

Modélisation dynamique de systèmes génétiques de régulation

I. L'induction de l'opéron lactose d'*Escherichia coli* : élaboration d'un modèle

C. CHEVALET, Florence CORPET, M. GILLOIS et A. MICALI *

*I.N.R.A., Laboratoire de Génétique cellulaire
Centre de Recherches de Toulouse, B.P. 12, F 31320 Castanet-Tolosan*

** Institut de Mathématiques, Université de Montpellier II
Place Eugène-Bataillon, F 34060 Montpellier*

Résumé

Dans ce premier article, on élabore un modèle mathématique pour décrire le mécanisme de l'induction de l'opéron lactose d'*Escherichia coli*. Les différentes interactions moléculaires sont successivement représentées par des équations différentielles qui traduisent l'évolution dans le temps : des probabilités des états possibles de la région de contrôle de l'opéron (opérateur et promoteur), de concentrations des enzymes codées par le système (perméase et β -galactosidase), et des concentrations des substrats et produits de ces enzymes (lactose, inducteur, glucose et galactose intracellulaires). L'ensemble constitue un système de dix équations différentielles du premier ordre, présentant des non-linéarités et des arguments retardés. La cohérence de l'ensemble se justifie par des considérations sur les hiérarchies observées entre les nombres de molécules des diverses espèces moléculaires, de un gène à plusieurs millions de molécules de sucre, et entre les vitesses absolues des réactions d'association et de dissociation entre molécules. Les valeurs numériques des paramètres du modèle sont évaluées pour une bactérie de type sauvage. La plupart des paramètres peuvent être obtenus d'après des expériences, réalisées *in vitro*, ou *in vivo* sur des parties du système, mais certains paramètres inconnus doivent être estimés en utilisant le modèle. Dans ce modèle, les paramètres ont une interprétation biologique immédiate, mais l'imprécision de la détermination numérique de certains, et la complexité du modèle, empêchent de pouvoir énoncer des propriétés qualitatives générales.

L'analyse de ces propriétés structurelles, de leur dépendance par rapport aux valeurs des paramètres, fait l'objet de l'article suivant, où l'on montre, notamment par des méthodes numériques, comment on peut analyser le comportement de souches mutantes, prévoir celui de génotypes inconnus, et proposer, à l'aide du modèle, de nouvelles expériences.

Mots-clés : Modèle mathématique, opéron lactose, Escherichia coli, métabolisme du lactose, transcription.

Summary

Dynamical modeling of genetic systems of regulation

1. The induction of Escherichia coli lactose operon : statement of a model

This series of papers is devoted to the building of dynamical models describing the expression of polygenic traits, when the molecular mechanisms are known from earlier biochemical and genetical analysis. The aims of these models are to provide a synthetic account of a huge body of analytical works, to make possible the simulation of further experiments, and to develop methods appropriate to the description of complex characters.

This first paper is concerned with the elaboration of a dynamical system modeling the induction of the lactose operon of *Escherichia coli*. The different molecular mechanisms involved in the functioning of the operon are reviewed and are given a mathematical translation that takes account of the main biochemical and genetical facts, and of the achievements of previous theoretical works. The subsystems modeled is this way are the interactions between regulatory molecules (repressor, catabolite activator protein, inducers and anti-inducers, and cyclic AMP) and the control region of the operon (operator and promoter), and the biological activities of the proteins encoded by the structural genes (permease and β -galactosidase). The biosynthesis of these proteins, which is not under the control of this operon, is modeled in a non-specific way, although it is taken advantage of the known short life-time of the messenger RNAs. Dynamical aspects of the catabolic repression by extracellular glucose are not included in the present analysis, since there are no available data that would allow to derive kinetic equations for this phenomenon.

Deriving a single mathematical model that represents the whole biological system, from the juxtaposition of the preceding sub-models, is justified by some hierarchical properties. They involve the numbers of molecules in a cell (one gene, a few regulatory molecules, thousands of enzymes, millions to billions of substrates and products), and the absolute velocities of different reactions. The full model is a system of ten first-order differential equations. Five independent equations describe the states of the control region, they are linear but the coefficients depend upon the level of intracellular inducer. Two linear equations yield the rates of synthesis of the proteins, they involve delayed arguments since the arising of new active enzymes depends on the initiating of transcription a few minutes before. Two non-linear equations give the rates of change in the concentrations of intracellular lactose and inducer, they are based on the kinetic equations of the enzymes. Finally, one equation gives the rate at which glucose and galactose are produced, it is the output of the system.

The parameters introduced in the model are estimated. Most of them can be derived from *in vitro* measurements (kinetic coefficients of association and dissociation between operator, repressor, and inducer; kinetic parameters of the enzymes; half life of enzymes), or from *in vivo* experiments dealing with sub-systems only (delays between transcription initiation and arising of enzymatic activity; kinetic coefficients of permeation; amplification coefficient of proteins biosynthesis). Other ones, that involve the interaction of the catabolite activator protein with the promoter, must be estimated because this interaction has not yet been quantitatively studied.

The model has the advantage that its parameters have a clear meaning, each one may be related either to some biophysical property of a gene or of its product, or to some environmental characteristics. However, as it is non linear, its general properties are expected to depend on the numerical values of the parameters, which are not always known with good accuracy. The mathematical analysis of the structural properties of the model, and of their possible dependence on the parameters, will be shown in the next paper. Also, the comparative study of the wild type operon and of various mutant strains, by numerical methods, will show how the model may be used to simulate new genotypes, to predict some responses, or to estimate some parameters that are difficult to get from a direct experiment.

Key-words : Mathematical models, lactose operon, Escherichia coli, lactose metabolism, transcription.

Orientation générale

L'essor de la génétique et de la biologie moléculaires au cours des dernières décennies, a mis en lumière les mécanismes principaux du fonctionnement des gènes, de leurs modes d'expression, et des systèmes de régulation qu'ils contrôlent ou dont ils dépendent. Dans l'univers bactérien surtout, les régulations génétiques de nombreuses voies de synthèse ou de dégradation ont été identifiées, et caractérisées par des travaux analytiques de génétique et de biochimie. On dispose, pour ces systèmes, d'une description verbale cohérente des mécanismes mis en jeu, et de concepts généraux qui permettent d'imaginer le fonctionnement d'un nouveau système et de concevoir les expériences qui conduiront à son identification. Parallèlement, se développent les méthodes physico-chimiques de mesure des interactions entre les molécules participant à ces mécanismes de régulation. Il devient ainsi concevable de tenter une synthèse quantitative des connaissances analytiques accumulées sur ces systèmes bactériens.

Vis-à-vis d'un système donné, une telle démarche peut être l'occasion de faire une mise au point, de discerner, dans la masse des résultats partiels, quelles sont les caractéristiques essentielles, et de déceler, le cas échéant, des incohérences ou des incompatibilités entre certaines mesures. La construction d'un modèle quantitatif peut aussi être un outil pour séparer plusieurs hypothèses, ou tout au moins pour aider à la conception d'expériences qui permettent de trancher ; un modèle peut enfin être utilisé pour mesurer des paramètres qui ne sont pas directement accessibles à l'expérience. Cependant, et indépendamment des systèmes que l'on peut ainsi étudier, cette démarche présente un intérêt plus général, car elle a pour objet le fonctionnement conjoint d'un ensemble de gènes. La perspective est une meilleure compréhension de l'hérédité des caractères polygéniques, qui nécessite l'élaboration de méthodes de description quantitative et qualitative des systèmes génétiquement contrôlés. Une telle méthodologie doit satisfaire à plusieurs conditions :

(i) être un résumé assez fidèle du fonctionnement du système dans les différentes conditions du milieu, ce qui suppose un accord qualitatif et une conformité quantitative du modèle aux résultats expérimentaux ;

(ii) être capable de traduire une variabilité génétique, par un choix des paramètres assurant une relation claire entre un gène et ses effets ;

(iii) être formulée en termes dynamiques, car les effets d'un ensemble de gènes sont exprimés par des associations moléculaires, des cinétiques enzymatiques, des taux d'incorporation, ou de dégradation. La notion d'effet quantitatif ne peut se traduire par un nombre que si un certain équilibre dynamique s'établit ;

(iv) avoir une structure mathématique assez simple pour permettre la caractérisation des propriétés qualitatives du modèle, c'est-à-dire pour exprimer en termes généraux la structure du système.

Plusieurs types de modèles ont été proposés pour répondre à ces exigences. Ils se rattachent à deux types de méthodes, algébriques et analytiques. Les premiers cherchent à décrire les liaisons structurelles entre les divers éléments d'un système métabolique soumis à des régulations (gènes, enzymes, substrats des enzymes), dans le langage de la théorie des catégories (ROSEN, 1972), ou par une formalisation booléenne (THOMAS, 1973, 1979) qui distingue deux états pour chacun des éléments (en fonction ou non, présent ou non, en quantité suffisante ou non). Ces modèles

sont essentiellement statiques et qualitatifs, ils décrivent seulement des états stationnaires, même si les méthodes booléennes, complétées par la notion de délais de transition, permettent une certaine simulation numérique de la dynamique des processus. La variabilité génétique n'est en revanche guère prise en compte dans ces travaux. L'autre approche est fondée sur une représentation par des systèmes dynamiques, et traduit les variations de concentrations en ARN messagers, en enzymes, en substrats, par des équations différentielles liées. Certains essais sont très généraux, comme celui de GOODWIN (1963) qui a tenté de construire une mécanique statistique des processus moléculaires de la biologie en tenant compte des mécanismes de régulation, et comme les études concernant la dynamique des flux dans un réseau métabolique en fonction des activités élémentaires des enzymes (KACSER & BURNS, 1973, 1981). D'autres travaux traitent au contraire de systèmes particuliers, ils s'appuient sur les données structurelles et sur des mesures spécifiques acquises par l'expérimentation. Les systèmes les plus étudiés à cet égard sont les premières étapes de la glycolyse, contrôlées par la phospho-fructo-kinase (HIGGINS, 1964 ; SEL'KOV, 1968 ; DEMONGEOT, 1981) et l'opéron lactose du colibacille (KNORRE, 1968, 1973 ; GOODWIN, 1969 ; BABLOYANTZ & SANGLIER, 1972 ; SANGLIER & NICOLIS, 1976). Ces travaux ont été conduits — comme les premiers essais théoriques de GOODWIN (1963) — avec pour objectif la recherche de structures biologiques susceptibles d'engendrer des oscillations entretenues et d'être des archétypes d'horloges biologiques. Ces modèles de l'opéron lactose, comparés à d'autres études théoriques concernant les seuls mécanismes d'interaction moléculaire au niveau des gènes de régulation (VON HIPPEL, REVZIN, GROSS, WANG, 1974 ; MANDECKI, 1979 ; MANABE, 1981), ont l'intérêt, dans la perspective décrite plus haut, de prendre en compte l'ensemble des mécanismes mis en jeu : perméation puis hydrolyse du lactose, induction et répression des gènes de structure, synthèse coordonnée des enzymes. Contrairement à l'étude de MANDECKI (1979), ces modèles globaux font peu de place à l'étude des souches mutantes, dont l'analyse a été déterminante dans la compréhension du système ; ils comportent également des simplifications injustifiées, qui concernent les vitesses relatives des différentes étapes, et traduisent d'une façon erronée les mécanismes de l'induction et de la répression catabolique (GILLOIS, TABARY, CHEVALET, 1981).

L'objet de ce travail est de proposer, dans une première partie, un modèle détaillé du mécanisme de l'induction de l'opéron lactose. Nous n'avons pas recherché *a priori* une expression mathématiquement simple de l'ensemble du système, mais tenté d'établir un ensemble de relations qui soient toutes justifiées par les études génétiques et biochimiques réalisées sur les différentes parties du système. La cohérence de l'ensemble des équations écrites n'est donc pas acquise d'emblée dans cette démarche, elle est en soi un problème. Sa résolution, et la discussion des résultats obtenus, font l'objet de la deuxième partie du travail.

Description de l'opéron lactose d'*E. coli*

La synthèse des trois enzymes de l'opéron lactose, la β -galactosidase Z, la β -galactoside perméase Y et la thiogalactoside transacétylase A, est coordonnée. Dans une souche sauvage, cette synthèse est réprimée, c'est-à-dire réduite à un taux très faible, en l'absence de lactose dans le milieu de culture et elle peut être induite en présence de ce substrat. Néanmoins, en présence de lactose, la synthèse reste très faible en

présence de glucose. Elle n'est amplifiée que si la concentration externe en glucose est faible (répression catabolique, diauxie). L'addition de glucose à une culture sur lactose provoque également une répression rapide de la synthèse des trois enzymes de l'opéron (répression transitoire).

Les mécanismes en jeu, dans ce système, sont de trois types : (i) le métabolisme du lactose contrôlé par les enzymes de l'opéron ; (ii) la répression et l'induction au niveau du gène opérateur ; (iii) la répression catabolique et l'activation de la transcription au niveau du gène promoteur. Ces mécanismes sont représentés, symboliquement, sur la figure 1.

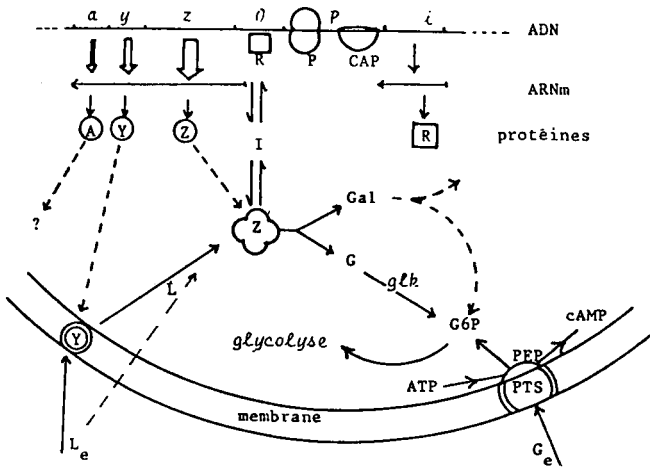


FIG. 1

Les principaux éléments de l'opéron lactose d'Escherichia coli.
Main elements of Escherichia coli lactose operon

(i) Le métabolisme du lactose

Le lactose, présent dans le milieu de culture, peut pénétrer dans la bactérie de façon active par un complexe membranaire incluant la perméase Y, ou de façon passive.

Dans la cellule, le lactose est hydrolysé par l'enzyme Z en glucose et galactose. Mais cette enzyme peut aussi convertir le lactose en allolactose par transgalactosidation. Ce dernier est lui-même substrat de l'enzyme Z et il est l'inducteur naturel de l'opéron.

Le glucose, produit de l'hydrolyse du lactose, est ensuite phosphorylé en glucose-6-phosphate, point de départ de la chaîne glycolytique. Le galactose, autre produit de l'hydrolyse du lactose, est utilisé dans les synthèses membranaires, il peut aussi être converti en glucose-6-phosphate par les enzymes de l'opéron galactose.

L'enzyme A, thiogalactoside transacétylase, n'a pas de fonction physiologique clairement identifiée.

(ii) *La répression et l'induction de l'opéron*

Situé en amont des gènes de l'opéron lactose, le gène *i* est constitutif et transcrit à un faible niveau. Il code pour la protéine R, le répresseur de l'opéron lactose. Ce répresseur peut se lier, d'une part à l'ADN et d'autre part à certains galactosides qui en sont des effecteurs. En absence d'effecteur, la protéine R se fixe sur l'opérateur *o* de l'opéron lactose. Cette association interdit la transcription des gènes de l'opéron en empêchant l'ARN-polymérase de s'associer efficacement au site d'initiation. En présence d'allolactose, inducteur naturel, le complexe R-*o* se dissocie et la protéine R forme avec l'allolactose un complexe qui se fixe beaucoup moins bien sur l'opérateur : la transcription des gènes de l'opéron peut avoir lieu.

(iii) *La répression catabolique et l'activation de la transcription*

Le mécanisme de la répression catabolique n'a été identifié que récemment. Il fait intervenir, d'une part la régulation de la synthèse d'AMP cyclique et d'autre part, un facteur protéique, le CAP. Cette protéine CAP peut fixer l'AMP cyclique, et le complexe cAMP-CAP, susceptible de se lier au promoteur de l'opéron, active la transcription des gènes de structure en présence de lactose ou d'un inducteur non métabolisable, mais en l'absence du complexe cAMP-CAP, le taux de transcription reste faible. L'activité de transcription est modulée par la concentration interne en AMP cyclique.

I. Introduction

L'élaboration du modèle complet de l'induction d'une bactérie réprimée s'appuie sur la représentation des éléments suivants du système :

(i) *Les interactions entre molécules et gènes de régulation*

Il s'agit des interactions entre, d'une part, l'opérateur et le promoteur de l'opéron et, d'autre part, le répresseur, la protéine CAP et l'ARN polymérase. Dans une cellule, les nombres de molécules en jeu sont très petits : un ou deux gènes et zéro ou une molécule régulatrice de chaque espèce, compte tenu des associations non spécifiques de ces protéines avec l'ADN. En conséquence, la formulation mathématique de ces interactions est probabiliste. A chaque instant, la région de contrôle est caractérisée par les probabilités d'existence de ses différents états.

(ii) *La synthèse des enzymes*

Ce mécanisme est complexe mais il peut être représenté de façon condensée en tenant compte des faits suivants : la durée de vie des ARN messagers est très courte (une minute, en moyenne), il n'y a pas de régulation post-transcriptionnelle connue chez les eucaryotes, les durées globales des synthèses enzymatiques depuis

l'initiation de la transcription jusqu'à l'apparition de l'activité enzymatique, sont connues. Ceci permet d'exprimer les taux de synthèse à un instant t en fonction des probabilités des états de la région de contrôle en un instant antérieur $t-\tau$.

(iii) *Le métabolisme du lactose*

Les cinétiques d'apparition des produits de l'hydrolyse et de l'isomérisation du lactose sont écrites d'après les études concernant la perméase et la β -galactosidase. Il faut cependant supposer, pour pouvoir utiliser les équations usuelles de vitesse, que les nombres de molécules d'enzymes soient petits par rapport aux nombres des molécules de substrats.

L'irréversibilité de la perméation active du lactose, et celle de l'hydrolyse du lactose dans des conditions normales des concentrations intracellulaires permet d'étudier isolément le système ainsi construit, si l'on suppose constantes la concentration extérieure en lactose et la concentration intracellulaire en cAMP. Le modèle obtenu décrit donc les phénomènes de l'induction et de la répression. Il prend aussi en compte les données sur la répression catabolique, mais de façon stationnaire : on ne pourra pas décrire les cinétiques de transition entre ces cas limites. En particulier on ne pourra pas encore simuler le phénomène de diauxie, d'autant que le système d'équations établi dans cette première étude ne prend pas explicitement en compte les divisions cellulaires ni la croissance d'une population. Le modèle se réfère à un opéron, il peut décrire une population stationnaire à croissance lente (il n'y a alors en général qu'un opéron par cellule). Dans une population à croissance rapide où chaque cellule possède plusieurs chromosomes en division simultanément, le modèle restera valable si l'on admet que le volume cellulaire est proportionnel au nombre d'opérons présents, ou au nombre de sites d'initiation de la réplication : dans ces conditions, en effet, le rapport entre le nombre de molécules d'un substrat dans une cellule et sa concentration intracellulaire restera sensiblement constant, et les équations relatives aux différents opérons pourront être sommées.

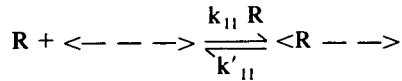
II. Modèle de la région de contrôle

La région de contrôle de l'expression de l'opéron lactose comprend trois sites de fixation pour les protéines qui régulent le niveau de la transcription : l'opérateur o , site du répresseur R , et sur le promoteur, les deux sites voisins où se fixent l'ARN-polymérase P et la protéine CAP . Chacun de ces trois sites peut être libre ou occupé par la molécule qui lui est spécifique ; chaque état possible de cette région sera représenté par une suite de trois symboles : un site libre est désigné par un tiret, et un site occupé, par le symbole de la protéine liée à l'ADN.

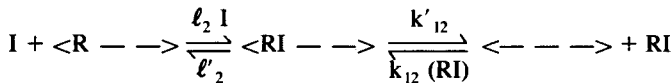
A. *Les interactions entre le répresseur et l'opérateur*

En l'absence de polymérase et de protéine CAP , la région de contrôle peut être dans l'un des deux états, $\langle R \text{ --- } \rangle$ ou $\langle \text{---} \text{---} \text{---} \rangle$, selon que le répresseur est lié ou non à l'opérateur. L'équilibre et la cinétique de cette liaison ont été étudiés

par RIGGS, SUZUKI & BOURGEOIS (1970 a), RIGGS, NEWBY & BOURGEOIS (1970 b), RIGGS, BOURGEOIS & COHN (1970 c), et correspondent à l'équation :

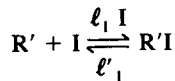


dans laquelle k_{11} et k'_{11} sont des constantes cinétiques, et R représente la concentration en répresseur libre. Le répresseur est un tétramère qui possède, sur chaque monomère, un site de liaison avec certains galactosides. Il a été envisagé, puis établi par des méthodes biochimiques (BEYREUTHER, 1978) que le répresseur, non lié à l'ADN, pouvait subir des transitions allostériques entre deux formes dont l'une, stabilisée par les galactosides inducteurs, présenterait une moindre affinité pour l'opérateur (MONOD, WYMAN & CHANGEUX, 1965). Ce mécanisme est cependant insuffisant pour expliquer la rapidité de l'induction *in vivo* ; il faut admettre l'existence d'un complexe ternaire entre l'opérateur, le répresseur, et l'inducteur, tel que ce dernier déstabilise le complexe répresseur-opérateur :



où I représente un inducteur, et où $k'_{12} > k'_{11}$ (GILBERT & MULLER-HILL, 1967 ; RIGGS *et al.*, 1970 b). Selon les mesures effectuées *in vitro*, le coefficient k'_{12} , correspondant à l'IPTG (*) ou à l'allolactose est environ mille fois plus élevé que k'_{11} ; inversement le lactose est un anti-inducteur qui stabilise le complexe opérateur-répresseur et est caractérisé par un rapport k'_{13}/k'_{11} d'environ un cinquième (JOBE & BOURGEOIS, 1973 ; BARKLEY & BOURGEOIS, 1978). Les coefficients cinétiques d'association sont en revanche insensibles à la présence d'un effecteur sur le répresseur (*ibid.*), soit : $k_{12} = k_{11} = k_1$. Le nombre de molécules de répresseur par opérateur est estimé à 10 ou 20, mais la majeure partie de ces molécules est liée de façon non spécifique à l'ADN, liaison d'ailleurs insensible aux effecteurs du répresseur : KAO-HUANG, REVZIN, BUTLER, O'CONNER, NOBLE & VON HIPPEL (1977) estiment qu'au plus 7 p. 100 de ce nombre total de molécules peut se trouver libre dans la cellule, et donc susceptible d'être mis en jeu dans les liaisons spécifiques.

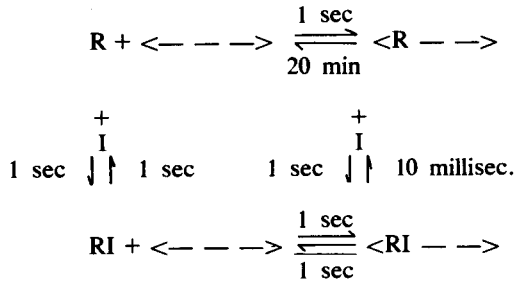
Les quatre sites de liaison d'un effecteur sur le répresseur libre présentent une faible coopérativité ; en les considérant comme indépendants, les caractéristiques de l'association :



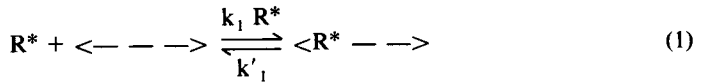
ont pu être déterminées, R' désignant un monomère (BARKLEY & BOURGEOIS, 1978). Sur le complexe répresseur-opérateur, le nombre de sites pourrait être réduit, les résultats de BARKLEY, RIGGS, JOBE & BOURGEOIS (1975) sont compatibles avec l'existence d'un seul site efficace. Les mesures ainsi effectuées indiquent une importante perte d'affinité des effecteurs pour le répresseur quand celui-ci est lié à l'opé-

(*) IPTG : isopropyl- β -D-Thiogalactoside (inducteur gratuit).

rateur. En admettant que la différence est principalement due à une modification de la constante de dissociation, soit $l'_2 > l'_1$ et $l_2 \simeq l_1$, des temps de réaction vraisemblables sont les suivants :



Ainsi les transitions $\left\langle R \text{---} \text{---} \right\rangle \rightarrow \left\langle RI \text{---} \text{---} \right\rangle$ et $\left\langle RI \text{---} \text{---} \right\rangle \rightarrow \left\langle R \text{---} \text{---} \right\rangle$ sont beaucoup plus fréquentes que les transitions concurrentes $\left\langle R \text{---} \text{---} \right\rangle \rightarrow \left\langle \text{---} \text{---} \right\rangle$ et $\left\langle RI \text{---} \text{---} \right\rangle \rightarrow \left\langle \text{---} \text{---} \right\rangle$, respectivement. Il est donc légitime de simplifier ce schéma en admettant que les formes $\left\langle R \text{---} \text{---} \right\rangle$ et $\left\langle RI \text{---} \text{---} \right\rangle$ s'équilibrent instantanément en fonction de la concentration en effecteur I ; parallèlement, l'égalité des coefficients d'association $k_{11} = k_{12} = k_1$ permet de considérer le répresseur non lié à l'ADN comme une seule entité R^* . Le schéma se résume ainsi à :



où $\left\langle R^* \text{---} \text{---} \right\rangle$ représente l'une des formes $\left\langle R \text{---} \text{---} \right\rangle$ ou $\left\langle RI \text{---} \text{---} \right\rangle$ selon la concentration en I. Si un autre effecteur est en compétition, par exemple le lactose L, $\left\langle R^* \text{---} \text{---} \right\rangle$ peut également représenter l'association $\left\langle RL \text{---} \text{---} \right\rangle$. La constante cinétique de dissociation, k'_1 , est fonction des concentrations en I (et éventuellement en L) selon l'expression :

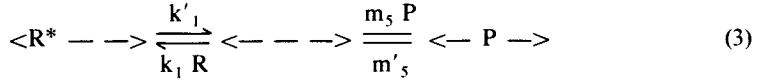
$$k'_1 = \frac{k'_{11} + k'_{12} \frac{l_2 I}{l'_2} + k'_{13} \frac{l_3 L}{l'_3}}{1 + \frac{l_2 I}{l'_2} + \frac{l_3 L}{l'_3}} \tag{2}$$

Cette expression de la constante de dissociation apparente k'_1 en fonction de I est conforme aux résultats expérimentaux de BARKLEY *et al.* (1975).

B. L'initiation de la transcription par l'ARN polymérase

Le site de fixation de l'ARN-polymérase est contigu à celui du répresseur, sans recouvrement. Cependant, la fixation de la polymérase ou celle du répresseur sont deux éventualités exclusives. Le répresseur ne peut pas dissocier de l'ADN une polymérase qui est déjà associée et l'ARN-polymérase ne peut pas former le complexe « ouvert »,

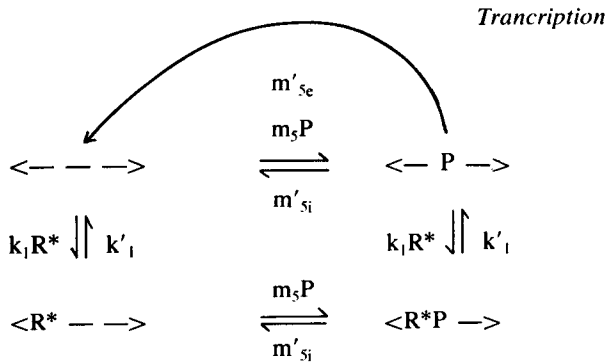
préalable à la transcription, en présence du répresseur (REZNIKOFF, 1976 ; REZNIKOFF & ABELSON, 1978). Nous formulons donc le modèle en respectant cette exclusion mutuelle. Au schéma simplifié (1) nous devons adjoindre l'état où P est liée au promoteur, en absence de répresseur. Les transitions possibles entre états sont les suivantes :



Le paramètre m'_5 décrit la cinétique de libération du site de fixation de la polymérase, mais cette libération peut représenter plusieurs événements. La polymérase peut quitter son site sans que la transcription n'ait débuté, elle peut aussi, en formant le complexe ouvert, commencer la transcription du gène opérateur et du gène z. Ces deux éventualités ont des caractéristiques cinétiques différentes. La première peut correspondre à une transition très rapide si les conditions sont telles que le complexe ouvert a une probabilité faible de se former ; au contraire la seconde nécessite un temps minimum de l'ordre de la seconde, puisque le gène opérateur comprend une quarantaine de bases. Nous caractérisons ces deux événements en décomposant le paramètre m' en deux parties, l'une m'_i décrivant la séparation inefficace de la polymérase et susceptible de dépendre des conditions extérieures au site de reconnaissance de la polymérase, l'autre m'_e décrivant l'initiation de la transcription, et essentiellement constant puisqu'il caractérise le temps de transcription de l'opérateur ($m'_e \simeq 1 \text{ s}^{-1}$). On aura :

$$m' = m'_i + m'_e$$

Le rapport m'_e/m' représente la probabilité que la libération du promoteur par la polymérase conduise à l'initiation de la transcription du messenger. MANABE (1981) a supposé que la polymérase était susceptible de se lier à l'ADN en présence du répresseur, sans toutefois pouvoir provoquer dans cette condition la formation du complexe ouvert. Ce point de vue revient à compléter le schéma de la façon suivante, en identifiant la transition conduisant à la transcription et en supposant l'égalité des coefficients :



Il ne semble pas que cette adjonction modifie beaucoup la cinétique du système, car le facteur limitant, même en présence d'inducteur, est vraisemblablement la quan-

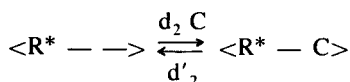
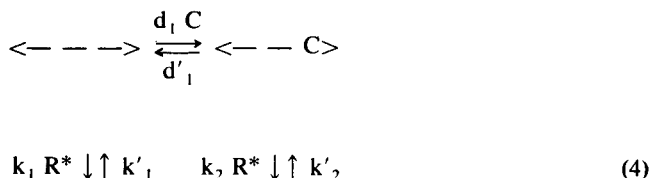
tité k'_1 . La situation pourrait être différente en cas de carence en polymérase car mP deviendrait faible et le passage par un état $\langle R^*P \rangle$ pourrait accélérer la fréquence de la transcription.

C. L'activation de la transcription par le complexe cAMP-CAP

En présence d'AMP cyclique en concentration suffisante dans la cellule, la protéine CAP stimule l'expression des opérons sensibles à la répression catabolique. Pour l'opéron lactose, l'effet du complexe cAMP-CAP est double : le taux de synthèse de l'ARN messager de l'opéron est accru (PERLMAN, CHEN, de COMBRUGGHE, EMMER, COTTESMAN, VARMUS & PASTAN, 1970 ; MAJORS, 1975 a) ; et l'effet du facteur *rho* sur la terminaison de la transcription est atténué ou même supprimé (ULLMAN, JOSEPH & DANCHIN, 1979). L'activation de l'opéron par la protéine CAP est contrôlée par le niveau intracellulaire en AMP cyclique (PERLMAN *et al.*, 1970 ; ARDITTI, ERON, ZUBAY, TOCCHINI-VALENTINI, CONNAWAY & BECKWITH, 1970). *In vivo*, ce niveau intracellulaire de cAMP est corrélé négativement avec l'incorporation de sucres par le système de phosphotransférase dépendant du phospho-énol-pyruvate (PTS). En effet ce complexe membranaire de perméation comprend aussi l'enzyme adénylate cyclase. Quand la bactérie utilise le glucose comme source de carbone, celui-ci est phosphorylé par le système PTS, à partir du phospho-énol-pyruvate qui ne peut être utilisé pour activer la cyclase : le niveau d'AMP cyclique demeure faible et la protéine CAP reste inactive. Au contraire, en absence de sucre pénétrant par le système PTS, la cyclase peut être activée et catalyser la synthèse d'AMP cyclique à partir de l'ATP : la protéine CAP est alors sous sa forme active (PETEKOVSKY, 1977).

Il est généralement admis que le complexe C, de la protéine CAP activée par l'AMP cyclique, agit par sa fixation en un site du promoteur, voisin du site de reconnaissance de l'ARN polymérase. Ce site a été identifié et précisé essentiellement par les études génétiques de mutants insensibles à l'activation par la protéine CAP, et par l'étude comparée des séquences nucléotidiques (REZNIKOFF & ABELSON, 1978). Les études biochimiques ont mis en évidence une association non spécifique entre CAP et ADN, favorisée par l'AMP cyclique (TAKAHASHI, BLAZY & BAUDRAS, 1979), ainsi qu'une association spécifique avec les promoteurs des opérons lactose et galactose (MAJORS, 1975 b) qui paraît difficile à caractériser en raison de la forte coopérativité de la liaison non spécifique (TAKAHASHI *et al.*, 1979). L'interprétation la plus simple du rôle du complexe C est une modification de la structure du site de reconnaissance de l'ARN polymérase, liée à la fixation de C sur un site voisin. Cependant, l'existence de l'association non spécifique entre C et l'ADN, avec des caractéristiques thermodynamiques analogues à celles de l'interaction entre ARN polymérase et ADN, et la mise en évidence d'une interaction entre le complexe C et la polymérase holoenzyme, libre ou liée à l'ADN (BLAZY, TAKAHASHI, BAUDRAS, 1980), permettent d'envisager un déplacement simultané de la protéine CAP et de la polymérase pendant la transcription (TAKAHASHI, 1979), et suggèrent une interprétation de l'effet anti-polaire de la CAP (ULMANN *et al.*, 1979).

En ce qui concerne les états possibles de la région de contrôle, seule la fixation de C, sur son site spécifique, doit être prise en compte. Le schéma (1) s'étend de la façon suivante :

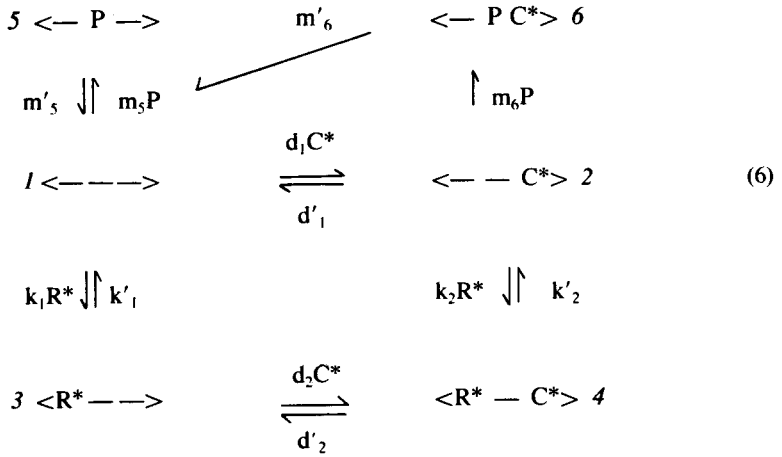


On ne sait rien des interactions entre les sites de fixation spécifique du répresseur et du complexe C, de sorte qu'*a priori*, les coefficients k_2 et k'_2 sont distincts de k_1 et k'_1 , avec une expression de k'_2 analogue à celle de k'_1 :

$$k'_2 = \frac{k'_{21} + k'_{22} \frac{t_2 I}{l'_2} + k'_{23} \frac{t_3 L}{l'_3}}{1 + \frac{t_2 I}{l'_2} + \frac{t_3 L}{l'_3}} \quad (5)$$

en admettant que l'affinité des effecteurs pour un répresseur lié à l'opérateur n'est pas modifiée par la présence du complexe C sur le promoteur. Ces coefficients, non plus que d_1 , d_2 , d'_1 , d'_2 , ne sont connus. De même que l'affinité des inducteurs pour le répresseur est modifiée quand le répresseur est lié à l'opérateur, les caractéristiques de la liaison entre le cAMP et la CAP sont modifiées quand celle-ci est associée à l'ADN (TAKAHASHI, BLAZY, BAUDRAS, 1980). La CAP, dimère, a deux sites de fixation du cAMP, qui présentent une coopérativité variable ; mais la liaison n'a pas été étudiée quand la CAP est liée au promoteur. En admettant, comme pour l'interaction entre le répresseur et l'opérateur, que les constantes d'association d_1 et d_2 sont les mêmes pour la CAP libre et pour le complexe cAMP-CAP, et que les équilibres entre les formes libre et liée de la CAP s'établissent rapidement, on peut représenter par un même symbole C^* ces différentes formes. Les coefficients de dissociation d'_1 et d'_2 peuvent alors dépendre de la concentration en cAMP ; de même il faut considérer que les coefficients de dissociation k'_{21} , k'_{22} et k'_{23} , dans l'équation (5) dépendent aussi de la concentration en cAMP. A ce jour, les études biochimiques réalisées ne permettent pas d'évaluer la dépendance de ces coefficients en fonction de la concentration interne en cAMP.

Enfin, la fixation de la polymérase en présence de C conduit au schéma (6) où l'on suppose que le complexe C libère le promoteur en même temps que la polymérase. Le paramètre m'_6 a le même sens que le paramètre m'_5 , mais l'activation du taux de transcription par la protéine CAP se traduit par une composante inefficace, m'_{6i} , plus petite que m'_{5i} . Comme les coefficients d'_1 et d'_2 , le coefficient m'_6 doit en général être supposé fonction de la concentration en cAMP.



D. *Modèle mathématique*

Numérotons de 1 à 6 les états de la région de contrôle, comme il est indiqué dans le schéma (6). Notons $x_i(t)$ la probabilité qu'à l'instant t la région de contrôle soit dans l'état i ($i = 1, \dots, 6$). Nous étudions la cinétique du système dans un petit intervalle de temps $[t, t + h]$, pour donner ensuite les équations différentielles correspondantes. Pendant cet intervalle de temps, la probabilité pour que la région change une fois d'état est proportionnelle à h , et au nombre de molécules qui peuvent s'associer, si le changement d'état correspond à une association. La probabilité pour qu'elle change deux fois d'état est de l'ordre h^2 . Par exemple, la probabilité de passer de l'état 1 à l'état 2 pendant le temps h est $d_1 C^* h$. Nous pouvons ainsi écrire la probabilité $x_1(t + h)$:

$$x_1(t + h) = x_1(t) + d'_1 h x_2(t) + k'_1 h x_3(t) + m'_5 h x_5(t) + m'_6 h x_6(t) - d_1 C^* h x_1(t) - k_1 R^* h x_1(t) - m_5 P h x_1(t) + O(h^2).$$

On écrit de même les transitions conduisant aux autres états à l'instant $t + h$, selon le schéma (6).

En supposant que l'on ait affaire à des fonctions différentiables (de classe C^1), la limite pour $h \rightarrow 0$, des expressions ci-dessus, nous donne un système dynamique qui s'écrit sous forme matricielle

$$\frac{dX}{dt}(t) = \mathcal{A}(I(t), A(t)) \cdot X(t),$$

$X(t)$ étant le vecteur de composantes $x_i(t)$ ($i = 1, \dots, 6$).

Les éléments de la matrice \mathcal{A} sont donnés sur le tableau 1. Ils sont fonctions des nombres de molécules C^* , R^* et P ainsi que de la concentration en inducteur I et de la concentration A en cAMP par l'intermédiaire des coefficients k'_1 , k'_2 , d'_1 , d'_2 et m'_6 . On remarque que la somme des colonnes de la matrice \mathcal{A} est toujours égale à zéro et que, sous cette forme, la matrice \mathcal{A} est singulière. Pour avoir un système régulier il faut ajouter la condition :

$$\sum_{i=1}^6 x_i(t) = 1, \text{ pour tout } t.$$

TABLEAU 1

Eléments de la matrice \mathcal{A} de transition entre les états de la région de contrôle.
Elements of the transition matrix \mathcal{A} between states of the control region (scheme 6).

Etat final	Etat initial					
	1	2	3	4	5	6
1	$-d_1 C^* - k_1 R^*$ $-m_5 P$	d'_1	k'_1	0	m'_5	m'_6
2	$d_1 C^*$	$-d'_1 - k_2 R^*$ $-m_6 P$	0	k'_2	0	0
3	$k_1 R^*$	0	$-k'_1$ $-d_2 C^*$	d'_2	0	0
4	0	$k_2 R^*$	$d_2 C^*$	$-k'_2$ $-d'_2$	0	0
5	$m_5 P$	0	0	0	$-m'_5$	0
6	0	$m_6 P$	0	0	0	$-m'_6$

Ainsi formulé, ce modèle de la région de contrôle de l'opéron lactose prend en compte l'essentiel des aspects déjà analysés par MANDECKI (1979) et MANABE (1981). Il est plus complet que celui de MANDECKI, par l'introduction de la concentration en inducteur dans les expressions des taux de transition entre états, qui est nécessaire pour analyser le processus de l'induction naturelle. D'autre part, l'analyse que nous avons faite de la relation entre les transitions entre états de la région de contrôle, et le taux d'initiation de la transcription (décomposition en deux termes des paramètres m'_5 et m'_6) constitue un résumé de l'étude détaillée que MANABE a faite des étapes du processus de transcription. Enfin ce modèle pourra intégrer des résultats quantitatifs concernant la liaison entre le cAMP et la protéine CAP, libre ou liée à son site spécifique, et, par la distinction des deux états, $\langle -P \rangle$ et $\langle -PC \rangle$, il permet d'envisager la prise en compte de l'effet anti-polaire de la CAP sur la transcription des gènes de structure.

III. Le modèle de la synthèse enzymatique

Si la région de contrôle est dans l'état j ($j = 5, 6$), à l'instant t , elle le quitte entre t et $t + h$ avec la probabilité $m'_j h$, et cette transition s'accompagne d'une initiation de la transcription avec la probabilité m'_e/m'_j (nous supposons que la composante efficace m'_e ne dépend pas de l'état initial). L'effet éventuel de polarité est exprimé par l'introduction de probabilités, α_{jz} et α_{jy} , pour que la transcription, une fois commencée, se poursuive au moins jusqu'à la fin des gènes z et y , respectivement (donc : $\alpha_{jy} \leq \alpha_{jz}$).

Soit $P_z(t)h$ la probabilité pour qu'entre les instants t et $t+h$ un nouvel ARN messager qui contiendra z commence à être transcrit. Nous avons :

$$P_z(t)h = \sum_{j=5}^6 x_j(t) \alpha_{jz} m'_e h + 0(h^2).$$

Les ribosomes se fixent sur les ARN messagers avant que ceux-ci soient entièrement synthétisés. L'espérance du nombre d'ARN messagers contenant z , qui peuvent être traduits à l'instant t s'écrit donc :

$$\int_{-\infty}^t P_z(\tau) e^{-\gamma(t-\tau)} d\tau$$

où $\frac{1}{\gamma}$ est l'espérance de vie d'un ARN messager, de l'ordre de la minute (KEPES, 1963).

Soit K_z^* le nombre moyen de monomères d'enzyme Z traduits par seconde à partir d'un ARN messager, et soit θ_z le délai qui sépare le début de la traduction, et l'apparition d'une enzyme active. Si $Z(t)$ désigne l'espérance du nombre de sites d'enzyme active à l'instant t , et $K'_z Z(t)$ le nombre de sites détruits par seconde, nous avons :

$$\frac{dZ}{dt}(t) = K_z^* \int_{-\infty}^t P_z(\tau - \theta_z) e^{-\gamma(t-\tau)} d\tau - K'_z Z(t).$$

Si nous supposons que la fonction P_z peut être approchée par son développement linéaire, en tout point, sur un intervalle de temps de l'ordre de la durée de vie d'un ARN messager ($1/\gamma$), l'intégrale ci-dessus peut être remplacée par un retard supplémentaire de $1/\gamma$. En posant alors :

$$\tau_z = \theta_z + \frac{1}{\gamma}, K_z = K_z^* \frac{1}{\gamma} \alpha_{6z} m'_e,$$

$$\text{et } \alpha_z = \alpha_{5z}/\alpha_{6z},$$

nous avons :

$$\frac{dZ}{dt}(t) = K_z (\alpha_z x_5(t - \tau_z) + x_6(t - \tau_z)) - K'_z Z(t).$$

Le paramètre α_z représente l'effet de polarité relatif entre les chaînes d'ARN messagers initiées à partir des états sans ou avec CAP activée.

De même, si $Y(t)$ désigne le nombre de molécules de perméase actives à l'instant t , on a :

$$\frac{dY}{dt} = K_y (\alpha_y x_5(t - \tau_y) + x_6(t - \tau_y)) - K'_y Y(t),$$

les constantes τ_y , K_y , α_y et K'_y étant définies comme pour la β -galactosidase.

On note que les constantes α_y et α_z sont plus petites que 1 car les messagers synthétisés en présence de CAP active sont plus longs qu'en son absence. Cet effet étant plus sensible sur les gènes distaux, nous avons $\alpha_y \leq \alpha_z \leq 1$, dans la mesure où les expériences d'ULLMANN *et al.* (1979), qui montrent que la CAP supprime tout effet de polarité entre les enzymes Z et A, permettent de supposer que $\alpha_{oy} = \alpha_{oz}$.

IV. Le métabolisme du lactose

Le lactose peut pénétrer dans la cellule par deux moyens : 1) Une entrée passive due à la différence de concentration en lactose entre les milieux extérieur et intérieur. Il lui est associé une sortie passive. 2) Une entrée active grâce à un complexe membranaire contenant la β -galactoside perméase, produit du gène *y*. Le nombre de molécules de lactose pénétrant dans la cellule grâce à ce complexe est proportionnel au nombre de perméases $Y(t)$, à chaque instant t , et dépend de la concentration extérieure L_e en lactose, selon la loi de Michaelis (KEPES, 1978).

Si $L(t)$ désigne le nombre de molécules de lactose présentes dans la cellule à l'instant t , le nombre de molécules échangées entre l'extérieur et l'intérieur dans l'intervalle de temps $[t, t + h]$ peut s'écrire :

$$\{k_s(L_e - L(t)) + \frac{k_c L_e}{L_e + K_M} Y(t)\} h + 0(h^2)$$

où k_s est une constante associée au transport passif, K_M est la constante de Michaelis de la perméase et k_c la constante catalytique de la perméase pour le lactose.

A l'intérieur de la cellule, le lactose est métabolisé grâce à la β -galactosidase. Nous avons considéré que les seuls substrats possibles étaient le lactose et l'allolactose et qu'ils étaient en faible quantité. Dans ces conditions, le mécanisme de l'enzyme est représenté sur la figure 2 (HUBER, WALLENFELS & KURZ, 1975 ; HUBER, KURZ & WALLENFELS, 1976).

Le nombre total de sites, ou monomères, de l'enzyme étant désigné par Z , la cinétique de ce mécanisme est décrite, de façon approchée, par les équations suivantes, pourvu que la concentration totale en enzymes soit petite par rapport aux concentrations des substrats (HEINEKEN, TSUCHIYA & ARIS, 1967 ; CHEVALET, GILLOIS & MICALI, 1978, 1981 ; GÖBBER & SEELIG, 1975) :

$$\frac{dL}{dt} = -Z \frac{(a + b)L}{L + \alpha I + \beta}$$

$$\frac{dI}{dt} = Z \frac{bL - \alpha cI}{L + \alpha I + \beta}$$

$$\frac{dG}{dt} = \frac{d \text{Gal}}{dt} = Z \frac{aL + \alpha cI}{L + \alpha I + \beta}$$

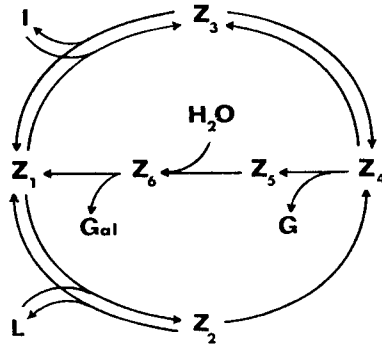


FIG. 2

Mécanisme enzymatique de la β -galactosidase d'E. coli.
 Enzymatic mechanism of E. coli β -galactosidase.

- Z₁ : Enzyme libre.
Free enzyme.
- Z₂ : Complexe enzyme-lactose.
Enzyme-lactose complex.
- Z₃ : Complexe enzyme-allolactose.
Enzyme-allolactose complex.
- Z₄, Z₅, Z₆ : Complexes intermédiaires.
Intermediary complexes.
- L, I, G, Gal : Lactose, allolactose, glucose, galactose.

TABLEAU 2

Paramètres cinétiques de la β -galactosidase d'Escherichia coli.
 Kinetic parameters of Escherichia coli β -galactosidase.

Substrat	Constante de Michaelis	Constantes catalytiques	
		Glucose	Allolactose
Lactose	β	a	b
Allolactose	$\frac{\beta}{\alpha}$	c	

Les constantes qui interviennent peuvent être déterminées expérimentalement en étudiant les vitesses initiales lorsqu'il n'y a qu'un seul substrat. Dans ces conditions, la réaction suit une cinétique du type Michaelis-Menten.

Ainsi, quand il y a uniquement du lactose, on peut écrire :

$$\frac{dI}{dt}(0) = Z(0) \frac{bL(0)}{L(0) + \beta}$$

$$\frac{dG}{dt}(0) = Z(0) \frac{aL(0)}{L(0) + \beta}$$

Quand il y a uniquement de l'allolactose, on a :

$$\frac{dG}{dt}(0) = Z(0) \frac{cI(0)}{I(0) + \beta/\alpha}$$

Ceci nous permet d'établir le tableau de correspondance pour les différentes constantes (tabl. 2).

V. Etablissement et ajustement du modèle de l'induction de l'opéron lactose

Les équations que nous avons écrites décrivent les mécanismes spécifiques de l'opéron lactose, et constituent un modèle complet. Ce système d'équations est cependant insuffisant pour décrire un phénomène comme la diauxie. Pour cela il faut représenter toutes les voies conduisant à la production du glucose-6-phosphate qui est le point de départ commun de la chaîne de la glycolyse. Cette complétion du modèle suppose donc : (1) la prise en compte de la cinétique de l'enzyme galactokinase *glk* ; (2) l'étude de l'induction de l'opéron galactose ; (3) la cinétique d'incorporation du glucose par le système PTS, celle de la production de cAMP et l'activation des deux opérons lactose et galactose par le complexe cAMP-CAP (figure 3).

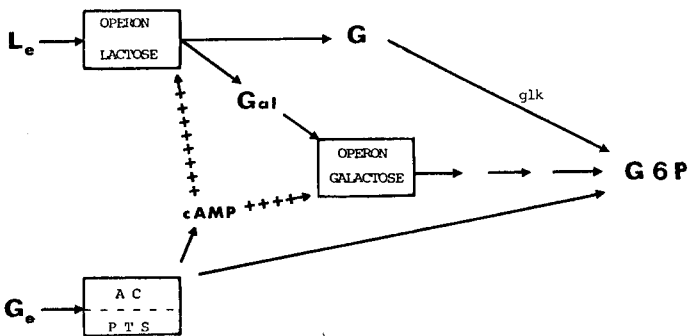


FIG. 3

Schéma de l'utilisation du glucose ou du lactose comme source de carbone. AC désigne l'adénylate cyclase, PTS le système membranaire de perméation et de phosphorylation du glucose, G-6-P, le glucose-6-phosphate.

Scheme of the regulations involved in the use of either glucose or lactose as the carbone source. AC : adenylate cyclase enzyme, linked to the permeation and phosphorylation system PTS ; G6-P : glucose-6-phosphate.

TABLEAU 3

*Définitions et notations des paramètres et des variables du modèle.
Definitions and notations of parameters and variables in the model.*

1. INTERACTIONS ENTRE LES MOLÉCULES RÉGULATRICES ET LES GÈNES DE CONTRÔLE

Répresseur : R ; inducteur : I ; complexe activateur cAMP-CAP : C ; polymérase : P.

11. Paramètres

Interaction	Définition	Symbole	
Inducteur et répresseur lié à l'opérateur.	Coefficient d'équilibre de l'association.	I_2/I'_2	
Répresseur et opérateur.	Coeff. d'association. Coeff. de dissociation ($I = 0$). Coeff. de dissociation ($I = \infty$). Coeff. de dissociation apparent (formules (2) et (5)).	Sans C k_1 k'_{11} $K_1 k'_{11}$ $k'_1(I)$	Avec C k_2 k'_{21} $K_2 k'_{21}$ $k'_2(I)$
Complexe activateur et promoteur.	Coeff. d'association. Coeff. de dissociation.	Sans R d_1 d'_1	Avec R d_2 d'_2
Polymérase et promoteur.	Avec R, la liaison est supposée impossible. Sans R, et : Coeff. d'association. Coeff. de dissociation. Probabilité d'initiation de la transcription par unité de temps à partir d'un état lié.	Sans C m_5 m'_5	Avec C m_6 m'_6
		m'_6	

12. Variables

Les probabilités $x_i(t)$ des six états possibles de la région de contrôle (schéma (6)). La transcription peut être initiée seulement à partir des états 5 ou 6.

2. TRANSCRIPTION, ET SYNTHÈSE DES ENZYMES PERMÉASE ET β -GALACTOSIDASE

21. Paramètres

- délais globaux de la synthèse des enzymes, entre le début de la transcription et l'apparition de l'activité : τ_y et τ_z ;
- effets relatifs de la polarité de la transcription en absence du complexe activateur : α_y et α_z ;
- efficacités globales du processus de traduction : K_y et K_z .

22. *Variables*

Concentrations intracellulaires en perméase (Y) et en β -galactosidase (Z), exprimées en nombres de sites actifs par cellule.

3. *ACTIVITÉS ENZYMATIQUES*31. *Paramètres*

Perméase : constante de Michaelis : K_m
 constante catalytique : k_c
 coefficient de dégradation : K'_y

β -galactosidase : constantes de Michaelis : β et β/α
 constantes catalytiques : a, b, c (tabl. 2)
 coefficient de dégradation : K'_z

32. *Variables*

L_e : concentration du lactose dans le milieu
 L : concentration du lactose dans la cellule
 I : concentration de l'allolactose (inducteur) dans la cellule
 $dG/dt = dGal/dt$: vitesse de production du glucose et du galactose, par hydrolyse du lactose.

Le modèle écrit peut néanmoins être étudié isolément, pourvu que l'on se restreigne aux deux conditions suivantes :

(i) L'enzyme *glk* et l'opéron galactose fonctionnent normalement, de sorte que les concentrations internes en glucose et galactose demeurent peu élevées : la production de ces sucres par la β -galactosidase est alors un mécanisme irréversible et les équations en L et I ne dépendent pas des niveaux de glucose et de galactose.

(ii) La concentration interne en cAMP demeure constante, cela se justifie si l'on suppose que le milieu de culture ne contient pas de sucre pénétrant par le système PTS.

A. *Structure générale du modèle*

Les différentes variables qui permettent de décrire l'état de la cellule à l'instant t sont les suivantes : le vecteur $X(t)$ des probabilités de chaque état de la région de contrôle de l'opéron lactose, le nombre $Y(t)$ de perméases fonctionnelles, le nombre $Z(t)$ de monomères de β -galactosidase fonctionnels, le nombre $L(t)$ de molécules de lactose intracellulaire et le nombre $I(t)$ de molécules d'inducteur (allolactose) intracellulaire. La signification des symboles, paramètres et variables utilisés dans ces équations est rappelée de façon synthétique dans le tableau 3.

Au cours du temps, ces variables vérifient les équations suivantes :

$$\frac{dX}{dt}(t) = \mathcal{A}(I(t)) \cdot X(t)$$

$$\frac{dY}{dt}(t) = K_y (\alpha_y x_5 (t - \tau_y) + x_6 (t - \tau_y)) - K'_y Y(t)$$

$$\frac{dZ}{dt}(t) = K_z (\alpha_z x_5 (t - \tau_z) + x_6 (t - \tau_z)) - K'_z Z(t) \quad (7)$$

$$\frac{dL}{dt}(t) = \frac{k_c L_e}{K_M + L_e} Y(t) + k_s (L_e - L(t)) - Z(t) \frac{(a + b) L(t)}{L(t) + \alpha I(t) + \beta}$$

$$\frac{dI}{dt}(t) = Z(t) \frac{bL(t) - \alpha cI(t)}{L(t) + \alpha I(t) + \beta}$$

Ce système assure un flux, du lactose extérieur L_e , vers les deux produits d'hydrolyse, glucose G et galactose Gal . Cette sortie du système est donnée par l'équation, dépendant du précédent système mais n'interférant pas avec lui :

$$\frac{dG}{dt} = \frac{dGal}{dt} = Z(t) \frac{aL(t) + \alpha cI(t)}{L(t) + \alpha I(t) + \beta} \quad (8)$$

Un modèle similaire, plus simple, décrit le mécanisme d'induction par un inducteur gratuit, l'IPTG ou l'allolactose pour une souche z^- . Dans ce cas, l'inducteur pénètre par la perméase, comme le lactose dans la situation naturelle, et n'est pas métabolisé par la β -galactosidase. L'équation en $Z(t)$ disparaît, et les deux équations en $L(t)$ et en $I(t)$ sont remplacées par une nouvelle équation en $I(t)$, qui correspond aux premiers termes de l'équation précédente en $L(t)$. Ce modèle de l'induction gratuite s'écrit :

$$\frac{dX}{dt}(t) = \sigma(I(t)) \cdot X(t)$$

$$\frac{dY}{dt}(t) = K_y (\alpha_y x_5 (t - \tau_y) + x_6 (t - \tau_y)) - K'_y Y(t) \quad (9)$$

$$\frac{dI}{dt}(t) = \frac{k_c^\circ E}{E + K_m^\circ} Y(t) + k_s^\circ (E - I(t))$$

où k_c° , k_m° et k_s° sont les paramètres cinétiques de la perméation de l'inducteur gratuit, et E représente la concentration de cet inducteur dans le milieu.

La structure du modèle (7) peut être considérée comme assez générale pour des opérons régulant un réseau métabolique :

(i) La région de contrôle apparaît comme un automate commandant la mise en place d'une fonction selon les conditions physiologiques ; les modalités de ce contrôle peuvent varier (régulation positive ou négative, contrôle par le substrat ou par le produit du réseau commandé), mais les mécanismes mis en jeu, c'est-à-dire la définition des états de l'automate par des protéines régulatrices admettant les substrats ou les produits comme effecteurs, semblent assez généraux. Les différentes combinaisons possibles

entre les sites de régulation, sur l'ADN, ne devraient pas modifier fondamentalement ces traits généraux, ni les propriétés qualitatives des réponses.

(ii) Les enzymes, produits primaires des gènes régulés, sont ici, directement, les « robots de la chaîne de fabrication ». Cette relation directe entre produit primaire du gène et entité active est sans doute la plus simple. Les mécanismes de régulation post-transcriptionnelle, chez les eucaryotes, et les processus d'amplification par activations d'enzymes présentes mais non fonctionnelles, sont d'autres modalités de l'action directe d'une induction.

(iii) Le réseau métabolique est ici très rudimentaire, et ne présente pas, normalement, de comportement réversible. Des mécanismes particuliers, comme la rétroinhibition de la première enzyme par le produit final (dans l'opéron tryptophane par exemple), ou comme la régulation allostérique d'une enzyme, peuvent interférer avec le mécanisme génétique d'induction. Il est clair que ce genre de problème ne peut être discuté dans le cas de l'opéron lactose.

Le modèle complet (7) ou (9) est un système d'équations différentielles présentant trois types de non-linéarités : les cinétiques enzymatiques introduisent des fonctions homographiques des variables, la synthèse enzymatique fait apparaître des arguments retardés, enfin les équations d'évolution des états de la région de contrôle constituent un système linéaire dans les variables x_i ($i = 1, \dots, 6$), mais dont certains coefficients de la matrice \mathcal{A} (tabl. 1) sont des fonctions des variables I et L (relations (2) et (5)). Par conséquent, les propriétés qualitatives des solutions peuvent dépendre des valeurs numériques des paramètres, et pas seulement de la forme des équations. Selon les valeurs des paramètres, le système peut présenter un ou plusieurs états d'équilibre, ce ou ces états d'équilibre peuvent être stables ou instables, enfin le système peut posséder des solutions stationnaires non constantes ou même des comportements chaotiques. Il en résulte que le modèle n'est vraiment déterminé, et susceptible de décrire le processus biologique, que si l'on détermine tous ses paramètres.

Enfin, ce modèle est semi-probabiliste : le traitement des états de la région de contrôle est celui d'un processus stochastique ; au contraire les équations de cinétique enzymatique sont de caractère déterministe, les concentrations en substrats étant considérées comme des variables continues. La liaison entre ces deux modalités de représentation est assurée par les équations décrivant la synthèse enzymatique, qui concernent les espérances des nombres d'enzymes synthétisées par unité de temps. Cette simplification n'est valable que si les conditions sont telles que les fluctuations aléatoires des taux de synthèse sont petites par rapport aux valeurs moyennes, seules considérées ici. Le modèle ne permet donc de décrire les premières étapes du processus d'induction que d'une façon assez approchée.

B. Détermination des valeurs numériques des paramètres du génotype sauvage

Certains paramètres du modèle ont fait l'objet d'études systématiques, il s'agit : des caractéristiques cinétiques de la β -galactosidase (HUBER *et al.*, 1975, 1976) ; des coefficients cinétiques des réactions d'association et de dissociation entre l'opérateur, le répresseur et ses effecteurs (mise au point dans : BARKLEY & BOURGEOIS, 1978) ; et des délais (τ_y et τ_z) s'écoulant entre le début de la transcription, et l'apparition des activités de la perméase et de la galactosidase (WEST & STEIN, 1973). Les valeurs correspondant au type sauvage sont indiquées dans le tableau 4-1.

TABLEAU 4

Valeurs numériques des paramètres du modèle.
(cf. paragraphe 5-2).

Numerical values of the parameters.
(see the text, paragraph 5-2).

Tableau 4-1 : Valeurs mesurées in vitro.

Symbole	Valeur	Unité	Symbole	Valeur	Unité
R*	0,2	mol (1)	k ₁	6	mol ⁻¹ s ⁻¹
k' ₁₁	6 · 10 ⁻⁴	s ⁻¹	K ₁	500	—
K ₃	0,2	—	l ₂ /l' ₂	5 · 10 ⁻⁶	mol ⁻¹
l ₃ /l' ₃	2 · 10 ⁻⁸	mol ⁻¹	τ _z	240	s
τ _y	300	s	K' _z	1,1 · 10 ⁻³	s ⁻¹
a	64	s ⁻¹	b	56,5	s ⁻¹
c	97,5	s ⁻¹	β	2,5 · 10 ⁶	s ⁻¹
α	2,1	—			

Tableau 4-2 : valeurs évaluées indirectement.

Symbole	Valeur	Unité	Symbole	Valeur	Unité
k ₂	6	mol ⁻¹ s ⁻¹	k _c	55	s ⁻¹
K _M	10 ⁵	mol	k _s	10 ⁻³	s ⁻¹
K _Z	50	mol s ⁻¹	K _Y	25	mol s ⁻¹
K' _Y	5,5 · 10 ⁻⁴	s ⁻¹	m' _e	1	s ⁻¹

Tableau 4-3 : Valeurs estimées à partir du modèle.

Symbole	Valeur	Unité	Symbole	Valeur	Unité
k' ₂₁	6 · 10 ⁻⁴	s ⁻¹	K ₂	5 000	—
m' ₀	1	s ⁻¹	m ₀ P = m ₀ P	0,75	s ⁻¹
d ₁ C* = d ₂ C*	2	s ⁻¹	d' ₁	0,2	s ⁻¹
d' ₂	0,2	s ⁻¹			
Variante A			Variante B		
m' ₅	25	s ⁻¹	m' ₅	1	s ⁻¹
α _z	1	—	α _z	0,04	—
P̄ _Z (2)	0,27	s ⁻¹	P̄ _Z	0,25	s ⁻¹
Paramètres non estimés					
K ₄	...	—	α _y	...	—

(1) Les concentrations sont exprimées en molécules (ou sites actifs), par cellule bactérienne. Pour un volume cellulaire de 1,5 μm³, une concentration molaire correspond à 10⁹ mol.

Concentrations are expressed as numbers of molecules (mol) per cell. With a cell volume of about 1.5 μm³, one molar concentration involves 10⁹ mol.

(2) Le symbole P̄_Z désigne le taux de transcription du gène z, sous induction.

This symbol, P̄_Z, stands for the transcription rate of the z gene under full induction.

D'autres paramètres n'ont pas été mesurés *in vitro* comme les précédents, mais peuvent être évalués indirectement, ce sont les paramètres caractérisant la cinétique de la perméation du lactose, et ceux décrivant l'efficacité globale des mécanismes de transcription et de traduction (tabl. 4-2). Les paramètres de perméation ont fait l'objet d'études assez nombreuses, mais dont les résultats ne sont pas toujours transposables à la situation naturelle. Le coefficient, k_s , de perméation passive a été bien étudié, mais surtout pour des thiogalactosides (KEPES, 1978), de même les coefficients k_c et K_m de la perméase ont été évalués soit pour des analogues du lactose, soit indirectement par mesure des produits d'hydrolyse du lactose dans une souche mutante ne possédant pas l'enzyme *glk*. L'emploi de substrats différents du lactose, de souches mutantes, ou de vésicules rend la détermination de ces paramètres très approchée : seul l'ordre de grandeur doit en être tenu pour certain (KEPES, 1978 ; MALONEY & WILSON, 1978 ; WRIGHT, RIEDE & OVERATH, 1981). Les coefficients K_y , K'_y et K_z sont de même évalués approximativement. Seul le coefficient K'_z a été évalué *in vitro* (WATSON, 1976), les autres en ont été déduits en s'appuyant sur les faits suivants, admis pour une bactérie sauvage complètement induite : les nombres de sites actifs de β -galactosidase et de perméase sont d'environ 10 000 par cellule (i.e. par opéron) (WATSON, 1976) ; le taux d'initiation des transcriptions est d'environ un messenger toutes les quatre secondes (WATSON, 1976 ; MANABE, 1981) ; il existe une polarité naturelle de la transcription qui se traduirait par la synthèse d'une perméase pour deux monomères de galactosidase (WATSON, 1976).

Enfin, tous les coefficients définissant les transitions de la région de contrôle sont inconnus en dehors de ceux déjà mentionnés. Nous les avons estimés en ajustant le modèle sur les résultats de quelques expériences correspondant à des situations extrêmes combinant des mutants i^- (sans répresseur) et crp^- (sans protéine CAP) avec un niveau nul ou saturant de l'inducteur non métabolisable IPTG. Cela définit six conditions, qui ont fait l'objet d'expériences (JACOB & MONOD, 1961 ; HOPKINS, 1974), et qui corres-

pondent à des cas où le système $\frac{dX}{dt} = \mathcal{A}(I).X$ est autonome. L'état d'équilibre

de la région de contrôle donne alors les taux d'initiation de la transcription, dont les valeurs relatives peuvent être comparées aux valeurs relatives des niveaux d'activité de la β -galactosidase mesurés dans les expériences. Les six conditions sont représentées dans le modèle par les valeurs de R^* , C^* et I données au tableau 5.

Comme nous l'avons déjà observé pour les paramètres k_{11} et k_{12} , ou admis, pour t_1 et t_2 , nous avons postulé l'égalité des coefficients d'association semblables, soit $k_2 = k_1$, $d_1 = d_2$ et $m_5 = m_6$. Le nombre de paramètres étant par ailleurs plus élevé que le nombre d'observations, nous avons d'abord essayé d'ajuster le modèle en supposant qu'il n'y avait pas d'interaction entre l'opérateur et le site de la protéine CAP,

$$\text{soit } k'_{11} = k'_{21}, K_2 = \frac{k'_{22}}{k'_{21}} = \frac{k'_{12}}{k'_{11}} = K_1 \text{ et } d'_1 = d'_2.$$

Le résultat important est qu'aucun ajustement n'est possible dans ces conditions, quelles que soient les valeurs dC^* , mP , m'_5 , m'_6 , $d'_1 = d'_2$ et α_z . La recherche de l'ajustement a montré que toutes ces restrictions pouvaient être maintenues, sauf une,

$K_2 = K_1$, alors que le maintien de cette restriction et la levée des autres n'apportait aucune amélioration : il était impossible de rendre compte de l'accroissement de l'activité galactosidasique après induction. MANDECKI (1979) avait abouti à la même conclusion, notre analyse renforce sa conclusion dans la mesure où l'introduction d'un effet de polarité (α_z) ne permet pas de mieux ajuster le modèle, si $K_2 = K_1$; comme cet auteur nous obtenons l'estimation, $K_2/K_1 \simeq 10$, signifiant qu'en présence du complexe C, l'effet déstabilisateur de l'inducteur sur le complexe répresseur-opérateur est renforcé.

TABLEAU 5

Expressions relatives de l'activité de la β -galactosidase selon le niveau d'induction et le génotype.

Relative expressions of β -galactosidase activity as function of the genotype and of the induction level.

Génotype	crp ⁻ i ⁻	crp ⁻ i ⁺	crp ⁻ i ⁺	crp ⁺ i ⁻	crp ⁺ i ⁺	crp ⁺ i ⁺
Induction	—	+	...	—	+
C*	0	0	0	1	1	1
R*	0	0,2	0,2	0	0,2	0,2
I	0	∞	...	0	∞
Résultats expérimentaux	70	0,07	20	1 200	1	1 000
Modèle variante A	107	0,06	22	1 250	1,04	1 000
Modèle variante B	68	0,06	21	1 180	1,12	1 000

Les deux effets décrits du complexe cAMP-CAP sur l'activation de la transcription (paramètres m'_5 et m'_6) et sur la polarité (paramètres α_y et α_z) n'ont pas pu être séparés : deux ensembles extrêmes de valeurs ($m'_5 = 25 \text{ s}^{-1}$, $\alpha_z = 1$ ou $m'_5 = 1 \text{ s}^{-1}$, $\alpha_z = 0,04$, avec $m'_6 = m'_e = 1 \text{ s}^{-1}$) donnent des résultats semblables. L'évaluation quantitative des rôles respectifs des deux phénomènes nécessiterait de reprendre les mêmes expériences en mesurant parallèlement les activités de la β -galactosidase et de la thiogalactoside transacétylase.

Les constantes de la liaison de la protéine CAP avec le promoteur ne sont pas connues, et la valeur efficace de C* non plus. Les calculs d'ajustement ont montré que les expressions relatives théoriques étaient peu sensibles aux valeurs absolues de $d_1C = d_2C$ et de $d'_1 = d'_2$, les meilleurs résultats étant obtenus avec $dC/d' \simeq 10$; l'adéquation du modèle à un taux de transcription vraisemblable conduit ensuite à des valeurs de ces paramètres proches de l'unité (tabl. 4-3).

VI. Discussion

L'établissement du précédent modèle est fondé sur les principales propriétés qualitatives de l'opéron lactose, mises en évidence par les études génétiques et biochimiques. La transcription quantitative de ces propriétés s'appuie sur quelques principes généraux :

(i) la loi d'action de masse est valable pour toutes les interactions moléculaires prises en compte ; cela vaut autant pour les réactions de diffusion le long de l'ADN que dans l'espace intracellulaire ;

(ii) quand la constante d'équilibre d'une réaction est modifiée par un effecteur, on suppose que seule la constante de dissociation est altérée ;

(iii) les hiérarchies entre nombres de molécules (opérateur < molécules régulatrices < enzymes < substrats et effecteurs), ainsi que les hiérarchies de vitesse entre différentes réactions concurrentes sont utilisées pour remplacer des équations différentielles par des relations algébriques ;

(iv) les cinétiques enzymatiques sont caractérisées par des équations complètes obtenues par application de l'hypothèse d'un état stationnaire entre les différentes formes des complexes enzymes-substrats ; en toute rigueur, cette approche est valable seulement si le système est clos et si les concentrations des substrats sont grandes par rapport à celles des enzymes ;

(v) le processus complexe de la synthèse protéique est résumé par une équation relative à la valeur moyenne du taux de synthèse enzymatique en fonction du taux d'initiation de la transcription à un instant précédent. La courte durée de vie de l'ARN messager dans ce système permet de ne pas introduire une variable supplémentaire (la concentration en lac-mRNA), mais les fluctuations de ce taux de synthèse autour de la valeur moyenne calculée peuvent jouer un rôle important quand elles sont grandes par rapport à la concentration totale de l'enzyme. Pour cette raison et la précédente (iv), le modèle décrira mal les étapes initiales de l'induction.

Pour déterminer les valeurs numériques des paramètres, nous avons autant que possible utilisé des données recueillies *in vitro* sur des parties isolées du système (interactions répresseur-opérateur-inducteur, cinétiques enzymatiques, stabilité de la β -galactosidase), ou des résultats d'expériences *in vivo* conçues pour la détermination du paramètre considéré (délais de synthèse des enzymes, coefficients de perméation passive des galactosides). Au contraire, d'autres évaluations résultent de la combinaison d'expériences distinctes, qui n'avaient pas été conçues pour que leurs résultats soient exploités conjointement : estimation des paramètres cinétiques de la perméase (expériences de perméation interprétées avec une évaluation du nombre de perméases faite dans une autre souche), estimation des coefficients d'amplification de la traduction (évaluation des nombres d'enzymes dans une souche induite, estimation du taux de transcription après induction, et mesure de la durée de vie de la β -galactosidase), estimation de la durée de vie de la perméase (déduite des résultats indirects précédents). Enfin les paramètres complètement inconnus ont été évalués à l'aide du modèle, des paramètres connus concernant la région de contrôle, de l'estimation du taux de transcription sous induction, et des résultats de deux séries d'expériences indépendantes. Les expériences dont nous avons réuni les résultats devraient être refaites dans des conditions strictement équivalentes ; ne disposant pas de ces résultats, nous devons étudier l'incidence de cette grande incertitude sur les propriétés du système.

En revanche, les paramètres du modèle, en dehors de ceux qui décrivent des fonctions générales non liées génétiquement à l'opéron lactose (synthèse protéique, perméation passive), possèdent une signification génétique directe, telle que la modification d'un paramètre désigne quel est le gène qui doit muter pour réaliser cette modification, et quel est le mécanisme altéré. Cela ne donne pas le moyen de sélectionner la mutation, mais peut en orienter la recherche, cela peut aussi permettre de désigner un mutant déjà connu ; inversement il est possible avec le modèle de simuler le comportement d'un génotype réunissant plusieurs mutations biochimiquement caractérisées. Le modèle que nous avons écrit présente ainsi un avantage sur les résumés proposés par GOODWIN (1969), KNORRE (1968), BABLOYANTZ & SANGLIER (1972), SANGLIER & NICOLIS (1976) : il permet de relier clairement des valeurs numériques de paramètres à des variations génétiques. Ce point est essentiel pour comprendre le sens biologique de résultats mathématiques. En contrepartie, cette clarté de l'écriture a pour corollaire la complexité mathématique du modèle. En termes mathématiques, la question qui se pose ainsi est celle de la stabilité structurelle du modèle : nous avons justifié son élaboration par des arguments empruntés à la biochimie et à la biologie moléculaire ; mais il reste à justifier que ce modèle est capable de décrire un comportement observable, que les propriétés qualitatives de ses solutions demeurent inchangées quand on fait varier les paramètres dans des domaines assez larges. La deuxième partie de ce travail répond à cette question, en montrant d'une part que les solutions correspondant aux différentes souches mutantes connues présentent les mêmes caractéristiques topologiques (un état stationnaire constant et stable), et en désignant d'autre part quels sont les paramètres génétiques ou de milieu dont les variations ne risquent pas de modifier le comportement du système, et quels sont ceux, au contraire, auxquels le système est le plus sensible, quantitativement et qualitativement.

En écrivant un modèle quantitatif de l'opéron lactose, nous avons proposé une synthèse des connaissances acquises sur ce système. Cette synthèse réunit les principaux faits expérimentaux qui ont contribué à la compréhension de l'ensemble, et notamment elle tient compte des comportements des mutants qui ont éclairé les mécanismes de la régulation. Ce travail a montré que les mécanismes de certaines parties du système demeuraient mal compris, ou mal quantifiés (interactions entre la protéine CAP et le promoteur, paramètres de la perméase du lactose, relation entre la cinétique d'incorporation du glucose et la synthèse de l'AMP cyclique). L'incidence de ces incertitudes sur la validité du modèle peut être jugée *a posteriori* par la comparaison des prédictions du modèle aux résultats expérimentaux. La même comparaison permettra de savoir si les approximations admises étaient justifiées, ou si, au contraire, certains détails avaient une importance primordiale. Cela ne devrait pas être le cas pour l'ensemble des résultats concernant les séquences des acides nucléiques et des protéines, dont la description — dans l'état actuel des connaissances — n'apporte aucune information sur les propriétés biochimiques des macromolécules correspondantes.

Reçu pour publication le 26 juillet 1982.

Accepté pour publication le 9 décembre 1982.

Références bibliographiques

- ARDITTI R.R., ERON L., ZUBAY G., TOCCHINI-VALENTINI G., CONNAWAY S., BECKNITH J., 1970. *In vitro* transcription of the *lac* Operon genes. *Cold Spring Harbor Symp. Quant. Biol.*, **35**, 425-432.
- BABLOYANTZ A., SANGLIER M., 1972. Chemical instabilities of « all or none » type in β -galactosidase induction and active transports. *F.E.B.S. letters*, **23**, 364-366.
- BARKLEY M.D., BOURGEOIS S., 1978. Repressor recognition of operator and effectors. In : *The operon* (ed. J.H. Miller and N.S. Reznikoff), Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 177-220.
- BARKLEY M.D., RIGGS A.D., JOBE A., BOURGEOIS S., 1975. Interaction of effective ligands with *lac* repressor and repressor-operator complex. *Biochemistry*, **14**, 1700-1712.
- BLAZY B., TAKAHASHI M., BAUDRAS A., 1980. Binding of CRP to DNA-dependent RNA polymerase from *E. coli* : modulation by cAMP of the interactions with free and DNA-bound holo and core enzyme. *Molec. Reports*, **6**, 39-43.
- BEYREUTHER K., 1978. Chemical structure and functional organization of *lac* repressor from *Escherichia coli*. In : *The operon* (ed. J.M. Miller, W.S. Reznikoff), Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 123-154.
- CHEVALET C., GILLOIS M., MICALI A., 1978. Sur un modèle en cinétique enzymatique. *C.R. Acad. Sc. Paris*, **287**, 169-172.
- CHEVALET C., GILLOIS M., MICALI A., 1981. Sur un modèle en cinétique enzymatique, II. In : *Modèles mathématiques en biologie* (Chevalet C. et Micali A., éd.), Lecture notes in biomathematics, vol. **41**, Springer Verlag, Berlin, Heidelberg, New York, 28-39.
- DEMONGEO J., 1981. Existence de solutions périodiques pour une classe de systèmes différentiels gouvernant la cinétique de chaînes enzymatiques oscillantes. Application à la glycolyse. In : *Modèles mathématiques en biologie* (Chevalet C. et Micali A., éd.), Lecture notes in biomathematics, vol. **41**, Springer Verlag, Berlin, Heidelberg, New York, 40-62.
- GILBERT W., MULLER-HILL, 1967. The *lac* operator is DNA. *Proc. Nat. Acad. Sci. U.S.A.*, **58**, 2415-2418.
- GILLOIS M., TABARY F., CHEVALET C., 1981. Biologie et mathématiques. In : *Modèles mathématiques en biologie* (Chevalet C. et Micali A., éd.), Lecture notes in biomathematics, vol. **41**, Springer Verlag, Berlin, Heidelberg, New York, 3-13.
- GÖBBER F., SEELIG F.F., 1975. Conditions for the application of the steady-state approximation to systems of differential equations. *Journal of Mathematical Biology*, **2**, 79-86.
- GOODWIN B.C., 1963. *Temporal organization in cells*, Academic Press, London.
- GOODWIN B.C., 1969. Control dynamics of β -galactosidase in relation to the bacterial cell cycle. *Eur. J. Biochem.*, **10**, 515-522.
- HEINEKEN F.G., TSUCHIYA H.M., ARIS R., 1967. On the mathematical status of the pseudo-steady state hypothesis of biochemical kinetics. *Mathematical Biosciences*, **1**, 95-113.
- HIGGINS J., 1967. Oscillating reactions. *Ind. Eng. Chem.*, **5**, 18-62.
- VON HIPPEL P.H., REVZIN A., GROSS C.A., WANG A.C., 1974. Non-specific DNA binding of genome regulating proteins as a biological control mechanism : I. The *lac* operon : Equilibrium aspects. *Proc. Nat. Acad. Sci. U.S.A.*, **71**, 4808-4812.
- HOPKINS J.D., 1974. A new class of promoter mutations in the lactose operon of *Escherichia coli*. *J. Mol. Biol.*, **87**, 715-724.
- HUBER R.E., WALLENFELS K., KURZ C., 1975. The action of β -galactosidase (*Escherichia coli*) on allolactose. *Can. J. Biochem.*, **53**, 1035-1038.
- HUBER R.E., KURZ G., WALLENFELS K., 1976. A quantitation of the factors which affect the hydrolyse and transgalactosidase activities of β -galactosidase (*E. coli*) on lactose. *Biochemistry*, **15**, 1994-2001.
- JACOB F., MONOD J., 1961. On the regulation of gene activity. *Cold Spring Harbor Symp. Quant. Biol.*, **28**, 193-210.

- JOBE A., BOURGEOIS S., 1973. *Lac* repressor operator interaction : VIII. Lactose is an anti-inducer of the *lac* operon. *J. Mol. Biol.*, **75**, 303-313.
- KACSER H., BURNS J.A., 1973. The control of flux. *Symp. Soc. Exp. Biol.*, **27**, 65-104.
- KACSER H., BURNS J.A., 1981. The molecular basis of dominance. *Genetics*, **97**, 639-666.
- KAO-HUANG Y., REVZIN A., BUTLER A.P., O'CONNOR P., NOBLE D.N., VON HIPPEL P.H., 1977. Non specific DNA binding of genome regulating proteins as a biological control mechanism : measurement of DNA bound *Escherichia coli lac* repressor *in vivo*. *Proc. Nat. Acad. Sci. U.S.A.*, **74**, 4228-4232.
- KEPES A., 1963. Kinetics of induced enzyme synthesis. Determination of the mean life of galactosidase specific messenger RNA. *Biochim. Biophys. Acta*, **76**, 293-309.
- KEPES A., 1978. Aspects moléculaires des fonctions membranaires. Masson, Paris.
- KNORRE W.A., 1968. Oscillations of the rate of synthesis of β -galactosidase in *Escherichia coli* ML 30 and ML 308. *Biochim. Biophys. Res. Comm.*, **31**, 812-817.
- KNORRE W.A., 1973. Oscillations in the epigenetic system : biophysical model of the β -galactosidase control system. In : *Biological and biochemical oscillators* (Ed. B. Chance, E. Pye, A. Ghosh, B. Hess), Academic Press, London, 449-457.
- MAJORS J., 1975 a. Initiation of *in vitro* mRNA synthesis from the wild-type *lac* promoter. *Proc. Nat. Acad. Sci. U.S.A.*, **72**, 4394-4398.
- MAJORS J., 1975 b. Specific binding of CAP factor to *lac* promoter DNA. *Nature*, **256**, 672-674.
- MALONEY P.C., WILSON T.H., 1978. Metabolic control of lactose entry in *E. coli*. *Biochem. Biophys. Acta*, **511**, 487-498.
- MANABE T., 1981. Mathematical expression for the rate of initiation and certain prediction on the non-competitive binding operon : the *E. coli* lactose operon. *J. Theor. Biol.*, **89**, 271-302.
- MANDECKI W., 1979. A kinetic model for interaction of regulatory protein and RNA polymerase with the control region of the *lac* operon of *E. coli*. *J. Theor. Biol.*, **81**, 105-122.
- MONOD J., WYMAN J., CHANGEUX J.P., 1965. On the nature of allosteric transitions : a plausible model. *J. Mol. Biol.*, **12**, 88-118.
- PERLMAN R., CHEN B., DE COMBRUGGHE B., EMMER M., GOTTESMAN M., VARMUS H., PASTAN I., 1970. The regulation of *lac* operon transcription by cyclic adenosine 3',5'-monophosphate. *Cold Spring Harbor Symp. Quant. Biol.*, **35**, 419-423.
- PETERKOVSKY A., 1977. Regulation of *Escherichia coli* adenylate cyclase by phosphorylation-dephosphorylation. *Trends Biochem. Sci.*, **2**, 12-14.
- REZNIKOFF W.S., 1976. Formation of the RNA polymerase *lac* promoter open complex. In : *RNA polymerase* (ed. R. Losick and M. Chamberlin), Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 441-455.
- REZNIKOFF W.S., ABELSON J.N., 1978. The *lac* promoter. In : *The operon* (ed. J.H. Miller and N.S. Reznikoff), Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 221-243.
- RIGGS A.D., SUZUKI H., BOURGEOIS S., 1970 a. *Lac* repressor-operator interaction. I - Equilibrium studies. *J. Mol. Biol.*, **48**, 67-83.
- RIGGS A.D., NEWBY R.F., BOURGEOIS S., 1970 b. *Lac* repressor-operator interaction. II - Effect of galactosides and other ligands. *J. Mol. Biol.*, **51**, 303-314.
- RIGGS A.D., BOURGEOIS S., COHN M., 1970 c. *Lac* repressor-operator interaction. III - Kinetic studies. *J. Mol. Biol.*, **53**, 401-417.
- ROSEN R., 1972. Some relational cell models : the metabolism-repair systems. In : *Foundations of mathematical biology* (R. Rosen, ed.), Academic Press, vol. II, 217-253.
- SANGLIER M., NICOLIS G., 1976. Sustained oscillations and threshold phenomena in an operon control circuit. *Biophysical Chemistry*, **4**, 113-121.
- SEL'KOV, 1968. Self oscillations in glycolysis 1. A simple kinetic model. *European J. Biochem.*, **4**, 79-86.

- TAKAHASHI M., BLAZY B., BAUDRAS A., 1979. Non-specific interactions of CRP from *E. coli* with native and denaturated DNAs : control of binding by cAMP and cGMP and by cation concentration. *Nucl. Acids Res.*, **7**, 1699-1712.
- TAKAHASHI M., 1979. Recherche des bases moléculaires de la répression catabolique chez *E. coli* : étude des interactions cAMP-dépendantes de la CRP, facteur positif de régulation de la transcription. Thèse 3^e cycle, Université Paul Sabatier, Toulouse.
- THOMAS R., 1973. Boolean formalization of genetic control circuits. *J. Theor. Biol.*, **42**, 563-585.
- THOMAS R. (ed), 1979. Kinetic logic. *Lecture notes in biomathematics*, vol. **29**, Springer Verlag, Berlin, Heidelberg, New York.
- ULLMANN A., JOSEPH E., DANCHIN A. Cyclic AMP as a modulator of polarity in polycistronic transcriptional units. *Proc. Nat. Acad. Sci. U.S.A.*, **76**, 3194-3197.
- WATSON J.D., 1976. Molecular biology of the gene. 3rd. ed., Benjamin, Menlo Park, California.
- WEST I.C., STEIN W.D. The kinetics of induction of β -galactoside permease. *Biochem. Biophys. Acta*, **308**, 161-167.
- WRIGHT J.K., RIELDE I., OVERATH P., 1981. Lactose carrier protein of *Escherichia coli* : interaction with galactosides and protons. *Biochemistry*, **20**, 6404-6415.