. 7. New York: Academic Press.
Monod, J. (1958). *Rec. Trav. Chim. des Pays-Bas*, **77**, 569.
Monod, J. (1959). *Angew. Chem.* **71**, 685.
Monod, J. & Audureau, A. (1946). *Ann. Inst. Pasteur*, **72**, 868.
Monod, J. & Cohen-Bazire, G. (1953). *C.R. Acad. Sci., Paris*, **236**, 530.
Monod, J. & Cohn, M. (1952). *Advanc. Enzymol.* **13**, 67.
Monod, J. & Cohn, M. (1953). In *Symposium on Microbial Metabolism*. VIth Intern. Cong. of Microbiol., Rome, p. 42.
Monod, J., Pappenheimer, A. M. & Cohen-Bazire, G. (1952), *Biochim. biophys. Acta*, **9**, 648.
Monod, J. & Wollman, E. L. (1947). *Ann. Inst. Pasteur*, **73**, 937.
Naono, S. & Gros, F. (1960*a*). *C.R. Acad. Sci., Paris*, **250**, 3527.
Naono, S. & Gros, F. (1960*b*). *C.R. Acad. Sci., Paris*, **250**, 3889.
Neidhardt, F. C. & Magasanik, B. (1956*a*). *Nature*, **178**, 801.
Neidhardt, F. C. & Magasanik, B. (1956*b*). *Biochim. biophys. Acta*, **21**, 324.
Novick, A. & Szilard, L., in Novick, A. (1955). *Ann. Rev. Microbiol.* **9**, 97.
Pardee, A. B. (1957). *J. Bact.* **73**, 376.
Pardee, A. B., Jacob, F. & Monod, J. (1959). *J. Mol. Biol.* **1**, 165.
Pardee, A. B. & Prestidge, L. S. (1959). *Biochim. biophys. Acta*, **36**, 545.
Pardee, A. B. & Prestidge, L. S. (1961). In preparation.
Perrin, D., Bussard, A. & Monod, J. (1959). *C.R. Acad. Sci., Paris*, **249**, 778.
Perrin, D., Jacob, F. & Monod, J. (1960). *C.R. Acad. Sci., Paris*, **250**, 155.
Pollock, M. (1950). *Brit. J. Exp. Pathol.* **4**, 739.
Pollock, M. & Perret, J. C. (1951). *Brit. J. Exp. Pathol.* **5**, 387.
Pontecorvo, G. (1958). *Trends in Genetic Analysis*. New York: Columbia University Press.
Rickenberg, H. V., Cohen, G. N., Buttin, G. & Monod, J. (1956). *Ann. Inst. Pasteur*, **91**, 829.
Riley, M., Pardee, A. B., Jacob, F. & Monod, J. (1960). *J. Mol. Biol.* **2**, 216.
Rotman, B. & Spiegelman, S. (1954). *J. Bact.* **68**, 419.
Siminovitch, L. & Jacob, F. (1952). *Ann. Inst. Pasteur*, **83**, 745.
Stanier, R. Y. (1951). *Ann. Rev. Microbiol.* **5**, 35.
Szilard, L. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 277.
Torriani, A. M. (1960). *Biochim. biophys. Acta*, **38**, 460.
Umbarger, H. E. (1956). *Science*, **123**, 848.
Vogel, H. J. (1957*a*). *Proc. Nat. Acad. Sci., Wash.* **43**, 491.
Vogel, H. J. (1957*b*). In *The Chemical Basis of Heredity*, p. 276. Baltimore: Johns Hopkins Press.
Volkin, E. & Astrachan, L. (1957). In *The Chemical Basis of Heredity*, p. 686. Baltimore: Johns Hopkins Press.
Went, F. C. (1901). *J. Wiss. Bot.* **36**, 611.
Wijesundera, S. & Woods, D. D. (1953). *Biochem. J.* **55**, viii.
Willson, C., Perrin, D., Jacob, F. & Monod, J. (1961). In preparation.
Wollman, E. L. & Jacob, F. (1959). *La Sexualité des Bactéries*. Paris: Masson et Cie.
Yanofsky, C. (1960). *Bact. Rev.* **24**, 221.
Yanofsky, C. & Lennox, E. S. (1959). *Virology*, **8**, 425.
Yarmolinsky, M. B. & Wiesmeyer, H. (1960). *Proc. Nat. Acad. Sci., Wash.* in the press.
Yates, R. A. & Pardee, A. B. (1956). *J. Biol. Chem.* **221**, 757.
Yates, R. A. & Pardee, A. B. (1957). *J. Biol. Chem.* **227**, 677.
Yčas, M. & Vincent, W. S. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 804.
Zabin, I., Képès, A. & Monod, J. (1959). *Biochem. Biophys. Res. Comm.* **1**, 289.

PRINTED IN GREAT BRITAIN AT THE UNIVERSITY PRESS ABERDEEN