KINETICS OF
RNA POLYMERASE β SUBUNIT SYNTHESIS
AND
ACID END PRODUCT TRANSPORT
IN
Escherichia coli

Thesis by

Douglas D. Axe

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1990

(Submitted May 24, 1990)

He is no fool who gives what he cannot keep to gain what he cannot lose.

James Elliot

ACKNOWLEDGEMENTS

I am very grateful to a number of people who have had a part in making my experience at Caltech both rewarding and enjoyable. My first thanks go to my thesis advisor, Dr. Jay Bailey, for encouraging new approaches and ideas and for providing ample resources to fuel them. The lattitude that Jay has granted me has contributed immeasurably to my development as a researcher.

The Bailey clan has made its mark on me as well. Though we eventually head in different directions, we remain comrades. Thanks to Roger Hart and Mike Kosinski for making 336 a happening lab and for the use of their refrigerator, to Neilay Dedhia for kindly proofreading portions of this thesis, to Peter Licari for loaning me, over a long period of time, enough money to buy a Porsche (I opted for 187,304 Snickers bars instead), and to Lou Boschelli for stretching my vocabulary and spotting me on the bench (occasionally). Thanks are due also to Chaitan Khosla for sharing his molecular biology expertise with me, and for numerous lunchtime discussions covering a wide range of topics. Dane Wittrup is another good friend to whom I am grateful. Dane and I were kind of like Huck Finn and Tom Sawyer. Many a crazy thing we did together: late-night tennis ball wars on the third floor of Spalding, eating moldy agar from petri dishes, and pranks that will remain undisclosed for a period of time defined by the statute of limitations. Without such diversions, life would have been bleak indeed.

The first year of graduate school was a bit trying at times, but lasting friendships often emerge from tribulations. Justin Ip, Peter Campo, Dane Wittrup, and Hung Nguyen have earned a status in my mind that can only be accorded someone who has fought along side me in the same trench. Space pre-

vents me from mentioning all the folks whose friendship and/or assistance has meant something to me over these years... thank you! Explicit thanks to Kevin Leubke for more than one thoughtful discussion and for the part he played in making apartment 302 such a fun place to live.

As this thesis marks the end of my formal education, I cannot help but look back at the road I have trod, and having done that, I cannot help but recognize that, were it not for the loving support of two very dear people, I could not possibly have made the journey. If a record had been kept, it would surely show that I owe my parents more than I can ever repay. It wasn't kept; what was given was given out of love. The values they instilled in me from my early childhood have in large part shaped me into the person I am. Never pushing and ever encouraging, they gently guided me until I was tall enough to see where I was going. Thank you, Mom and Dad.

I leave Pasadena with much more than a degree. A young woman I met a year after my arrival has become my closest companion and my wife. Anita's unending generosity and earnest love have been a constant inspiration to me. Her buoyant spirit has kept me afloat when I might otherwise have sunk. Whether our courtship has prolonged my stay at Caltech somewhat, I cannot be sure, but I am certain that her friendship has made my stay immeasurably more pleasant. B.T.T.E.

Scientists are typically very careful about assigning credit for discoveries and achievements. If my work relies on the work of another, it would be presumptuous of me to ignore that fact. Lest I be guilty of this, I should express my gratitude to and dependence upon the One who made both the universe and the mind with which I seek to comprehend it. I have accomplished nothing of significance apart from Him, nor will I ever.

To Mom and Dad

ABSTRACT

An approach to modeling the regulation of synthesis of crucial bacterial proteins has been developed. The unique features of this approach are that it focuses on maintenance of a steady state rather than on transitions between steady states, and that stochastic fluctuations in the number of transcripts per cell are treated as the perturbation. It has been used to investigate various models of translational regulation of RNA polymerase β subunit synthesis. The simplest autogenous regulatory mechanism, binding of a single RNA polymerase molecule to the *rpoBC* mRNA, appears to provide inadequate control. A more sophisticated mechanism, sequential binding of multiple polymerase molecules in a cooperative manner, was shown to dramatically improve the control characteristics.

The interaction between RNA polymerase and *rpoBC* mRNA was examined experimentally. RNA polymerase was incubated with RNA that is identical to a small portion of the native *rpoBC* message. Gel mobility-shift assays were performed to detect complexes. The relative amount of RNA in different complexes was determined by radiolabeling the transcript. Data obtained in this way indicate that cooperative binding is occuring.

^{31}P NMR studies of intact *E. coli* cells suggested a difference in membrane function between a plasmid-containing strain and the plasmid-free host. Similar ^{31}P NMR experiments were complemented with ^{13}C NMR experiments to examine the transport of lactate and acetate through the cytoplasmic membrane during anaerobic glycolysis. Methods were developed to measure cytoplasmic and extracytoplasmic solution volumes and intra- and extracellular acid concentrations using ^{13}C NMR. The results demonstrated significant differences

in the transport of the two acids. Acetate was determined to permeate the membrane at comparable rates in the dissociated and undissociated forms. The mode of lactate transport in cells that are actively glycolyzing was found to be different from that of cells that have exhausted their supply of glucose. Lactate thus appears to be transported by a system that is sensitive to some indicator of glycolytic activity. It also appears to diffuse across the membrane in both forms.

A kinetic approach was used to deduce constraints on the unidirectional fluxes for a general protein-mediated, ATP-independent transport process. The conclusion was that under certain circumstances, the Ussing-Teorell flux ratio equation applies to protein-mediated transport. From the analysis, an expression was derived for the driving force of such a transport process, and the relationship between this and the net flux is discussed.

TABLE OF CONTENTS

CHAPTER I.

A THEORETICAL APPROACH TO STUDYING THE REGULATION OF SYNTHESIS OF CRUCIAL PROTEINS IN BACTERIA

The material in this chapter has been submitted for publication in the *Journal of Theoretical Biology.*

Summary

Mathematical modeling is an important step in evaluating hypotheses concerning mechanisms by which gene expression is regulated. The complexity of even the simplest cells, however, precludes rigorous mathematical description. While simplifying approximations always lead to a discrepancy between model characteristics and the properties of the real system, sufficient agreement for obtaining useful results can often be achieved. Modeling the regulation of synthesis of crucial proteins poses particular problems in this regard. A small displacement in the concentration of such a protein causes a shift in the growth properties of the entire cell. Consequently, transient behavior of a crucial-protein concentration modeled in the context of an unperturbed cell may have no discernible relationship with transients in an actual cell.

Here we present a method for making quantitative comparisons of regulatory mechanisms hypothesized as governing the synthesis of a crucial protein in a bacterial cell. One of the distinguishing features of the approach is its focus on maintenance of a steady state rather than on transitions between steady states. Stochastic fluctuation of mRNA concentration is treated as an intrinsic perturbation. The model of gene expression developed herein accounts for such details as the time required for polypeptide chain elongation, the mRNA length distribution, and the distribution of endonucleolytic cleavage sites along the length of the transcript. Specific models of transcriptional and translational regulation are inserted into this expression framework in order to determine their control responses. The framework is of a sufficiently general form that the method should be applicable to expression of any gene.

1. Introduction

Since the landmark work of François Jacob and Jacques Monod on the *lac* operon (1961), considerable progress has been made toward elucidating the mechanisms that govern the regulation of gene expression in prokaryotes. One might reasonably expect that better understanding of the regulatory mechanisms associated with individual operons or genes, coupled with detailed knowledge of biochemistry, will ultimately lead to an integrated, mechanistic understanding of the processes that constitute the basis of life in its simplest forms. Initial explanation of phenomena occurring in highly complex systems inevitably requires a reductionistic approach; the focus shifts from understanding the *whole* to understanding relatively simple *parts* thereof. Eventually, however, an integrated picture must be constucted from the fragmented one. The shortcomings of reductionism become apparent at this stage.

In many instances it is very reasonable to separate a part from the whole and to treat it as a distinct subsystem. Concerning gene expression, one would expect that synthesis of a macromolecule that does not perform an essential function (under given growth conditions) can be treated as a distinct subprocess, growth of the entire cell being the complete process. By considering cells in exponential growth, one can then eliminate the need to describe transient behavior outside of the subsystem. Thus, a mathematical description of the entire cell is not needed to describe transient behavior of the subsystem. This situation is illustrated schematically in Figure 1.

While the conditions under which this approach is valid have not yet been precisely expounded, one can identify numerous loci that are very promising candidates. For example, in the absence of a β-lactam antibiotic, β-lactamase does

not perform any metabolically significant function. Hence, provided that the rate of synthesis of β-lactamase is modest, one would not expect fluctuations in this rate to have a significant impact on any other process occurring in the cell (extreme overexpresssion of *any* protein will obviously have a substantial impact). As another example, consider expression of the gene coding for the *lac* repressor, *lacI*, in the absence of lactose and glucose. In this case, the *lac* repressor *does* have a metabolically significant function; it prevents the synthesis of proteins that are useful only in catabolizing lactose. The affinity of the repressor for its operator site is so high, however, that fluctuations in the concentration of repressor will have no significant effect, provided that the level of repressor remains above some low basal value. As these examples illustrate, treating gene expression as a distinct subprocess is qualitatively justified when variation of the expression level over the range of interest has no significant effect on other cellular processes.

In cases where this approach is valid, mechanisms regulating expression of the gene of interest can be modeled in a straightforward manner. The problem is that some of the most interesting genetic control problems involve genes for which the subsystem approach is invalid. In fact, one might attribute a portion of the interest in these genes simply to the fact that they influence many other cellular processes. We were confronted with this problem in the course of studying the synthesis of the major subunits of RNA polymerase in *E. coli* (Axe and Bailey, 1990). Needless to say, the cellular level of RNA polymerase strongly influences gene expression, thereby influencing virtually all aspects of metabolism. This suggests another reason for interest in the synthesis of crucial proteins; because they strongly influence cellular activities, their synthesis is tightly regulated, often by rather elaborate mechanisms.

In this work, we develop an approach to modeling the expression of genes

coding for these proteins. The success of a modeling effort is often judged by the magnitude of the discrepancy between model behavior and the observed behavior of the system being modeled. The very nature of the problem we address here precludes such evaluation. There is no way to precisely predict the behavior of a poorly understood, complex system unless, as just discussed, that behavior is confined to a relatively simple subsystem. The objective of this work, then, is not to facilitate simulation of the cellular response to a shift in the level of a crucial protein, but instead to construct a flexible theoretical framework that can be used to make *comparative* evaluations of specific transcriptional and translational control models.

While in some cases (Schlosser and Bailey, 1990) it may be possible to circumvent the problems inherent to the subsystem approach by experimentally characterizing the surrounding system (represented by the "black frame" in Figure 1), there is no analogous theoretical method. Our approach has instead been to devise a context in which the error resulting from the subsystem approach is reduced and, more importantly, the nature of the error is characterizable. The transient behavior computed in this way is a precise function of the control properties of the subsystem of interest and, hence, can be used to quantitatively compare the efficacy of different control strategies. The computed transients may not, however, accurately represent transients that might be observed in living cells. They represent a bound on the actual behavior that is more or less optimistic (from a control viewpoint), depending on how realistic it is to treat the subsystem as a distinct process.

2. Approach

The source of the error inherent in the separate subsystem approach is sim-

ply that perturbations are considered to be confined to the subsystem, whereas in reality they would propagate into the surrounding system. As already noted, the reason for the separate subsystem approach in the first place is that the surrounding system is too complex to model mechanistically. When the subsystem encompasses synthesis of a crucial protein, a shift from one subsystem steady state to a new subsystem steady state necessarily implies a shift in the *surrounding* system from one steady state to another. This can actually serve as the definition of a crucial protein. Concerning such a transition between steady states, neglecting propagated perturbations not only results in error, but it is difficult or impossible to assess the magnitude or even the direction of the error. This is a consequence of the fact that the error depends heavily on properties of the new steady state that would actually be attained by the surrounding system.

In order to eliminate this problem, we will consider maintenance of a single steady state rather than transitions between steady states. To study control dynamics we must, of course, consider perturbations from the steady state, but by restricting ourselves to cases where the initial and ultimate states are identical, we can circumvent the need to characterize a new exponential growth configuration. One would expect a model that rigorously accounts for propagation of a local disturbance to predict a slower return to the steady state than a model that confines the disturbance to the subsystem in which it arose. The discrepancy between the two would depend on the magnitude and nature of the disturbance. Hence, the approach taken here will yield a control response that represents an optimistic bound on the actual response.

Because we are confining transients to the subsystem of interest, we must identify an intrinsic perturbation within the subsystem. We suggest that stochastic fluctuation in the number of mRNA copies present in a cell constitutes the domi-

nant intrinsic perturbation. On average, only a few copies of a given message are present in a cell at a time (Neidhardt, 1987). This is due not to a low rate of transcription per se, but rather to a high rate of message degradation that results in a relatively low *net* rate of message synthesis. If, as would be expected, mRNA synthesis and degradation can be described as Poisson processes on the level of a single cell, then the magnitude of random fluctuation in mRNA copies is substantial. In the absense of transcriptional regulation, the distribution of mRNA copies among cells in a population is described by the Poisson formula:

$$P(n) = \frac{[\tau]^n}{n!} e^{-[\tau]}$$

where $[\tau]$ is the mean concentration of transcript τ (copies per cell), and $P(n)$ is the fraction of the population having exactly n copies of τ. Figure 2 illustrates the population heterogeneity that is expected when $[\tau]$ has a value of 5 copies per cell. Less than 18% of the cells have an mRNA content corresponding to the population mean; 1.8% have 10 copies, and 3.4% have only one.

It is important to recognize that while properties of a culture undergoing balanced exponential growth are time invariant (on a per biomass basis), properties of the individual cells comprising the population are in continual flux. The cell cycle certainly contributes to this, but, as we have just seen, random fluctuations alone account for extreme population heterogeneity. For balanced exponential growth to be possible at all, it must be true that cells that have been significantly displaced from the mean configuration tend to recover from the displacement. It seems likely that a very small fraction of the population might suffer lethal displacements, but given a distribution as broad as that depicted in Figure 2, successful recovery from displacements must be a routine part of bacterial life.

This forms the basis of the approach that is developed here. The model equa-

tions to be used describe mean properties of a large population, but the perturbation introduced in the initial conditions is based upon random fluctuation on the cellular scale. A hypothetical experimental analogy to the approach would be the following. Starting with a large culture of cells in balanced exponential growth, we somehow identify a subpopulation consisting of all cells that have n copies of transcript τ, where n is significantly different from the mean value for the entire population. We then monitor this subpopulation by frequently recording the number of τ molecules and π molecules in each cell, where π designates the protein resulting from translation of τ. Finally, we plot the average τ and π levels for the subpopulation as functions of time. If such an experiment could be performed, we would expect the mean π level to be displaced as a result of the displacement in τ. We would also expect the mean levels of both τ and π to eventually return to the values for the entire population. Plots of the transient behavior would characterize the performance of regulatory mechanisms governing synthesis of π, stronger regulation resulting in a smaller displacement in the level of π and in more rapid recovery.

3. Mathematical Description of Gene Expression

Implementation of the theoretical equivalent to the experiment described above necessitates a mathematical description of gene expression. The basic material balance equation for any molecular species x that is not transported across the cytoplasmic membrane is:

$$\frac{d[x]}{dt} = R_x - (\mu + \delta_x)[x] \tag{1}$$

where μ is the specific growth rate (min^{-1}), δ_x is the rate constant for degradation of species x (min^{-1}), R_x is the rate of synthesis of species x (M/min), and brackets

denote a concentration in moles per liter of cytoplasm. All terms in this equation refer to mean values for a large, genetically homogeneous bacterial population.

Gene expression can be modeled by using equations of the above form to describe transcript and protein concentrations. We will first develop the protein balance equation. Here, π denotes mature protein, and τ denotes mRNA that has not been cleaved within the coding region (henceforth referred to as *viable* mRNA). It would be possible to construct a rather detailed model of translation that includes factors like ribosome, aminoacyl-tRNA, and GTP concentrations, temperature, pH, and so forth, but this is unnecessary because our approach involves fixing all parameters outside the subsystem to their balanced-growth values. The rate at which polypeptide chains are initiated is proportional to the concentration of viable transcripts, all of the above parameters being fixed. Since it is possible for a transcript to be cleaved on the 3′ side of a translating ribosome, a polypeptide initiation event does not always yield a mature protein. Therefore the rate of π synthesis must include a factor representing the probability of successful translation. Furthermore, the time required to complete a polypeptide chain is sufficiently long that it cannot generally be ignored. The rate of π synthesis thus depends on *previous* cell conditions rather than on current conditions.

Incorporating these factors into an equation of the form (1), we obtain:

$$\frac{d[\pi]}{dt} = \left[\theta_2 \Phi_2 \xi[\tau]\right]_{t-\ell} - (\mu + \delta_\pi)[\pi] \tag{2}$$

where ξ is the probability that an initiated polypeptide chain will be successfully completed, and ℓ is the time required to translate the message. While the value of the constant Φ_2 depends on numerous factors that influence the rate of translation initiation, it can be interpreted as a measure of the affinity of the Shine-Dalgarno sequence for ribosomes. That is, if expression of more than one gene were considered

in the context of the same balanced-growth state, the respective Φ_2 values would be indicative of the ribosome affinities. If π influences translation in a manner that is sufficiently direct as to permit mathematical description, Φ_2 should be replaced by the appropriate expression involving $[\pi]$. The specific control mechanism enters the equation through the translational control function, θ_2. The mathematical expression for θ_2 may include any concentrations that are to be described by the subsystem model. During exponential growth, many proteins in *E. coli* have half-lives much longer than a generation (Mosteller *et al.*, 1980), so the degradation term, δ_π, can often be neglected. The fact that the rate of π synthesis depends on the probability of cleavage of the template carries significant implications for our mathematical description of $[\tau]$. These will be discussed after considering the mechanisms of mRNA degradation.

Message RNA is relatively short-lived in bacteria. A half-life of two or three minutes is typical (Pedersen and Reeh, 1978). While there is still much to be learned about this process, several important features have emerged (see, for example, Belasco and Higgins, 1988). The enzymes responsible for degrading mRNA, ribonucleases, can be divided into two groups. Exonucleases catalyze processive degradation of mRNA from the 3′ end. Stem-loop structures are known to be an effective barrier to exonuclease attack. Because stem-loop structures also play an important role in transcript termination, they are frequently present at the 3′ end of transcripts, though they often occur between coding regions in the transcript interior as well. Endonucleases, on the other hand, cleave mRNA internally in a manner that is sequence dependent, though probably not highly specific. Endonucleolytic cleavage appears to be the rate-determining step in message degradation in a number of cases (Belasco and Higgins, 1988). The explanation is presumably that stem-loop structures prevent exonucleolytic digestion from proceeding into a

coding region until endonucleolytic cleavage forms a new 3′ end on the 5′ side of the stem-loop structure.

The susceptibility of a viable transcript of a particular type to endonucleolytic cleavage is a function of the extent to which transcription has progressed. As transcription proceeds, new cleavage sites are formed in the nascent chain. The fractional survival of the template ahead of a translating ribosome is thus a function of the message length at the time of initiation. It follows that ξ is a function of the message length distribution. The mathematical implication of this is that τ must be treated as a series of discrete length classes, $\tau_1, \tau_2, \ldots, \tau_n$, rather than as a single variable. The value of n determines the number of equal-sized zones into which the full transcript length is divided. Endonuclease cleavage patterns for a specific transcript can be mapped onto these zones by specifying an individual degradative rate constant, $\delta_{\tau i}$, for each zone. Larger values of n thus permit finer "resolution" of cleavage sites along the transcript. A value of ten or twenty would probably be adequate for most purposes.

Consider an open reading frame L bases in length that is divided into n zones of length L/n, the first zone being at the 5′ end. The i^{th} length class, τ_i, will be defined as the set of viable transcripts with 3′ (elongating) ends that map onto the i^{th} zone. An additional class, τ_{n+1}, will include all full-length transcripts. For purely computational reasons, we will consider the endonuclease recognition sites to be situated near the 3′ ends of the zones. An equation of the form (1) can be written for each length class. The rate of transcript initiation appears only in the equation for the first class. If π is not directly involved in transcript initiation, the equation governing $[\tau_1]$ is:

$$\frac{d[\tau_1]}{dt} = \theta_1 \Phi_1 - \mu[\tau_1]. \tag{3}$$

Transcripts in the first length class are not susceptible to cleavage because the first cleavage site is at the 3′ end of the first zone. Φ_1 and θ_1 are the transcriptional analogs of Φ_2 and θ_2 of equation (2); Φ_1 is proportional to the affinity of the promoter for RNA polymerase, and θ_1 is the control function used to model specific transcriptional regulatory mechanisms. Equations for $\tau_2, \ldots, \tau_{n+1}$ are of the form:

$$\frac{d[\tau_i]}{dt} = -\left(\mu + \sum_{j=1}^{i-1} \delta_{\tau j}\right)[\tau_i]. \tag{4}$$

A few remarks will be made here regarding the method used to solve equations (3) and (4). First of all, the discrete zones and length classes are purely conceptual; in reality, the elongating ends of transcripts are distributed more or less uniformly over the coding region, irrespective of zone boundaries. This means that the constituency of a length class changes continually as younger transcripts enter and older transcripts exit. There is no provision for transcript elongation in equations (3) and (4). Equation (4) describes the effects of degradation and dilution on $[\tau_i]$, but it says nothing about the flux of transcripts from class τ_{i-1} into class τ_i or the flux from class τ_i into class τ_{i+1}. Transcript elongation will be represented in the method of solution rather than in the equations themselves.

The length-class distribution at $t = 0$ constitutes a grouping of transcripts. The new distribution at $t = \Delta t$ (where Δt is defined as the time required for transcription to proceed through a single zone) can be obtained by computing the evolution of each of these groups. For $0 < t < \Delta t$ the groups are staggered with respect to current length classes, but at $t = \Delta t$ the group that originally corresponded to the τ_i class now corresponds to the τ_{i+1} class. The equation describing the evolution of the group originally defined by τ_i $(i = 1, \ldots, n)$ over the interval Δt is thus a linear combination of the equations for length classes τ_i and τ_{i+1}:

$$E = E_i + \frac{t}{\Delta t}(E_{i+1} - E_i)$$

where E_i and E_{i+1} are the appropriate forms of equation (4), and E represents the equation to be solved. The genesis of the *next* τ_1 length class (which can be designated τ_0 until $t = \Delta t$) is described by equation (3) with $[\tau_0]$ replacing $[\tau_1]$ ($[\tau_0] = 0$ at $t = 0$). The aging of length class τ_{n+1} is described by equation (4) with $i = n + 1$. After solving all $n + 2$ equations over an interval of length Δt, concentration values are passed to the succeeding length class as follows:

$$[\tau_i] \rightarrow [\tau_{i+1}] \qquad \text{for } i = 0, \ldots, n-1$$

$$[\tau_n] + [\tau_{n+1}] \rightarrow [\tau_{n+1}]$$

(symbols on the left referring to values that were just computed), and the process is repeated.

Rewriting equation (2) in terms of transcript length classes, we obtain:

$$\frac{d[\pi]}{dt} = \left[\theta_2 \Phi_2 \sum_{i=1}^{n+1} \xi_i [\tau_i]\right]_{t-\ell} - (\mu + \delta_\pi)[\pi] \tag{5}$$

where ξ_i is the probability that translation initiated on a viable transcript in class τ_i will produce a mature protein, π. A mathematical description of ξ_i is developed in the Appendix. Given control functions θ_1 and θ_2, equations (3) and (5) together with n equations of the form (4) constitute a complete description of gene expression. Depending on the nature of the regulatory mechanism, it may not be possible to express θ_1 or θ_2 as an explicit function of the variables appearing in these equations. In that case additional equations would be required. A example of this sort is presented in section 5 (Model Application).

An important point regarding the differential equations presented here should be made before proceeding. All variables in these equations represent mean values for a large population. In light of the very broad distribution expected for $[\tau]$, it is reasonable to ask whether it is legitimate to describe reaction rates in a heterogeneous population in terms of mean concentrations. The first term on the right-hand

side of equation (2) involves only one concentration, $[\tau]$, explicitly (in the present discussion we are referring to equation (2) instead of equation (5) for the sake of clarity only; the point being made is equally applicable to equation (5)). However, as noted previously, Φ_2 implicitly depends on several factors including the concentration of free ribosomes, $[R_f]$. If we consider Φ_2 to be simply proportional to $[R_f]$, then an average rate is represented in equation (2) as a product of two average concentrations. Accurate representation of mass-action kinetics would require that this be equivalent to the average of the product of actual cellular concentrations. This is rigorously true only if R_f and τ are independently distributed (i.e., only if the cellular deviation of τ concentration from the mean value, $[\tau]$, does not correlate with the cellular deviation of R_f concentration from the mean value, $[R_f]$).

In *E. coli* growing with a forty-minute doubling time, there are roughly five thousand nontranslating ribosomes per cell (Bremmer and Dennis, 1987). Random fluctuation would not lead to substantial dispersion of a concentration this high. We would expect 95% of the cells in a population to have ribosome concentrations within 3% of the mean. It is quite possible, however, that both τ and R_f concentrations exhibit cell cycle dependence. If this is the case, the above condition is not met, and equation (2) would introduce error. The situation with the first term on the right-hand side of equation (3) is similar in that a product of averages is assumed to be equivalent to the average of the product. Here Φ_1 depends on the concentration of the gene of interest, which certainly varies with the cell cycle, and on the concentration of free RNA polymerase. If the latter exhibits cell-cycle dependence, equation (3) would also introduce error. Still, given that our objective is to compare control mechanisms rather than to simulate cell behavior, equations (3)-(5) constitute a reasonable framework. In summary, our mathematical description assumes that perturbations in the subsystem of interest are not propagated into

surrounding systems and that the surrounding system is in a steady state that does not exhibit cell-cycle periodicity. Both assumptions are idealizations that deviate to varying degrees from real cell behavior, but neither should greatly interfere with the intended application.

4. Modeling the Perturbation

The rationale behind treating random fluctuation of transcript concentration as the perturbation to the control system has been discussed. Here we describe the implementation of this in the model context that has now been developed. Note, first of all, that while the above model equations describe *mean* properties of a population, the perturbation we are proposing involves a specific feature of the distribution of cell states within the population. One way to characterize the perturbation would be to develop a population model that describes the distribution of transcripts, τ, among cells. A simpler appoach is equally suitable for our purposes.

We would like the perturbation to be specified in the initial conditions. That is, we want a method for determining non-steady-state values for model variables such that together they represent a perturbed state. Furthermore, a rational means of determining *comparably* perturbed states for different control models must be employed because our aim is to compare the responses of different models. Where extrinsic perturbations are concerned, it is reasonable to view *equal* ones as comparable. For intrinsic perturbations, on the other hand, it is more reasonable to view *equally likely* ones as comparable. Given that transcript initiation and degradation are Poisson processes, we can determine the likelihood that the actual mean rate of τ synthesis (or degradation) for an entire population would deviate by a specified amount from the expected mean rate. This probability is a function only of the

actual and expected rates and is determined by the Poisson relation. By specifying a probability, one can thus determine an actual (perturbed) rate for any expected rate, the probability being a measure of the degree of perturbation.

While the Poisson formula is theoretically applicable, it is inconvenient for this purpose because the terms involved become extremely large or extremely small when the expected rate is larger than about 50. Since the actual rate is restricted to integral values, the Poisson formula cannot be applied to populations that are large enough that the probability resembles a continuous function. This problem is easily overcome by using Stirling's approximation to $N!$ for large values of N:

$$\ln N! \approx (N + \tfrac{1}{2})\ln N - N + \tfrac{1}{2}\ln 2\pi. \tag{6}$$

By taking the natural logarithm of the Poisson formula and using equation (6) to substitute for the factorial term, we obtain:

$$\ln P \approx (R_a - R_e) + R_a(\ln R_e - \ln R_a) - \tfrac{1}{2}\ln (2\pi R_a) \tag{7}$$

where R_e is the expected number of occurrences of some event during a specified time interval, and P is the probability that the actual number of occurrences will be R_a. Equation (7) allows simple and accurate evaluation of $\ln P$ for very large values of R_a.

The equations developed in the previous section enable us to determine the expected rates of transcript initiation and of τ_i degradation for $i = 1, \ldots, n, n+1$. By specifying a value for P and applying equation (7), one can determine equally probable (or improbable) deviated rates for each of the expected rates. To determine initial conditions that characterize cells with a surplus of mRNA, one solves the governing differential equations (using the normal steady state for initial conditions) over a brief interval with expected rates replaced by actual rates, such that

the deviation in initiation rate is positive and deviations in degradation rates are negative. The extent of deviation obtained depends on the value chosen for P, with smaller values yielding more rapid deviation from the steady state. If the same value of P is used to compute deviated states for several different regulatory models, the states obtained will differ, but they will represent equally improbable deviations.

5. Model Application

To illustrate the application of this approach, we will consider simple cases of autogenous transcriptional and translational regulation of the synthesis of a hypothetical protein. An actual problem, regulation of *rpoB* expression, is treated in detail in the accompanying paper (Axe and Bailey, 1990). Gene regulation is said to be *autogenous* if the products of a gene are directly involved in regulating expression of that gene. Transcriptional control frequently involves repression or induction of an upstream promoter. The simplest model of autogenous transcriptional control consists of binding of a single protein molecule to a site that prevents binding of RNA polymerase to the promoter responsible for synthesizing the protein. In this case the transcriptional control function, θ_1, is defined as the fraction of promoters that are not blocked by protein binding. If the rate constants for association and dissociation are sufficiently high, the bound and unbound forms of the promoter can be described as being in equilibrium, and θ_1 is given by:

$$\theta_1 = \frac{1}{K[\pi] + 1} \tag{8}$$

where K is the equilibrium binding constant (M^{-1}).

Analogously, translational control can be achieved via binding of π to the 5′ end of the mRNA in such a way as to prevent translation. The rate at which transcripts

are produced is typically much greater than the rate at which new gene copies are produced, so rapid equilibration of transcript complexes is not nearly as feasible as rapid equilibration of promoter complexes. The translational control function, θ_2, is defined as the fraction of viable transcripts that are not bound, but in the absense of equilibrium we cannot express θ_2 as a function of π only. We instead replace equations (3) and (4) with:

$$\frac{d[\tau_1]_0}{dt} = \theta_1\Phi_1 + r[\tau_1]_1 - (k[\pi] + \mu)[\tau_1]_0 \tag{9}$$

$$\frac{d[\tau_i]_0}{dt} = r[\tau_i]_1 - \left(k[\pi] + \mu + \sum_{j=1}^{i-1}\delta_{\tau j}\right)[\tau_i]_0 \tag{10}$$

and

$$\frac{d[\tau_i]_1}{dt} = k[\pi][\tau_i]_0 - \left(r + \mu + \sum_{j=0}^{i-1}\delta_{\tau j}\right)[\tau_i]_1 \tag{11}$$

where $[\tau_i]_0$ and $[\tau_i]_1$ are, respectively, the concentrations of unbound and bound viable message in length class i, and k ($\mathrm{M}^{-1}\mathrm{min}^{-1}$) and r (min^{-1}) are the rate constants for association and dissociation. Equation (11) applies to all length classes (the value of $\delta_{\tau 0}$ is taken to be zero), whereas equation (10) applies to all but the first.

The rate of π degradation will be taken as zero, so the equation governing $[\pi]$ becomes:

$$\frac{d[\pi]}{dt} = \left[\Phi_2 \sum_{i=1}^{n+1} \xi_i[\tau_i]_0\right]_{t-\ell} - \mu[\pi] \tag{12}$$

after dropping θ_2 from equation (5) and replacing $[\tau_i]$ with $[\tau_i]_0$. The net effect of the changes incorporated into equations (9)-(12) is the replacement of the θ_2 function with n differential equations describing binding of π to its mRNA. These equations will be solved for four control configurations: no regulation ($K = k = 0$), transcriptional regulation only ($k = 0$), translational regulation only ($K = 0$), and combined regulation.

Values of physical parameters used for this illustration (Table 1) do not represent any particular gene but are typical values for *E. coli*. We have used $n = 20$ with the same $\delta_{\tau i}$ value for all twenty sites. This would be a reasonable approach to use for an actual transcript if nothing were known regarding cleavage sites. Transients have been computed for an initial positively displaced mRNA concentration corresponding to about ten copies per cell (five copies at steady state) and a negatively displaced concentration corresponding to about one copy per cell. While these represent significant deviations from the steady state, they are not highly improbable; roughly 7% of the cells in an exponentially growing culture are displaced at least this severely (see Figure 2). Over the course of a generation, approximately 40% of the cells will experience a displacement of this magitude.

In the absence of regulation, recovery is rather slow (Figure 3). The positive displacement in $[\tau]$ leads to an 8% increase in $[\pi]$. There is a substantial lag between the times that $[\tau]$ and $[\pi]$ reach their maximum values (the maximum $[\tau]$ value occurs at $t = 0$ and is outside the the range of the plot; $[\pi]$ reaches a maximum after roughly nine minutes of recovery). This is due in part to the time required for translation, ℓ, and in part to the relative turnover rates of protein and message. After two generations of recovery, the $[\pi]$ displacement is reduced to 2.5%. Here the introductory remarks should be reemphasized. We have operationally defined a crucial protein as a protein that, when displaced significantly in steady state concentration, leads to a significant displacement in the steady state growth configuration. While $[\pi]^*$ is not displaced here, the fact that $[\pi]$ is significantly displaced for more than a generation suggests strongly that the cells would experience a sizable global perturbation. Consequently, we would expect the actual response to be worse than the depicted response.

Computed responses for cells with transcriptional regulation only (Figure 4)

and translational regulation only (Figure 5) are much better than the zero-control responses. The $[\pi]/[\pi]^*$ trajectories in these two cases are very similar, indicating comparable effectiveness of the two control strategies. While the maximum $[\pi]$ displacements are nearly as great as those in the absence of regulation, after two generations of recovery they are reduced to less than 0.8%. As would be expected, transcriptional control leads to a major change in the mRNA response. In Figure 4B, $[\tau]$ overshoots $[\tau]^*$, causing a more rapid $[\pi]$ recovery.

Further improvement is observed when transcriptional and translational regulation are combined (Figure 6). Here $[\pi]$ recovers to within 0.2% of $[\pi]^*$ in two generations. Only two-thirds of a generation is required to reach the displacement observed after two generations of recovery without regulation. This level of control would certainly be adequate for synthesis of some bacterial proteins, though it is probably inadequate for others. More elaborate regulatory mechanisms can be devised to yield significantly improved responses (Axe and Bailey, 1990).

6. Discussion

Bacterial cells employ elaborate regulatory mechanisms to control the synthesis of proteins that are crucial to cell functions. A detailed picture of these mechanisms, though, does not generally emerge until long after the gross features are discovered. Modeling is an essential step in bridging the gap between a vague, qualitative picture and a precise, mathematical one. This work presents and demonstrates a methodology for making quantitative comparisons of regulatory mechanisms hypothesized as governing the synthesis of crucial proteins. The most salient features of the approach used are: first, that it focuses on maintenance of a steady state rather than on transitions between steady states, and second, that it employs random fluctua-

tion of mRNA concentration as an intrinsic perturbation to the steady state. The central role of mRNA in this approach has necessitated the development of a framework model that accounts for such details as a transcript length distribution and the distribution of endonucleolytic cleavage sites along the length of a transcript.

It should be noted that while this approach enables one to compare alternate models quantitatively, it does not specifically provide a means of judging the suitability of a given mechanism for controling expression of a given gene. The difficulty lies in quantifying the cellular sensitivity to fluctuations in the concentration of a crucial protein. Proteins for which an *a priori* estimate can be made probably constitute a rare minority. Experimental estimation of the distribution of a protein concentration over a population may be feasible in some instances. Still, in view of the fact that model responses will err on the optimistic side, it is safer to deem a hypothesized regulatory process unsuitable based on its response than to deem one suitable.

Acknowledgement:

This work was supported by the Energy Conversion and Utilization Technologies (ECUT) Program of the U.S. Department of Energy.

References

Axe, D. D. and Bailey, J. E. (1990). Translational regulation of *rpoB* expression in *Escherichia coli*: a theoretical study. *J. Theor. Biol.*

Belasco, J. G. and Higgins, C. F. (1988). Mechanisms of mRNA decay in bacteria: a perspective. *Gene* **72**, 15.

Bremer, H. and Dennis, P. P. (1987). Modulation of chemical composition and other parameters of the cell by growth rate. In: *Escherichia coli and Salmonella typhimurium.* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magsanic, B., Schaechter, M., and Umbarger, H. E., eds.) pp. 1527-1542. Washington, DC: American Society for Microbiology.

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Molec. Biol.* **3**, 318.

Mosteller, R. D., *et al.* (1980). Metabolism of individual proteins in exponentially growing *Escherichia coli. J. Biol. Chem.* **255**, 2524.

Neidhardt, F. C. (1987). Chemical composition of *Escherichia coli.* In: *Escherichia coli and Salmonella typhimurium.* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magsanic, B., Schaechter, M., and Umbarger, H. E., eds.) pp. 3-6. Washington, DC: American Society for Microbiology.

Pedersen, S. and Reeh, S. (1978). Functional mRNA half lives in *E. coli. Molec. Gen. Genet.* **166**, 329.

Schlosser, P. M. and Bailey, J. E. (1990). An integrated modelling-experimental strategy for the analysis of metabolic pathways. *J. Math. Biosci.*

Appendix

Evaluation of the Translational Success Rate, ξ_i

The rate at which mRNA is translated (bases per minute) very nearly matches the rate of RNA chain elongation at a doubling time of 40 minutes (Bremer and Dennis, 1987). Because calculation of ξ_i is simplified by assuming equal rates of transcription and translation, we will make that assumption here. Modification of the following equations to represent unequal transcription and translation rates, if necessary, is straightforward.

An open reading frame L bases in length is first divided into n equal-sized zones, the first zone at the 5′ end. For $i = 1, \ldots, n$, the i^{th} length class, τ_i, consists of all transcripts with elongating ends within the i^{th} zone (τ_{n+1} being the only class that has finished elongation). If translation is initiated on an elongating transcript, the length of template between the ribosome and RNA polymerase will remain constant until transcription is complete (since we are concerned only with the coding region, completion coincides with transcription of the stop codon). During this period, cleavage at any site in this target region will result in premature termination of translation. Consider the fate of polypeptide chains that are initiated on transcripts in the i^{th} length class at $t = 0$. Assuming that the lengths of transcripts in this class are uniformly distributed over the range of a zone, the rate at which target regions are cleaved is given by:

$$\frac{dN}{dt} = -\left(\frac{t\delta_{\tau i}}{\Delta t} + \sum_{j=1}^{i-1} \delta_{\tau j}\right) N \qquad (A1)$$

where $N(t)$ is the number of transcripts remaining at time t, and Δt is the time required to transcribe through a single zone. Equation (A1) applies during the interval $0 \leq t \leq \Delta t$.

By integrating equation (A1), we obtain the following expression for the fractional survival (or probability of survival) at $t = \Delta t$:

$$P_1 = \exp\left(-\Delta t\left(\frac{\delta_{\tau i}}{2} + \sum_{j=1}^{i-1}\delta_{\tau j}\right)\right). \qquad (A2)$$

The probability of survival from $t = (m-1)\Delta t$ to $t = m\Delta t$ is then:

$$P_2 = \exp\left(-\Delta t\left(\frac{\delta_{\tau(i+m-1)}}{2} + \sum_{j=m}^{i+m-2}\delta_{\tau j}\right)\right) \qquad (A3)$$

for $m = 1, 2, \ldots, n-i+1$. When $m = n-i+1$, transcription is in its final stage. The overall probability that the target regions will survive until transcription is complete is given by the product of the probabilities for each step:

$$P_3 = \exp\left(-\Delta t \sum_{m=1}^{n-i+1}\left(\frac{\delta_{\tau(i+m-1)}}{2} + \sum_{j=m}^{i+m-2}\delta_{\tau j}\right)\right). \qquad (A4)$$

This expression applies to all length classes except τ_{n+1}. For this class, transcription is already complete at $t = 0$, so $P_3 = 1$.

After transcription is complete, the target region shrinks from $i-1$ zones to zero as translation proceeds. The probability of surviving this phase is:

$$P_4 = \exp\left(-\Delta t \sum_{m=n-i+2}^{n}\left(\sum_{j=m}^{n}\delta_{\tau j}\right)\right). \qquad (A5)$$

This expression applies to all length classes except τ_1. For this class, completion of translation and completion transcription are simultaneous events, so $P_4 = 1$. Finally, the probabability that translation initiated on a τ_i transcript will yield a complete protein is given by the product of P_3 and P_4:

$$\xi_i = \exp\left(-\Delta t\left(\sum_{m=1}^{n-i+1}\left(\frac{\delta_{\tau(i+m-1)}}{2} + \sum_{j=m}^{i+m-2}\delta_{\tau j}\right) + \sum_{m=n-i+2}^{n}\left(\sum_{j=m}^{n}\delta_{\tau j}\right)\right)\right) \qquad (A6)$$

where the first m summation is dropped if $i = n+1$, and the second is dropped if $i = 1$.

TABLE I.

Values of Physical Parameters

Symbol	Meaning	Value
μ	specific growth rate	0.017 min^{-1} (40 min gen. time)
$[\tau]^*$	concentration of viable mRNA at steady state	1.4×10^{-8} M (5 per cell)[a]
$[\pi]^*$	steady state protein concentration	5.54×10^{-6} M (2000 per cell)[a]
$\delta_{\tau i}$	degradative rate constant for the i^{th} zone	0.014 min^{-1} (full-length $t_{\frac{1}{2}} = 2.5$ min)
δ_π	degradative rate constant for protein π	0
L	length of open reading frame	2000 bases
ℓ	lag time for synthesis of π	0.67 min[b]
K	equilibrium constant for promoter binding	1×10^7 M^{-1}
k	rate constant for association of π and τ	6×10^9 $M^{-1}min^{-1}$
r	rate constant for dissociation of π and τ	30 min^{-1}

[a]Assuming a cytoplasmic solution volume of $6. \times 10^{-16}$ L.

[b]At the specified growth rate, translation proceeds at a rate of about 1000 amino acid residues per minute (Bremer and Dennis, 1987).

Figure Captions

Figure 1.

Schematic illustration of the relationship between a subsystem and the remainder of the cell. The "black frame" represents the multitude of unknown or poorly characterized components and processes that comprise a bacterial cell. The subsystem consists of the processes involved in expression of a particular gene. As indicated by the bold arrows, these processes are dependent upon properties of the black frame. The dashed arrow indicates the possibility that black frame properties will be significantly influenced by the subsystem protein. If this influence is so small as to be negligible, we can model transient subsystem behavior in the context of a static (steady state) black frame without incurring serious errors. If the influence is not negligible, serious errors will result. This manner of representing subsystem/system interaction and the term "black frame" are borrowed from P. M. Schlosser (see Schlosser and Bailey, 1990).

Figure 2.

Poisson distribution (mean = 5). If the mean number of transcripts of a particular type per cell is 5, we would expect the actual number per cell to be distributed as shown (assuming no transcriptional control).

Figure 3.

Response curves for synthesis of π in the absence of regulation. (A) π concentration relative to the steady state concentration ($[\pi]^*$). (B) Total τ concentration relative to the steady state concentration, $[\tau]^*$.

Figure 4.

Response curves for the synthesis of π under transcriptional regulation only. Dashed curves represent the unregulated response (identical to solid curves in Figure 3). (A) π concentration relative to the steady state concentration. (B) Total τ concentration relative to the steady state concentration.

Figure 5.

Response curves for synthesis of π under translational regulation only. Dashed curves represent the unregulated response (identical to solid curves in Figure 3). (A) π concentration relative to the steady state concentration. (B) Total τ concentration relative to the steady state concentration (dashed curves indistinguishable from solid curves).

Figure 6.

Response curves for synthesis of π under both transcriptional and translational regulation. Dashed curves represent the response obtained with transcriptional regulation only (identical to solid curves in Figure 4). (A) π concentration relative to the steady state concentration. (B) Total τ concentration relative to the steady state concentration.

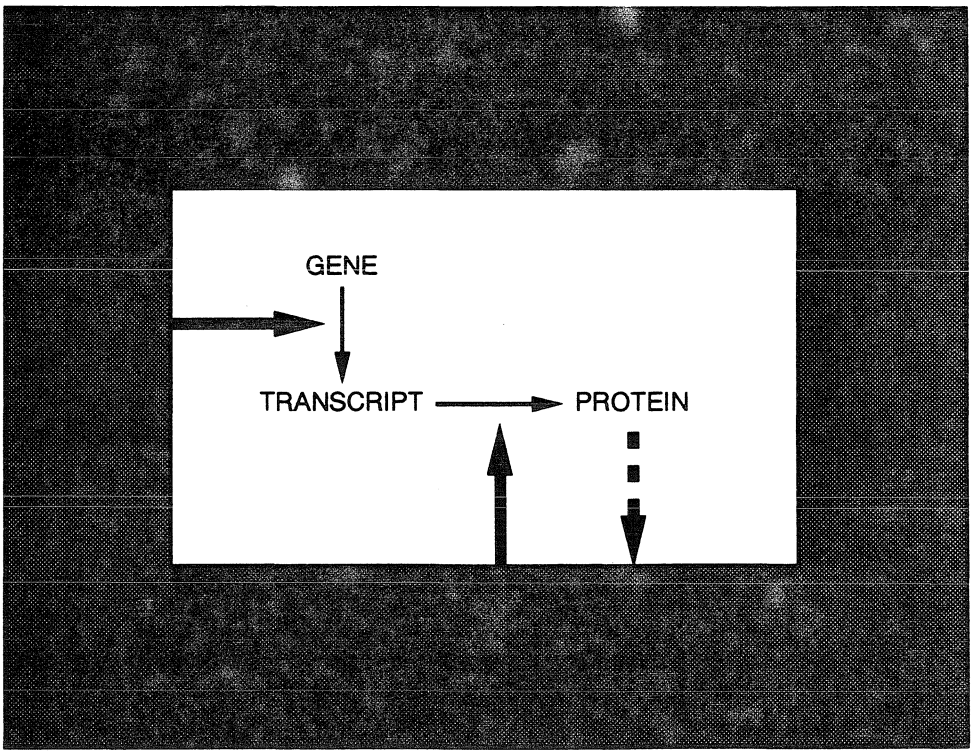

Figure 1

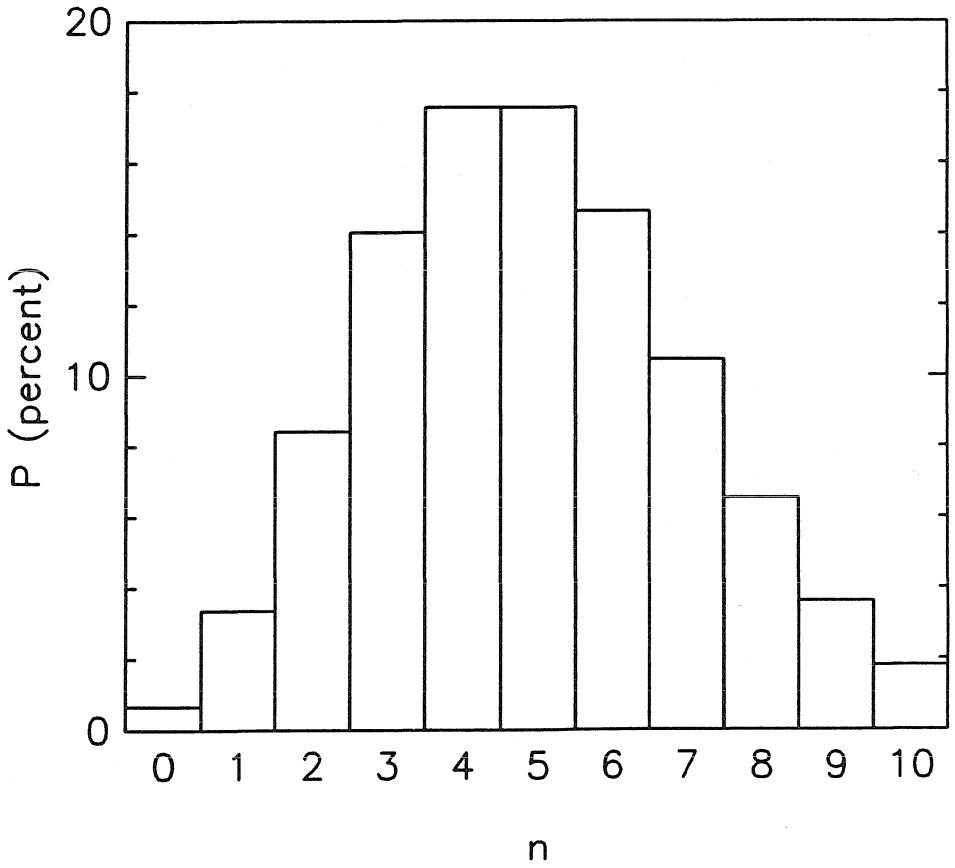

Figure 2

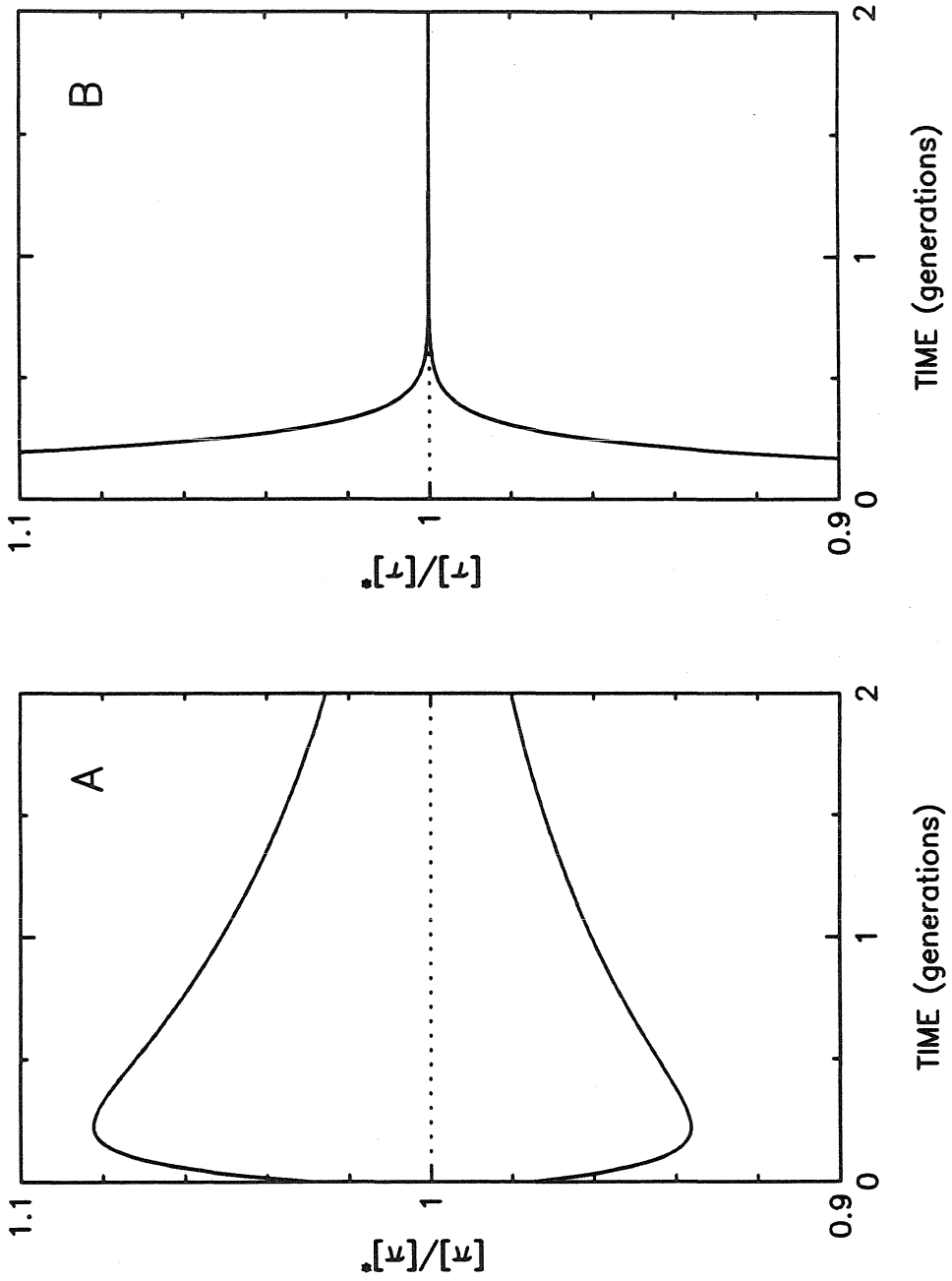

Figure 3

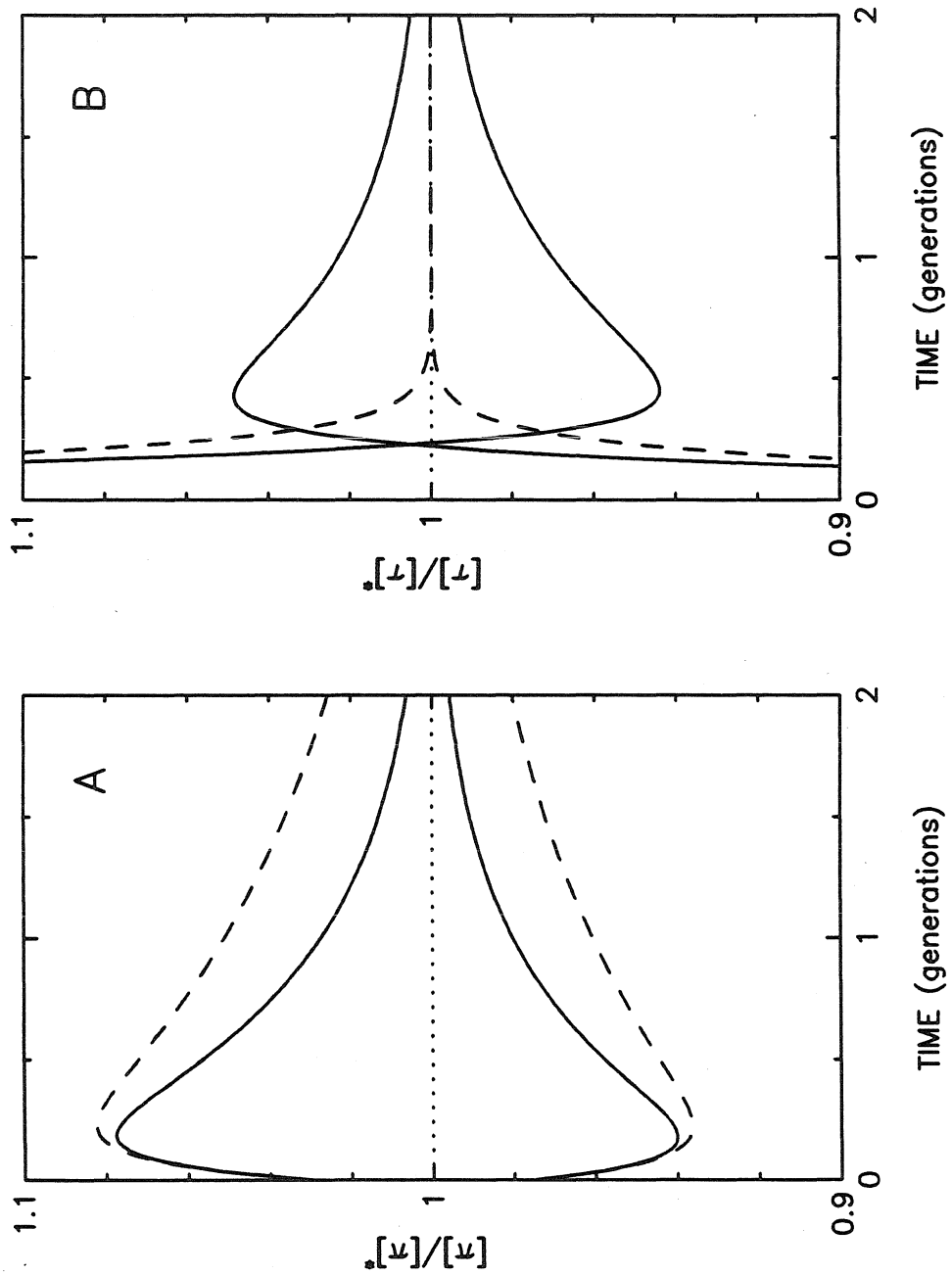

Figure 4

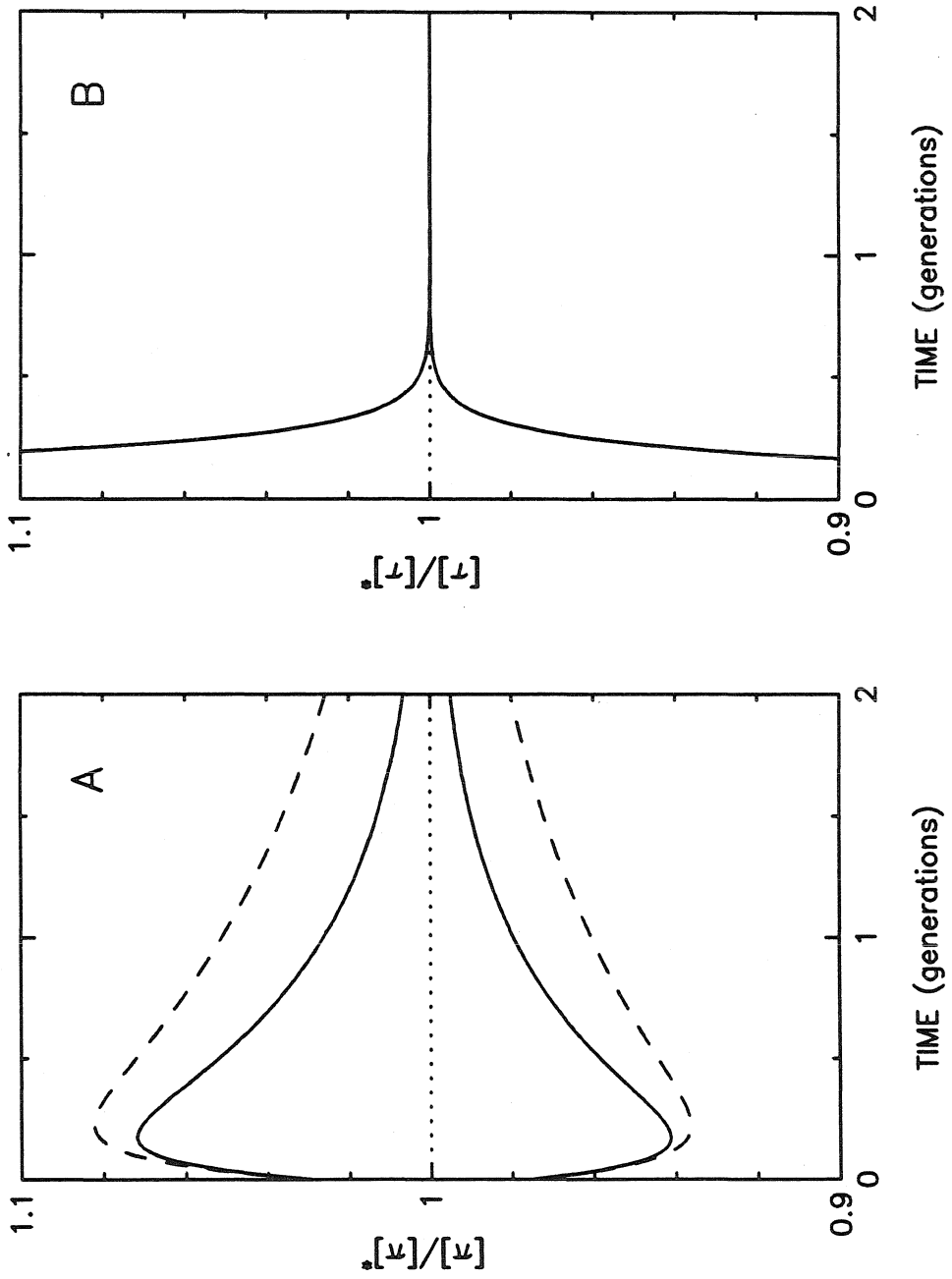

Figure 5

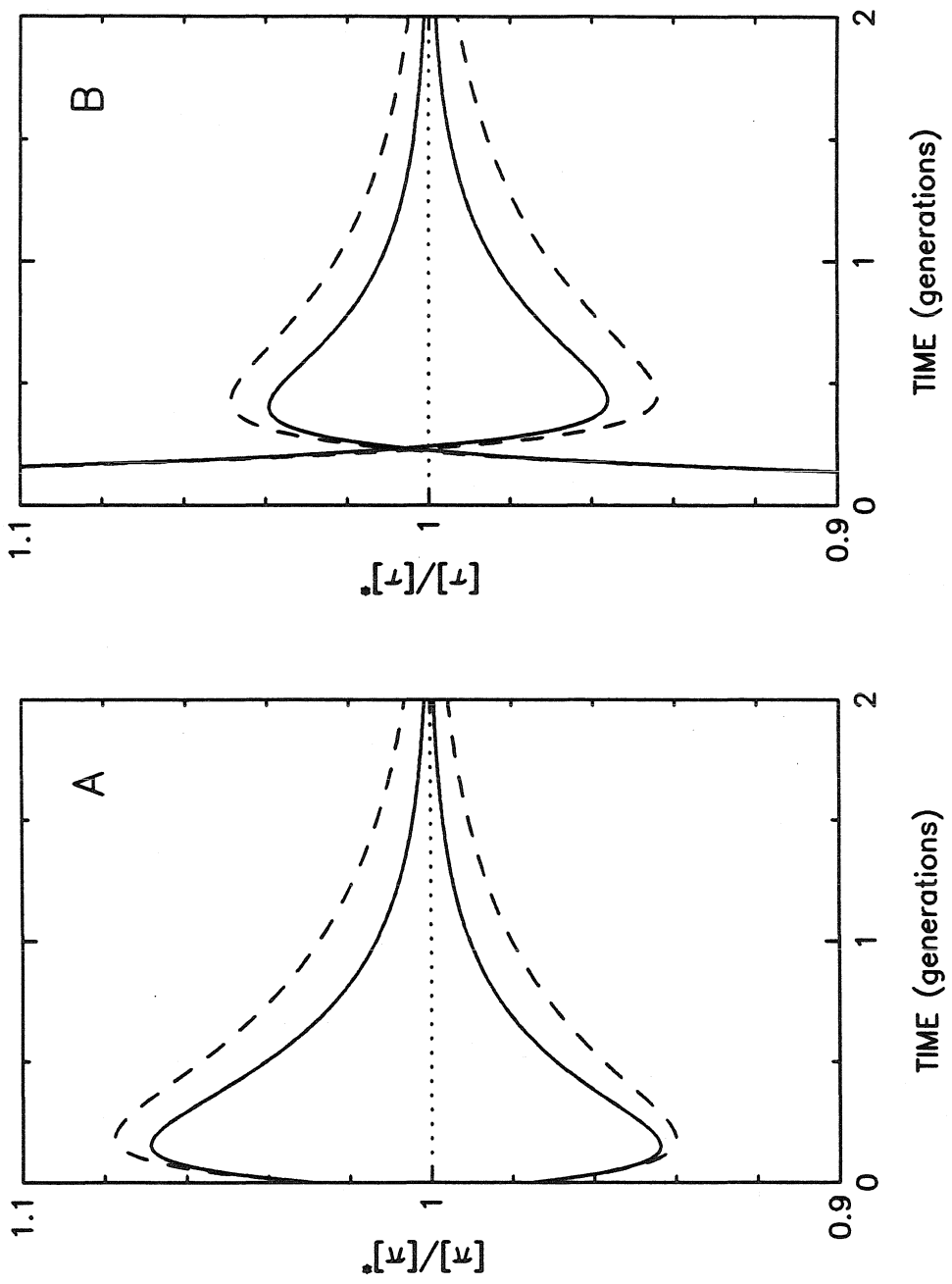

Figure 6

CHAPTER II.

TRANSLATIONAL REGULATION OF *rpoB* EXPRESSION IN *Escherichia coli:* A THEORETICAL STUDY

The material in this chapter has been submitted for publication in the *Journal of Theoretical Biology.*

Summary

Expression of *rpoB*, the gene coding for the β subunit of RNA polymerase, poses a significant control challenge to the bacterial cell because RNA polymerase is directly involved in *rpoB* expression and in the expression of all genes. The former implies counterproductive feedback, and the latter implies that there is little tolerance for fluctuation of RNA polymerase levels. Experimental results from many laboratories have demonstrated that *rpoB* expression is regulated at both the transcriptional and translational levels. Furthermore, it has been established that translational regulation is achieved by some form of autogenous repression. In this work, a theoretical method that has been developed for studying the regulation of synthesis of crucial proteins is applied to the problem of translational regulation of *rpoB*. A simple translational regulatory mechanism consisting of RNA polymerase binding to a site on *rpoBC* mRNA appears to provide inadequate control. A more sophisticated mechanism involving cooperative binding of multiple RNA polymerase molecules is proposed and found to improve control properties substantially.

1. Introduction

RNA polymerase is the enzyme responsible for synthesizing bacterial mRNAs, tRNAs, and rRNAs. In *E. coli* the form of this enzyme that catalyzes DNA-dependent RNA chain elongation (referred to as the *core* enzyme) is composed of two identical α subunits, a β subunit, and a β' subunit. Another subunit, σ^{70}, binds reversibly to the core enzyme to impart promoter specificity to the entire complex (the *holoenzyme*). In fact, it has become apparent that other σ factors exist which confer different promoter specificities to the holoenzyme (Grossman *et al.*, 1984; Hirschman *et al.*, 1985; Hunt and Magasanic, 1985), though σ^{70} is responsible for the bulk of transcription. The β and β' subunits together comprise 80% of the mass of the core enzyme. The genes coding for these subunits, *rpoB* and *rpoC* respectivly, occupy adjacent positions in an operon that also codes for two small ribosomal proteins (see Figure 1).

Expression of many bacterial genes is specifically regulated at the level of transcription. Even in cases in which translational control dominates, the rate of transcription significantly influences the rate of protein synthesis. Hence, a complete understanding of the mechanisms by which a bacterial cell simultaneously controls expression of over one thousand genes cannot be achieved until the regulation of RNA polymerase synthesis is well understood. A similar argument could be made for the importance of the regulation of ribosomal synthesis. This has been studied in considerable detail (Nomura *et al.*, 1984; Lindahl and Zengel, 1986), and rather elaborate mechanisms for coordinating the synthesis of the fifty-five macromolecules that comprise a ribosome have emerged. Though fewer details have been established with respect to synthesis of RNA polymerase, one might expect a bacterial cell to regulate this in a comparably sophisticated manner.

The topic of control of RNA polymerase synthesis was addressed as early as 1973 (Hayward *et al.*), and while many papers have appeared since then, a mechanism has not yet been established. A brief analysis of the work published in this area appears in the Appendix. The most important conclusions for the purposes of this work are summarized here. First, synthesis of RNA polymerase appears to be limitted by the availability of β and β' subunits. Secondly, it has been well established that β and β' subunit synthesis is specifically regulated in *E. coli*. Thirdly, the evidence strongly suggests that *rpoBC* regulation occurs at both the transcriptional and the translational steps. Fourthly, it has been demonstrated that the translational control mechanism is autogenous (i.e., an *rpoBC* gene product is directly involved in controlling translation of the *rpoBC* message), and finally, there is reason to believe that translational regulation plays a dominant role in the overall control scheme.

The last two points taken together pave the way for model formulation. The most straightforward model of autogenous translational regulation would involve binding of RNA polymerase, or some subunit combination thereof, to the *rpoBC* mRNA in such a way as to prevent translation. This has in fact been proposed (Ishihama and Fukuda, 1980). We were prompted to offer a more specific hypothesis after making the following observation. The published nucleotide sequence of *rpoB* (Ovchinnikov *et al.*, 1981) includes, in the vicinity of the Shine-Dalgarno sequence, the following stretch:

$$5' - \mathbf{TTGTCAGCGAGCT}\underline{\mathbf{GAGG}}\mathbf{AACCCTATGGT} - 3'$$

The Shine-Dalgarno sequence, GAGG, is underlined. The boldface sequences bear striking homology to the consensus *E. coli* promoter (TTGACA/16-18N/TATAAT), yet no promoter activity has been reported in this region. Since the Pribnow-box

homology includes the ATG start codon of *rpoB*, transcriptional activity at this promoter-like region could not possibly lead to synthesis of the β subunit. Our hypothesis is that holoenzyme binds specifically to this site and, in so doing, blocks the Shine-Dalgarno sequence from ribosomes.

Some comments should be made at this point regarding the implications of sequence-specific binding to a single-stranded nucleic acid. Binding of holoenzyme to a promoter is usually viewed as a two-step process (see McClure, 1985). The holoenzyme first binds loosely to duplex DNA at the promoter sequence. The holoenzyme-promoter complex then isomerizes to form an extremely stable complex where the DNA strands are separated in the vicinity of the promoter. Studies with the *lac* UV5 and T7 A3 promoters (Siebenlist *et al.*, 1980) have demonstrated that binding of RNA polymerase to a promoter primarily involves contacts with the non-template strand. The promoter-like sequence on *rpoBC* mRNA is the RNA analogue of this strand. It thus seems plausible that RNA polymerase might bind to a "pseudo-promoter" region on RNA with contacts analogous to those made with the nontemplate strand of an actual promoter. The physical process by which such a complex forms, though, would probably be quite different from the process of promoter binding.

The objective of this work is to quantitatively compare the efficacy of several proposed models of translational regulation of *rpoB* expression. Expression of *rpoB* is unusual in several respects. In the first place, since the rate of RNA polymerase synthesis is presumed to equal the rate of β subunit synthesis (see Appendix), we have a situation where expression of a gene is directly dependent on the gene product. In the absence of regulation, a decrease in the rate of β synthesis will result in a decrease in the concentration of RNA polymerase, which will result in a decrease in the rate of β synthesis. This means that the unregulated system is

naturally unstable. While this is unusual as far as gene expression is concerned, it is not particularly problematic from a modeling viewpoint.

The difficulty in modeling *rpoB* expression lies rather in the fact that all protein synthesis and hence all of metabolism is critically dependent upon RNA polymerase. That is, even a modest change in the level of RNA polymerase is expected to alter the growth state of a cell. The significance of this with respect to modeling is that one cannot treat synthesis of RNA polymerase as distinct from other cellular processes without incurring serious errors. On the other hand, mathematical description of the entire cell is obviously not feasible. An approach to studying regulation of gene expression in situations of this kind is described in the accompanying paper (Axe and Bailey, 1990). Briefly, the approach involves modeling maintenance of balanced exponential growth rather than transitions between growth configurations. We can thereby circumvent the need to characterize the response of the entire cell to a permanent shift in the concentration of a crucial protein. All cellular processes except synthesis of the protein of interest are considered to be at steady state. The quality of control is then characterized by computing the response of this expression system to an internal perturbation. Since random fluctuation in the number of transcripts coding for the crucial protein is apt to be the dominant perturbation, we have incorporated this into the approach.

An assumption inherent in this approach is that transitory perturbations in the processes directly involved in synthesis of the protein of interest do not propagate into surrounding cellular processes. If such propagation did occur, a rigorous approach would have to include a description of the transient response of the entire cell (the very thing we have attempted to avoid). In fact, this assumption is never wholly justified. It always represents an approximation that deviates more or less from reality depending on the magnitude of the perturbation and the nature of

the protein involved. Where synthesis of crucial proteins is concerned, neglecting propagation of a perturbation would almost certainly lead to prediction of an unrealistically rapid return to the steady state. That is, a model that confines the displacement to a small subsystem will predict a more rapid recovery than a model that describes the transition from a local displacement to a global one. The results of this modeling approach should be interpreted with this in mind. The magnitude of the discrepancy between model response and actual cell behavior will depend on how crude the above assumption is, but the direction of the error will always be toward more rapid recovery. Computed transients should then be interpreted as optimistic bounds on the true bahavior, being more realistic for smaller perturbations in the level of the crucial protein.

2. Equations Governing Expression

In the previous paper, synthesis of mRNA was described by:

$$\frac{d[\tau_1]}{dt} = \theta_1 \Phi_1 - \mu[\tau_1] \tag{1}$$

where τ_1 represents newly initiated transcripts (specifically, transcripts in the first length class; see Axe and Bailey, 1990 for a complete description). This equation must be modified for the current application because Φ_1 is actually a function of the RNA polymerase concentration, which is to be described explicitly. In light of the evidence that synthesis of β and β' determines the rate of RNA polymerase synthesis, and that *rpoB* and *rpoC* are expressed in parallel, we will consider the product of *rpoB* expression to be RNA polymerase. Equation (1) will therefore be rewritten as:

$$\frac{d[\tau_1]}{dt} = \theta_1 \hat{\Phi}_1[P_f] - \mu[\tau_1] \tag{2}$$

where $[P_f]$ is the concentration of free RNA polymerase, and $\hat{\Phi}_1$ can be interpretted as the strength of the upstream promoter, P_{L10} (see Figure 1).

The equation describing degradation of older transcripts:

$$\frac{d[\tau_i]}{dt} = -\left(\mu + \sum_{j=1}^{i-1} \delta_{\tau j}\right)[\tau_i] \tag{3}$$

is unchanged from the previous form, and the equation describing synthesis of protein:

$$\frac{d[P]}{dt} = \left[\theta_2 \Phi_2 \sum_{i=1}^{n+1} \xi_i[\tau_i]\right]_{t-\ell} - \mu[P] \tag{4}$$

differs only in that $[\pi]$ has been replaced by $[P]$, the total RNA polymerase concentration, and the degradation term has been dropped (RNA polymerase is stable in growing cells, Iwakura *et al.* 1974). The fraction of RNA polymerase that is freely diffusing in the cytosol is a function of such things as the stage of chromosomal replication and the induction/repression state of all operons that are regulated at the level of transcription. Since mean values of these over a population will be time invariant during exponential growth, $[P_f]$ will be a constant fraction of $[P]$. Equations (2),(3) and (4) form the basic framework within which translational regulation will be examined. θ_2 in equation (4) represents the translational control function. Different conceptual models of autogenous regulatory mechanisms will be expressed here in corresponding mathematical terms. The resulting expressions will be substituted for θ_2 in order to evaluate the control performance.

Before considering control mechanisms, it is instructive to examine the response of these equations in the absence of control. This is done by fixing the values of θ_1 and θ_2 and solving equations (2)-(4) using initial conditions that represent displaced mRNA levels (the method for determining initial conditions is explained in Axe and Bailey, 1990). Figure 2 shows the computed responses for positive and negative

mRNA displacements. The responses indicated by dashed lines were computed with the same parameter values, the only difference being that the protein product was hypothetically considered to be unrelated to RNA polymerase (i.e., the equations from the previous paper were used without modification). The effect of coupling the protein product to transcription is dramatic. The solid curves in Figure 2B indicate that while partial $[\tau]$ recovery is rapidly achieved, a large permanent displacement remains. In the uncoupled case, complete $[\tau]$ recovery is attained in a fraction of a generation. $[P]$ is also permanently displaced in the actual (coupled) case, whereas in the hypothetical case the original steady state value is recovered.

Two comments should be made regarding the interpretation of Figure 2. First, in view of the final points made in the Introduction, we would expect actual responses of cells lacking regulation at *rpoB* to be considerably worse than the model responses. Second, the equations used to obtain these curves describe mean properties of a large population in exponential growth. Concentrations of τ and P in individual cells would be distributed about these mean values. Hence, the proper interpretation of the solid curves is not that $[P]$ and $[\tau]$ will attain new steady state values after the displacement, but instead that $[P]$ and $[\tau]$ will *not* recover their original steady state values after a displacement. Because cells are continually subjected to the internal perturbation, the distribution about the mean would broaden with time (though the mean itself would remain fixed), and no steady state could exist. The conclusion to be drawn from these considerations is that, in the context of this approach, recovery of the initial steady state is a necessary but insufficient condition for balanced exponential growth to be realizable.

3. Transcriptional Regulation

Transcriptional regulation is not the primary concern of this work, yet since it has been established that *rpoB* expression is under transcriptional control, θ_1 in equation (1) should be replaced by an appropriate expression. The mechanism of transcriptional regulation has not been identified. One possibility is that it involves the transcription attenuator located in the intergenic region between *rplL* and *rpoB* (Figure 1). Roughly 75% of the transcripts initiated at P_{L10} are normally terminated at this site (Ralling and Linn, 1984; Downing and Dennis, 1987). It is conceivable that the termination frequency is modulated either directly or indirectly by the concentration of RNA polymerase.

Dennis and coworkers (1985) examined an *E. coli* mutant that weakly suppresses an amber lesion in *rpoB*. The rate of β subunit synthesis was found to be half the normal rate, relative to total protein synthesis, and the rate of *rpoB* mRNA synthesis was found to be 60% higher than normal, relative to total RNA synthesis. These numbers can be used to construct an empirical expression for θ_1:

$$\theta_1 = 0.25 + 0.3(1 - [P]/[P]^*). \tag{5}$$

Here $[P]^*$ is the steady state RNA polymerase concentration. On the supposition that transcriptional regulation is accomplished by modulating of the frequency of termination, we have scaled θ_1 such that the steady state value corresponds to the normal readthrough frequency. This is of no mathematical significance, however, because θ_1 can be multiplied by an arbitrary constant (Φ_1 would be multiplied by the inverse). Equation (5) represents no mechanism; it is simply a means of incorporating transcriptional control that roughly corresponds to actual cell behavior.

By fixing the value of θ_2 and using the above expression for θ_1, we have computed model responses corresponding to transcriptional control in the absence of

translational control (Figure 3). The control response is manifested in $[\tau]$ overshooting the value defined by the horizontal portion of the dashed curve (Figure 3B). The extent of the overshoot determines the rapidity and extent of $[P]$ recovery. If the transcriptional control function were slightly weaker, $[\tau]$ would not overshoot $[\tau]^*$, and $[P]$ would again fail to return to $[P]^*$. Instead, transcriptional regulation defined by equation (5) is alone sufficient to enable the population to recover the original steady state. While the maximum $[P]$ displacement is nearly as great as before, the displacement falls to less than a quarter of this after two generations of recovery. Still, it seems unlikely that a 3% to 6% displacement in RNA polymerase concentration over an entire generation would not lead to a sizable global perturbation. The responses in Figure 3, then, should also be viewed as optimistic.

4. Assigning Values to Kinetic Parameters

The translational regulatory models to be examined involve kinetic rate constants. Selection of values for these constants will be based upon two criteria. We will require first that values used be physically reasonable and second that they result in optimal control performance of the model being tested. Here we develop a general method for optimizing parameter sets. Physical limitations will be discussed in connection with specific regulatory models.

The control functions, θ_1 and θ_2, may depend on many cell parameters, but they must ultimately depend on the the concentration of the protein whose synthesis they control (in this case $[P_f]$). In cases where the protein of interest affects θ_1 or θ_2 indirectly, it may be difficult to express the dependence in mathematical terms. Autogenous regulatory mechanisms are less likely to pose problems in this respect. Whether or not an autogenous mechanism is being considered, it is often impossible

to express θ_2 as an explicit function of the protein concentration (an example of this kind follows; see also Axe and Bailey, 1990). The reason is that translational control typically involves binding to mRNA, and degradation of mRNA is so rapid that binding cannot generally be described by equilibrium equations. Consequently, θ_2 is history-dependent rather than being strictly dependent on the instantaneous concentration of the binding protein. The parameter optimization procedure to be used requires an explicit control function, so an approximation is usually necessary. One approach would be to assume rapid binding equilibrium. This enables one to express θ_2 explicitly, but in some cases (see section 6 for example) rapid equilibrium is physically unrealistic, and the results of the optimization would be incorrect.

A more accurate explicit representation of θ_2 can be obtained by assuming steady state behavior. This expression will be exact when the system is unperturbed and approximate otherwise. Hence, kinetic parameters selected in this way will result in optimal recovery from very small displacements. While it may be possible to "tune" parameter values for optimal recovery from a displacement of a specific magnitude and direction, it is unlikely that one can improve upon the steady state optimization as far as general control performance is concerned. This method will be used in section 6 (Translational Regulation (case 2)).

Consider a pair of explicit control functions, $\hat{\theta}_2([P_f])$ and $\check{\theta}_2([P_f])$, with specified kinetic parameters. We will describe $\hat{\theta}_2$ and $\check{\theta}_2$ as *similar* functions if and only if there exists a positive constant, κ, such that $\hat{\theta}_2([P_f]) = \check{\theta}_2(\kappa[P_f])$ for all values of $[P_f]$. The set of all θ_2 that are similar to a particular θ_2 will be referred to as a similarity class. The control performance of a specific control function is generally a function of $[P_f]$. That is, $\hat{\theta}_2$ generally exhibits better control properties at some $[P_f]$ values than it does at others. If $\underline{[P_f]}$ is the RNA polymerase concentration at which $\hat{\theta}_2$ is most acceptable as a control function, then $\kappa\underline{[P_f]}$ is the concentration

at which $\check{\theta}_2$ is most acceptable. The significance of this is that given a measure of control acceptablity that is invariant over an entire similarity class, we need only consider acceptability at the *class* level. Once the optimal class has been determined, we can easily determine the control function within that class that exhibits optimal control at $[P_f] = [P_f]^*$.

The function Ψ, defined as:

$$\Psi \equiv -\frac{[P_f]}{\theta_2}\frac{d\theta_2}{d[P_f]}, \tag{6}$$

is a suitable measure of control acceptability. When $\hat{\Psi}(\hat{\theta}_2, [P_f])$ is plotted as a function of $[P_f]$, the $[P_f]$ value corresponding to the maximum of $\hat{\Psi}$ is the free RNA polymerase concentration at which $\hat{\theta}_2$ exhibits the most stringent control, and the maximum $\hat{\Psi}$ value is a measure of the maximum stringency of the similarity class represented by $\hat{\theta}_2$. Thus, the general method for optimizing kinetic parameters is to select a group of nonsimilar parameter sets that adequately samples the entire spectrum of similarity classes, plot Ψ versus $[P_f]$ for each of these sets, identify the parameter set that yields the highest Ψ value, and finally compute parameters for the similar θ_2 function that exhibits a maximum Ψ at the desired steady state $[P_f]$.

5. Translational Regulation (Case 1)

The simplest model of autogenous translational regulation is illustrated schematically in Figure 4. RNA polymerase binds specifically to a site in the vicinity of the Shine-Dalgarno sequence on *rpoBC* mRNA, thereby preventing translation of the message. Binding is assumed to have no effect on message stability. The equations governing expression must now be modified to reflect the different com-

plexation states of *rpoBC* mRNA. Equation (2) becomes:

$$\frac{d[\tau_1]_0}{dt} = \theta_1 \hat{\Phi}_1 [P_f] + r_1 [\tau_1]_1 - (k_1 [P_f] + \mu)[\tau_1]_0, \tag{7}$$

and equation (3) is replaced by:

$$\frac{d[\tau_i]_0}{dt} = r_1 [\tau_i]_1 - \left(k_1 [P_f] + \mu + \sum_{j=1}^{i-1} \delta_{\tau j} \right) [\tau_i]_0 \tag{8}$$

and

$$\frac{d[\tau_i]_1}{dt} = k_1 [P_f][\tau_i]_0 - \left(r_1 + \mu + \sum_{j=0}^{i-1} \delta_{\tau j} \right) [\tau_i]_1 \tag{9}$$

where $[\tau_i]_0$ and $[\tau_i]_1$ are, respectively, the concentrations of unbound and bound *rpoBC* mRNA in length class i, and k_1 ($M^{-1}sec^{-1}$) and r_1 (sec^{-1}) are the rate constants for association and dissociation. Equation (9) applies to all length classes ($\delta_{\tau 0}$ is taken to be zero), whereas equation (8) applies to all but the first. The equation governing $[P]$:

$$\frac{d[P]}{dt} = \left[\Phi_2 \sum_{i=1}^{n+1} \xi_i [\tau_i]_0 \right]_{t-\ell} - \mu [P]. \tag{10}$$

differs from equation (4) only in that θ_2 is now implicit due to the fact that the summation excludes bound mRNA.

Physical limitations should be considered before attempting to optimize the values of k_1 and r_1. A theoretical upper bound on the rate constant for RNA polymerase/promoter association has been estimated by assuming the interaction to be controlled entirely by diffusion (von Hippel *et al.*, 1984). The calculated value (10^8 $M^{-1}sec^{-1}$) is two orders of magnitude lower than experimentally determined apparent rate constants for strong promoters (Bujard *et al.*, 1982; Chamberlin *et al.*, 1982), suggesting that a process other than free diffusion, such as linear diffusion along DNA, is involved in promoter recognition. Linear diffusion of RNA

polymerase along *rpoBC* mRNA seems unlikely in that it would imply a general interaction between RNA polymerase and mRNA. Since the estimate of von Hippel *et al.* would not be greatly different if the binding site were on RNA instead of DNA, we will consider association rate constants with values exceeding (10^8 $M^{-1}sec^{-1}$) to be physically unrealistic.

The rate constant for dissociation is a function of the binding energy, tight binding corresponding to a low r_1 value. Rate constants for dissociation of RNA polymerase from strong promoters can be as low as 10^{-5} sec^{-1} (Cech and McClure, 1980). An upper bound on r_1 can be estimated from the velocity of translational motion of freely diffusing RNA polymerase and the approximate distance over which binding occurs. We estimate the nonbinding limit of r_1 to be on the order of 10^{10} sec^{-1}. The value of r_1 would have to be much lower than this for the interaction to be termed "binding" in the usual sense. We will use a conservative upper limit of 3×10^1 sec^{-1} for r_1.

For the purpose of selecting optimal k_1 and r_1 values, we will initially assume that binding is in rapid equilibrium. If it proves impossible to approach the maximum Ψ value while satisfying this condition, the steady state equations can be used. When binding is equilibrated, $[\tau_i]_1$ is related to $[\tau_i]_0$ by:

$$\frac{[\tau_i]_1}{[P_f][\tau_i]_0} = K_1 \equiv \frac{k_1}{r_1} \tag{11}$$

where K_1 is the equilibrium constant for binding. Defining θ_2 as the fraction of *rpoBC* mRNA that is unbound, we have:

$$\theta_2 = \frac{1}{K_1[P_f] + 1} \tag{12}$$

which can be substituted into equation (4) to describe $[P]$. The form of control function represented by equation (12) has only one similarity class because the only

adjustable parameter, K_1, is a coefficient of $[P_f]$. For this class Ψ has the form:

$$\Psi = \frac{K_1[P_f]}{K_1[P_f]+1}. \tag{13}$$

Here we see by inspection that for a given value of K_1, Ψ asymptotically approaches 1 as $[P_f]$ approaches infinity. For a given $[P_f]^*$, optimal response would then be obtained when K_1 is infinite.

The values of k_1 and r_1 are restricted by the rapid equilibrium condition in addition to physical limitations. Binding equilibrium will be approached if both $k_1[P_f]^*$ and r_1 are much greater than the total degradative rate constant for the full-length transcript, 1.2×10^{-2} sec^{-1}. If k_1 assumes its maximum value, 1×10^8 $\text{M}^{-1}\text{sec}^{-1}$, then $k_1[P_f]^* = 1.3 \times 10^1 \gg 1.2 \times 10^{-2}$ sec^{-1}. Conflicting demands are being placed on r_1. Rapid equilibrium requires that r_1 be large while optimal response requires that it be small, but since "large" and "small" are to be judged relative to different numbers we can attempt a compromise. If r_1 is taken to be 0.5 sec^{-1} (substantially greater than 1.2×10^{-2} sec^{-1}), the equilibrium constant, K_1, has the value 2.0×10^8 M^{-1}. Setting $[P_f]$ equal to $[P_f]^*$ in equation (13), we obtain $\Psi = 0.96$, which indicates that these k_1 and r_1 values result in a control stringency that approaches the theoretical maximum for equilibrium binding.

It is worth considering whether nonequilibrium binding might yield an even better response. Under steady state conditions, equation (9) can be written as:

$$\frac{d[\tau]_1}{dt} = k_1[P_f][\tau]_0 - (r_1 + \mu + \bar{\delta}_1)[\tau]_1 = 0 \tag{14}$$

where $[\tau]_1$ is the total concentration of bound mRNA, and $\bar{\delta}_1$ is the average rate constant for degradation of bound mRNA. By rearranging we obtain:

$$\frac{[\tau]_1}{[P_f][\tau]_0} = \left(\frac{k_1}{r_1 + \mu + \bar{\delta}_1}\right). \tag{15}$$

This equation reduces to the form of the equilibrium relation (equation (11)) if r_1 is the dominant term in the denominator. Comparison of these equations reveals that if μ and $\bar{\delta}_1$ are not negligible, they effectively reduce the value of K_1. That is, at steady state θ_2 and Ψ are of the same forms as those derived for equilibrium binding (equations (12) and (13)), the only difference being that K_1 is reduced by the factor $r_1/(r_1 + \mu + \bar{\delta}_1)$. This will necessarily reduce the value of Ψ vis-a-vis the rapid equilibrium value. Hence the maximum possible value of Ψ is 1, and the kinetic values determined above can be used to compute the optimal response of the current model.

The rapid-equilibrium approximation was used only to optimize kinetic parameters. Having accomplished this, we can solve the full set of equations ((7)-(10)) to obtain response curves. The results (Figure 5) indicate that addition of translational regulation improves the control response significantly. Combined regulation results in a 50% reduction in $[P]$ displacement after one generation of recovery, and a 75% reduction after two generations. This is reflected in a diminished $[\tau]$ overshoot (Figure 5B).

6. Translational Regulation (Case 2)

Translational control can be made much more stringent if binding occurs at several interacting sites. Consider a sequential binding model where the n^{th} site is formed upon binding at site $n-1$, and dissociation occurs in the reverse order of binding. Figure 6 illustrates the mechanism when the total number of sites is four. The quadruply bound message cannot be translated; all other complexes are translationally active. Rate constants for association and dissociation are subject to the same physical limits used in section 5. As in the previous model, it is assumed

that binding has no effect on degradation.

The equation describing transcript initiation for the previous case (equation (7)) is equally applicable here. The equations describing binding and degradation for all mRNA length classes and complexation states are of the form:

$$\frac{d[\tau_i]_j}{dt} = k_j[P_f][\tau_i]_{j-1} + r_{j+1}[\tau_i]_{j+1} - \left(k_{j+1}[P_f] + r_j + \mu + \sum_{m=0}^{i-1}\delta_{\tau m}\right)[\tau_i]_j \quad (16)$$

where $[\tau_i]_j$ is the concentration of mRNA in length class i and complexation state j (i.e., with j RNA polymerase molecules bound), and k_j and r_j pertain to association and dissociation of the j^{th} polymerase molecule ($k_j, r_j \equiv 0$ for $j \notin \{1,2,3,4\}$). Equation (16) applies to all $[\tau_i]_j$ (again, $\delta_{\tau 0} \equiv 0$) except $[\tau_1]_0$, which is described by equation (7). $[P]$ is now described by:

$$\frac{d[P]}{dt} = \left[\Phi_2 \sum_{i=1}^{n+1}\left(\xi_i \sum_{j=0}^{3}[\tau_i]_j\right)\right]_{t-\ell} \quad (17)$$

since only quadruply bound mRNA is untranslatable.

Selection of an optimal set of kinetic parameters becomes a more difficult task as the complexity of the model increases. It will again be useful to start with the assumption of rapid binding equilibrium. As in the previous case, we can use equilibrium relationships to obtain an explicit θ_2 function. In this case the rapid-equilibrium expression for θ_2 is:

$$\theta_2 = \frac{1 + K_1[P_f] + K_1K_2[P_f]^2 + K_1K_2K_3[P_f]^3}{1 + K_1[P_f] + K_1K_2[P_f]^2 + K_1K_2K_3[P_f]^3 + K_1K_2K_3K_4[P_f]^4} \quad (18)$$

where K_j is the equilibrium constant for the j^{th} binding step. Ψ is obtained in a straightforward manner from equation (18), but in this case visual inspection avails us little. Here the methodology outlined in section 4 (Assigning Values to Kinetic Parameters) is very useful.

The K_j in equation (18) all have units of inverse concentration. It follows that proportional K_j sets yield similar θ_2 functions. That is, if $\check{K}_j = \alpha \hat{K}_j$ for $j = 1$ to 4, then $\check{\theta}_2([P_f]) = \hat{\theta}_2(\alpha[P_f])$. Consequently, the acceptability of all proportional K_j sets can be determined by plotting Ψ versus $[P_f]$ for a single set. The series of Ψ versus $[P_f]$ curves shown in Figure 7 was obtained by specifying K_j values that form a geometric progression, $K_j = \rho K_{j-1}$ (K_1 being arbitrary). Higher values of ρ clearly produce better control functions. Physically, this means that if the tendency for RNA polymerase to bind increases with each successive binding event, a sharper control response results. Mechanisms of this type will be referred to as *cooperative* sequential binding mechanisms. The maximum rapid-equilibrium value of Ψ for this model, 4, indicates that cooperative sequential binding can potentially improve the control response dramatically.

Actually, the control benefits of cooperativity do not require K_j values to form a geometric progression. The curves in Figure 8 were obtained by fixing K_1 and K_4 such that $K_4/K_1 = 10^4$, and varying the values of K_2 and K_3 from $K_2 = K_1$, $K_3 = K_4$ to $K_2 = K_4$, $K_3 = K_1$. The maximum value of Ψ is surprisingly insensitive to changes in K_2 and K_3. Provided that $K_j \geq K_{j-1}$, the ratio of K_4 to K_1 dominates in determining the acceptability of a similarity class.

In the previous model, the conditions under which binding approached equilibrium were mathematically simple. We were thus able to easily select k_1 and r_1 values to meet these criteria. Here, since the mathematical conditions for rapid equilibrium are much more complicated, we will instead use the more accurate steady state θ_2 expression to generate Ψ versus $[P_f]$ curves. Sets of k_j and r_j for which the equilibrium and steady state curves overlap would result in near-equilibrium binding. The curves in Figure 9 were generated by selecting the largest physically reasonable rate constants (representing the closest physically reasonable approach

to equilibrium) that give the same sets of equilibrium constants used to generate the curves in Figure 7. Comparison of Figures 9 and 7 leads to several significant conclusions. Most importantly, since the upper curves in Figure 9 fall well below the corresponding curves in Figure 7, we conclude that the optimal parameter set for the four-site model precludes equilibration of the message complexes. Furthermore, the realistic optimum is achieved at an intermediate value of ρ ($K_j = \rho K_{j-1}$) rather than a maximum value. The maximum physically realistic Ψ value ($\Psi = 2.7$) is about two-thirds the theoretical maximum Ψ value.

Notwithstanding the disparity between theoretical and practical optima, the cooperative four-site model represents a marked improvement over the single site model. Realistic, near-optimal parameter values were used to generate the responses shown in Figure 10. Dashed curves in this figure are identical to the solid curves in Figure 5. After a generation of recovery, the cooperative sequential binding mechanism reduces the $[P]$ displacement 75% relative to the single-site mechanism. At this point $[P]/[P]^*$ is within 0.5% of unity. After two generations, recovery is essentially complete (0.02% displacement).

7. Discussion

Expression of *rpoB* poses a significant control challenge to the bacterial cell for two principal reasons. First, the product of this gene is directly involved in its own synthesis. This alone poses a control problem because it reinforces any displacement from the steady state. Secondly, all gene expression, and consequently all of metabolism, is critically dependent on the product of *rpoB*. While the complexity of this second aspect precludes rigorous theoretical analysis, significant information has been obtained through careful analysis of the first aspect. Specifically, four-site

cooperative sequential binding of RNA polymerase to *rpoBC* mRNA can realistically be expected to yield a threefold increase in translational control stringency (Ψ) relative to single-site binding.

We would ultimately like to take a step beyond quantitative comparison of regulatory models and evaluate their acceptability in the context of a real cell. The function Ψ has been introduced as a relative measure of acceptability, but we have not specified a control criterion either in terms of Ψ or in terms of the computed response. In principle a control scheme that maximizes μ in a given environment could be considered optimal for that environment, but it is not presently possible to relate μ mechanistically to θ_2. We might alternatively attempt to estimate the actual control stringency from experimental data. If the stringency of a particular control model differs greatly from the estimate, it can be regarded as an unlikely candidate on empirical grounds. It is difficult to obtain reliable data for this purpose. Ideally, we would like to introduce a small perturbation in $[P]$ and measure the effect on the rate of polymerase synthesis while monitoring $[\tau]$. The inaccuracy of $[P]$ and $[\tau]$ measurements, however, is such that large perturbations are required to resolve changes. Gross displacements of $[P]$ are apt to cause metabolic adjustments that are large in magnitude and unpredictable in nature.

Bearing this in mind, we will use the results of Dennis *et al.* (1985) as an estimate of the stringency of translational regulation in *E. coli*. They found that in a strain that weakly suppresses an amber lesion in *rpoB* (10% suppression), the complete β subunit (as a fraction of total cell protein) is present at 47% of the normal level. The level of *rpoBC* mRNA (as a fraction of total cell mRNA) was found to be 60% higher than the normal level. By taking the ratio of combined β and β fragment synthesis to *rpoBC* mRNA level, they estimated the "translation efficiency" in the mutant strain to be three times the control value. Equating translation efficiency

with θ_2 and β level with RNA polymerase level, we can calculate values from the single-site and four-site models for comparison. The single-site model gives a value of 2.0, and the four-site model gives a value of 5.1. Though it would be unwise to draw any firm conclusions from this comparison, it suggests that the actual stringency of translational regulation may fall between the stringencies of the single-site and four-site models.

The cooperative sequential binding mechanism presented here is probably not the only feasible mechanism that can have a Ψ value significantly greater than unity. Nor is there anything particularly significant about four sites; fewer sites would result in lower stringency, and more would result in higher stringency. The approach taken here cannot be used to distinguish the actual regulatory mechanism from plausible alternatives. Its significance lies rather in its potential for distinguishing plausible alternatives from implausible ones. Given a measure either of a bacterial cell's tollerance to diplacements in RNA polymerase concentration or of the stringency of the control response exhibited in bacterial cells, one can use this approach to test the plausibility of a proposed control mechanism. Based on the data available at this time, it appears that the simplest autogenous translational control model, binding of RNA polymerase to a single site on *rpoBC* mRNA, provides inadequate control, while the control afforded by the four-site cooperative sequential binding model appears to be excessively stringent.

The results of this work have important implications for the design of experiments to test the hypothesis of cooperative sequential binding. The rate constant for dissociation of the j^{th} complex, r_j, is inversely proportional to the half-life of that complex. We have found that optimal response occurs when r_j decreases with increasing j. This means that the half-life increases as more RNA polymerase molecules bind. The longest half-life, however, was found to be on the order of ten

seconds in the case of four binding sites. Two- and three-site cooperative sequential models would not be very different in this regard. Approaches used routinely to identify specific binding of proteins to nucleic acids (eg. filter binding assay and mobility shift assay) do not normally resolve such labile complexes.

Finally, it should be noted that there is no essential relationship between the cooperative sequential binding hypothesis and the pseudo-promoter hypothesis discussed in the Introduction. The two can be consolidated by considering the pseudo-promoter to be incompetent as a binding site until previous sites are bound. The final binding event at the pseudo-promoter would then prevent translation. Nonetheless, neither is dependent on the other. It is our intention that hypotheses like that of cooperative sequential binding and the theoretical work from which they arise can have a symbiotic relationship with experimental efforts. Theories cannot be validated apart from observation, nor can observations be interpreted apart from theory.

Acknowledgement:

This work was supported by the Energy Conversion and Utilization Technologies (ECUT) Program of the U.S. Department of Energy.

References

Axe, D. D., and Bailey, J. E. (1990). A theoretical approach to studying the regulation of synthesis of crucial proteins in bacteria. *J. Theor. Biol.*

Bedwell, D. M., and Nomura, M. (1986). Feedback regulation of RNA polymerase subunit synthesis after the conditional overproduction of RNA polymerase in *Escherichia coli*. *Molec. Gen. Genet.* **204**, 17.

Blumenthal, R. M., and Dennis, P. P. (1978). Gene expression in *Escherichia coli* B/R during partial rifampicin-mediated restrictions of transcription initiation. *Molec. Gen. Genet.* **165**, 79.

Bremer, H., and Dennis, P. P. (1987). Modulation of chemical composition and other parameters of the cell by growth rate. In: *Escherichia coli and Salmonella Typhimurium.* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M., and Umbarger, H. E., eds.) pp. 1527-1542. Washington, DC: American Society for Microbiology.

Bujard, H., Niemann, A., Breunig, K.,Roisch, U., Dressel, A., *et al.* (1982). In: *Promoters: structure and function.* (Rodriguez, R. L., and Chamberlin, M. J., eds.) pp. 121-140. New York: Praeger.

Cech, C. L., and McClure, W. R. (1980). Characterization of ribonucleic-acid polymerase-T7 promoter binding complexes. *Biochem.* **19**, 2440.

Chamberlin, M. J., Rosenberg, S., and Kadesch, T. (1982). In: *Promoters: structure and function.* (Rodriguez, R. L., and Chamberlin, M. J., eds.) pp. 34-53. New York: Praeger.

Cromie, K. D., and Hayward, R. S. (1984). Evidence for rifampicin-promoted readthrough of a fully rho-dependent transcriptional terminator. *Molec. Gen. Genet.* **193**, 532.

Dennis, P. P., Nene, V., and Glass, R. E. (1985). Autogenous posttranscriptional

regulation of RNA polymerase β and β' subunit synthesis in *Escherichia coli*. *J. Bacteriol.* **161(2)**, 803.

Downing, W. L., and Dennis, P. P. (1987). Transcription products from the *rplKAJL-rpoBC* gene cluster. *J. Molec. Biol.* **194**, 609.

Grossman, A. D., Erickson, J. W., and Gross, C. A. (1984). The *htpR* gene product of *E. coli* is a sigma factor for heat shock promoters. *Cell* **38**, 383.

Hayward, R. S., Tittawella, I. P. B., and Scaife, J. G. (1973). Evidence for specific control of RNA polymerase synthesis in *Escherichia coli*. *Nature New Biol.* **243**, 6.

Hirschman, J., Wong, P. K., Sei, K., Keener, J., and Kustu, S. (1985). Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription *in vitro*: evidence that the *ntrA* product is a σ factor. *Proc. Nat. Acad. Sci USA* **82**, 7525.

Howe, K. M., Newman, A. J., Garner, I., Wallis, A., and Hayward, R. S. (1982). Effect of rifampicin on expression of *lacZ* fused to promoters or terminaters of the *E. coli rpoBC* operon. *Nucl. Acids Res.* **10(22)**, 7425.

Hunt, T. P., and Magasanik, B. (1985). Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. *Proc. Nat. Acad. Sci. USA* **82**, 8453.

Ishihama, A., and Fukuda, R. (1980). Autogenous and post-transcriptional regulation of *Escherichia coli* RNA polymerase synthesis. *Molec. Cell. Biochem.* **31**, 177.

Iwakura, Y., Ito, K., and Ishihama, A. (1974). Biosynthesis of RNA polymerase in *Escherichia coli*. 1. Control of RNA polymerase content at various growth rates. *Molec. Gen. Genet.* **133**, 1.

Kajitani, M., Fukuda, R., and Ishihama, A. (1980). Autogenous and post-transcrip- tional regulation of *E. coli* RNA polymerase synthesis *in vitro*.

Molec. Gen. Genet. **179**, 489.

Lindahl, L., and Zengel, J. M. (1986). Ribosomal genes in *Escherichia coli. Ann. Rev. Genet.* **20**, 297.

Lohman, T. M., Wensley, C. G., Cina, J., Burgess, R. R., and Record, M. T. (1980). Use of difference boundary sedimentation velocity to investigate non-specific protein-nucleic acid interactions. *Biochem.* **19**, 3516.

Mandelstam, J. (1958). Turnover of protein in growing and nongrowing populations of *Escherichia coli. Biochem. J.* **169**, 110.

Morgan, B. A., and Hayward, R. S. (1987). Direct evidence for rifampicin-promoted readthrough of the partial terminator t_{L7} in the *rpoBC* operon of *Escherichia coli. Molec. Gen. Genet.* **210**, 358.

Neidhardt, F. C. (1987). Chemical composition of *Escherichia coli.* In: *Escherichia coli and Salmonella Typhimurium.* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M., and Umbarger, H. E., eds.) pp. 3-6. Washington, DC: American Society for Microbiology.

Newman, A. J., Ma, J. C., Howe, K. M., Garner, I., and Hayward, R. S. (1982). Evidence that rifampicin can stimulate readthrough of transcriptional terminators in *Escherichia coli. Nucl. Acids Res.* **10(22)**, 4709.

Nomura, M., Gourse, R., and Baughman, G. (1984). Regulation of the synthesis of ribosomes and ribosomal components. *Ann. Rev. Biochem.* **53**, 75.

Ovchinnikov, Y. A., Monastyrskaya, G. S., Gubanov, V. V., *et al.* (1981). The primary structure of *Escherichia coli* RNA polymerase. Nucleotide sequence of the *rpoB* gene and amino acid sequence of the β-subunit. *Eur. J. Biochem.* **116**, 621.

Passador, L., and Linn, T. (1989). Autogenous regulation of the RNA polymerase β subunit of *Escherichia coli* occurs at the translational level in vivo. *J. Bacteriol.* **171**, 6234.

Peacock, S., Cenatiempo, Y., Robakis, N., Brot, N., and Weissbach, H. (1982). *In vitro* synthesis of the first dipeptide bond of the β subunit of *Escherichia coli* RNA polymerase. *Proc. Nat. Acad. Sci. USA* **79**, 4609.

Pedersen, S., and Reeh, S. (1978). Functional mRNA half lives in *E. coli*. *Molec. Gen. Genet.* **166**, 329.

Ralling, G., and Linn, T. (1984). Relative activities of the translational regulatory sites in the *rplKAJL rpoBC* gene cluster of *Escherichia coli*. *J. Bacteriol.* **158(1)**, 279.

Siebenlist, U., Simpson, R. B., and Gilbert, W. (1980). E. coli RNA polymerase interacts homologously with two different promoters. *Cell* **20**, 269.

von Hippel, P. H., Bear, D. G., Morgan, W. D., and McSwiggen, J. A. (1984). Protein-nucleic acid interactions in transcription: a molecular analysis. *Ann. Rev. Biochem.* **53**, 389.

Appendix

Control of RNA Polymerase Synthesis (Background)

Several investigators have reported a substantial increase in synthesis of the β and β' subunits of RNA polymerase relative to total protein synthesis after addition of rifampicin either to *in vitro* transcription/translation systems or to growing cells (Hayward *et al.*, 1973; Blumenthal and Dennis, 1978; Kajitani *et al.*, 1980; Ishihama and Fukuda, 1980). Without exception, the effect could be attributed to an increased level of *rpoBC* transcription; no translational effect was observed. The initial interpretation was that transcription of *rpoBC* is repressed by the β subunit or some complex thereof (inferred from the fact that rifampicin binds to this subunit).

Seemingly contradictory results were obtained when putative repressors were added to *in vitro* transcription/translation systems containing the *rpoBC* operon (no rifampicin). These experiments demonstrated specific repression of *rpoBC* expression at the level of *translation* either when RNA polymerase holoenzyme ($\alpha_2\beta\beta'\sigma$) was added (Kajitani *et al.*, 1980; Peacock *et al.*, 1982), or when $\alpha_2\beta$, an intermediate in the assembly of RNA polymerase, was added (Kajitani *et al.*, 1980). No evidence of transcriptional repression was found in these studies.

The conflict now appears to stem from misinterpretation of the rifampicin effect. Two small genes encoding ribosomal proteins are situated upstream of *rpoB* and are cotranscribed with *rpoBC* (Figure 1). The large intergenic region immediately upstream of *rpoB* contains a transcriptional attenuator that terminates roughly 75% of the transcripts initiated at P_{L10} (Downing and Dennis, 1987; Ralling and Linn, 1984). Morgan and Hayward (1987) have demonstrated that rifampicin promotes readthrough at this attenuator *in vivo*, thus explaining the

earlier observations with rifampicin. Furthermore, work in the same laboratory has suggested that the antitermination effect of rifampicin is *general* rather than unique to the *rplL-rpoB* attenuator (Newman *et al.*, 1982; Howe *et al.*, 1982; Cromie and Hayward, 1984).

It appears, then, that rifampicin has two physiologically distinct effects. Primarily, of course, it prevents initiation of transcription, and secondly it interferes with transcriptional termination. The latter is presumably a peripheral consequence of rifampicin binding to the β subunit and, hence, irrelevent in a bactericidal sense. Experiments with rifampicin probably failed to indicate translational regulation for one of two reasons: the translational regulatory mechanism is "blind" to rifampicin (i.e., there was no control response because there was no detectable perturbation), or rifampicin interferes with the regulatory mechanism.

While the effect of rifampicin on transcription of *rpoBC* appears to have little to do with a specific transcriptional regulatory mechanism, there is reason to believe that such a mechanism exists. Experiments with a bacterial strain that weakly suppresses an amber mutation in *rpoB* have demonstrated regulation at both transcriptional and translational levels (Dennis *et al.*, 1985). Experiments performed *in vitro* without rifampicin probably failed to indicate transcriptional regulation because the mechanism involves components or conditions that exist in whole cells but were absent from the *in vitro* systems. The fact that translational control has been demonstrated both *in vitro* and *in vivo* argues that RNA polymerase itself is the effector molecule. Furthermore, *in vivo* experiments have produced evidence that translational control is stronger than transcriptional control (Dennis *et al.*, 1985; Passador and Linn, 1989).

Several lines of evidence suggest that cellular levels of RNA polymerase are controlled via coordinated regulation of β and β' subunit synthesis. First, these

subunits are encoded by adjacent genes in an operon, suggesting parallel expression. Secondly, the α subunit is known to be present in excess under normal growth conditions (Bremer and Dennis, 1987), thus ruling out α subunit synthesis as the controlling step. Thirdly, Bedwell and Nomura (1986) found that overproduction of the β and β' subunits *in vivo* by introducing a plasmid carrying the *rpoB* and *rpoC* genes causes a substantial increase in synthesis of the σ subunit relative to total protein synthesis. This is consistent with the notion that σ synthesis is regulated so as to keep pace with core enzyme synthesis, but contrary to the notion that bacterial cells control RNA polymerase activity by regulating σ synthesis. Finally, and most importantly, there is ample evidence that *rpoBC* expression is specifically repressed when the concentration of RNA polymerase exceedes normal levels. Recent work by Passador and Linn (1989) has shown that overproduction of β and β' via plasmid-borne genes causes substantial reduction in the synthesis of β-galactosidase from fusions of *rpoB* to a *lacZ* reporter gene. In conjunction with the evidence that holoenzyme is the repressor molecule, this supports the idea that synthesis of β and β' is the rate determining step in synthesis of RNA polymerase.

TABLE I.

Values of Physical Parameters

Symbol	Meaning	Value	Reference
μ	specific growth rate	0.017 min^{-1} (40 min gen. time)	
$[\tau]^*$	steady state *rpoBC* mRNA concentration	1.4×10^{-8} M (5 copies per cell)	[a,b]
$[P]^*$	steady state RNA polymerase concentration	1.3×10^{-5} M	Bremer and Dennis, 1987[b]
$\sum \delta_{\tau i}$	deg. const. for full-length mRNA	0.69 min^{-1} ($t_{1/2} = 1$ min)	Pedersen and Reeh, 1978[c]
δ_p	RNA polymerase degradation constant	0	Mandelstam, 1958; Iwakura *et al.*, 1974[d]
$[P_f]/[P]^*$	free fraction of RNA polymerase	0.01	[e]
ℓ	lag time for RNA polymerase synthesis	1.3 min	Bremer and Dennis, 1987

[a]The average number of transcripts of a given kind is roughly 2.3 per cell (Neidhardt, 1987). We have used a somewhat larger value for *rpoBC* mRNA.

[b]Assuming a cytoplasmic solution volume of $6. \times 10^{-16}$ L.

[c]mRNA level was inferred from the protein synthesis rate.

[d]δ_P is so small compared to μ that it can be ignored in this work.

[e]There is reason to believe that $[P_f]/[P]^* \ll 1$ (Lohman *et al.*, 1980), but an *in vivo* measurement is not available.

Figure Captions

Figure 1.

The *rplJL-rpoBC* operon of *Escherichia coli*. Approximately 75% of the transcripts initiated at the P_{L10} promoter are terminated at the attenuator (atn) located in the intergenic region upstream of *rpoB*.

Figure 2.

Trajectories (solid curves) for expression of *rpoB* in the absence of regulation. Dashed curves refer to the hypothetical case where the protein is unrelated to RNA polymerase. (A) RNA polymerase concentration relative to the steady state concentration ($[P]^*$). (B) Total rpoBC mRNA concentration relative to the steady state concentration ($[\tau]^*$). Initial $[P]$ and $[\tau]$ values correspond to equally probable positive and negative displacements (computed by the method described in Axe and Bailey, 1990). Approximately 40% of the cells in an exponentially growing culture will experience a stochastic displacement of this magnitude during the course of a generation.

Figure 3.

Response curves for expression of *rpoB* under transcriptional regulation but in the absence of translational regulation. Dashed curves represent the unregulated response (identical to solid curves in Figure 2). (A) RNA polymerase concentration relative to the steady state concentration. (B) Total *rpoBC* mRNA concentration relative to the steady state concentration.

Figure 4.

Schematic representation of translational regulation via binding of RNA poly-

merase to a single site on the *rpoBC* mRNA. Bound mRNA is not translatable (Shine-Dalgarno sequence indicated by box). Binding is assumed to have no influence on degradation.

Figure 5.

Response curves for expression of *rpoB* under transcriptional regulation and translational regulation as represented by Figure 4. Dashed curves represent the response with transcriptional regulation only (identical to solid curves in Figure 3). (A) RNA polymerase concentration relative to the steady state concentration. (B) Total *rpoBC* mRNA concentration relative to the steady state concentration.

Figure 6.

Schematic representation of translational regulation via four-site cooperative sequential binding of RNA polymerase to *rpoBC* mRNA. Only the quadruply bound complex is untranslatable (Shine-Dalgarno sequence indicated by box). Binding is assumed to have no influence on degradation.

Figure 7.

Ψ versus RNA polymerase concentration where binding is assumed to be at equilibrium. Equilibrium binding constants, K_j, were chosen such that $K_j = \rho K_{j-1}$.

Figure 8.

Ψ versus RNA polymerase concentration where binding is assumed to be at equilibrium. Equilibrium binding constants, K_j, were chosen such that $K_4/K_1 = 10^4$ and (A) $K_3/K_2 = 10^4$, (B) $K_3/K_2 = 5 \times 10^2$, (C) $K_3/K_2 =$

2×10^1, (D) $K_3/K_2 = 1$, (E) $K_3/K_2 = 5 \times 10^{-2}$, (F) $K_3/K_2 = \times 10^{-3}$, (G) $K_3/K_2 = 10^{-4}$. In all cases $K_2/K_1 = K_4/K_3$. Curves A, B, and C are virtually indistinguishable.

Figure 9.

Ψ versus RNA polymerase concentration where steady state conditions are assumed (not equilibrium). Kinetic constants, k_j and r_j, were chosen such that $k_j/r_j = \rho k_{j-1}/r_{j-1}$. The ρ values used to generate these curves correspond to those used for Figure 7.

Figure 10.

Response curves for expression of *rpoB* under transcriptional regulation and translational regulation as represented by Figure 6. Dashed curves represent the response with transcriptional regulation and single-site translational regulation (identical to solid curves in Figure 5). (A) RNA polymerase concentration relative to the steady state concentration. (B) Total *rpoBC* mRNA concentration relative to the steady state concentration. Values of kinetic constants are as follows: (k_j in $M^{-1}sec^{-1}$) $k_1 = 1.48 \times 10^7$, $k_2 = k_3 = k_4 = 7.18 \times 10^7$, ($r_j$ in sec^{-1}) $r_1 = 33.3$, $r_2 = 15.9$, $r_3 = 1.57$, $r_4 = 0.156$.

Figure 11.

Polymerase concentration responses for each of the regulatory schemes considered (outside to inside): no regulation, transcriptional regulation only, transcriptional regulation with translational regulation via single-site binding, transcriptional regulation with translational regulation via four-site cooperative sequential binding.

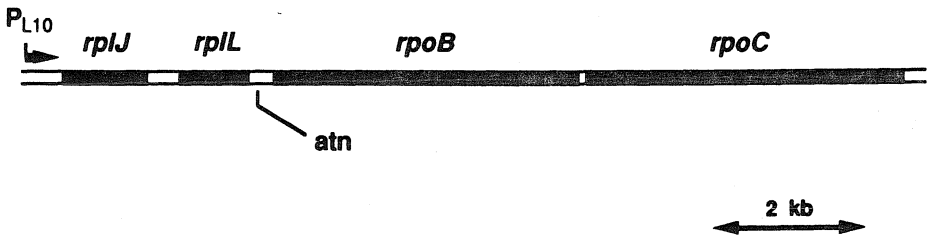

Figure 1

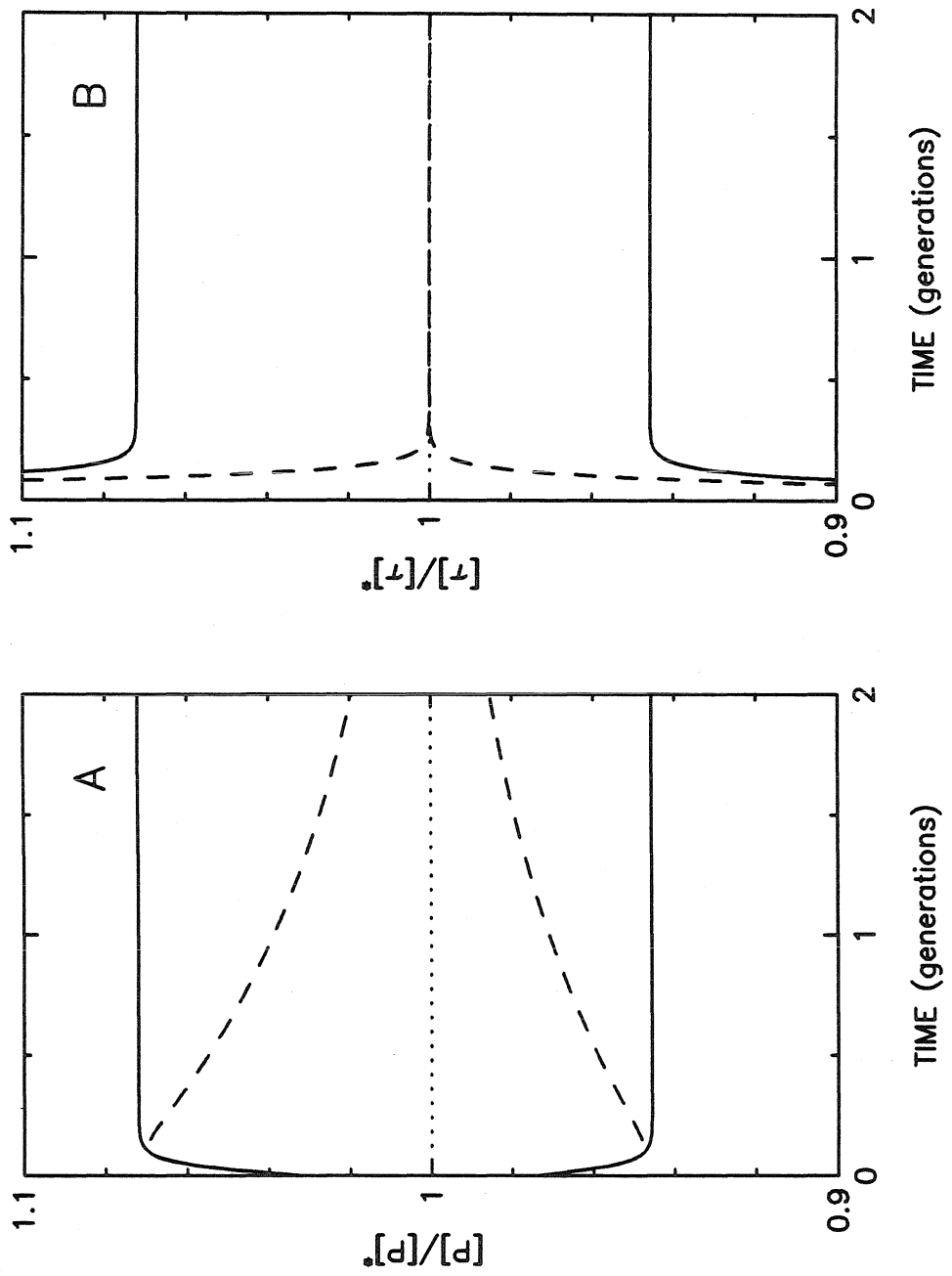

Figure 2

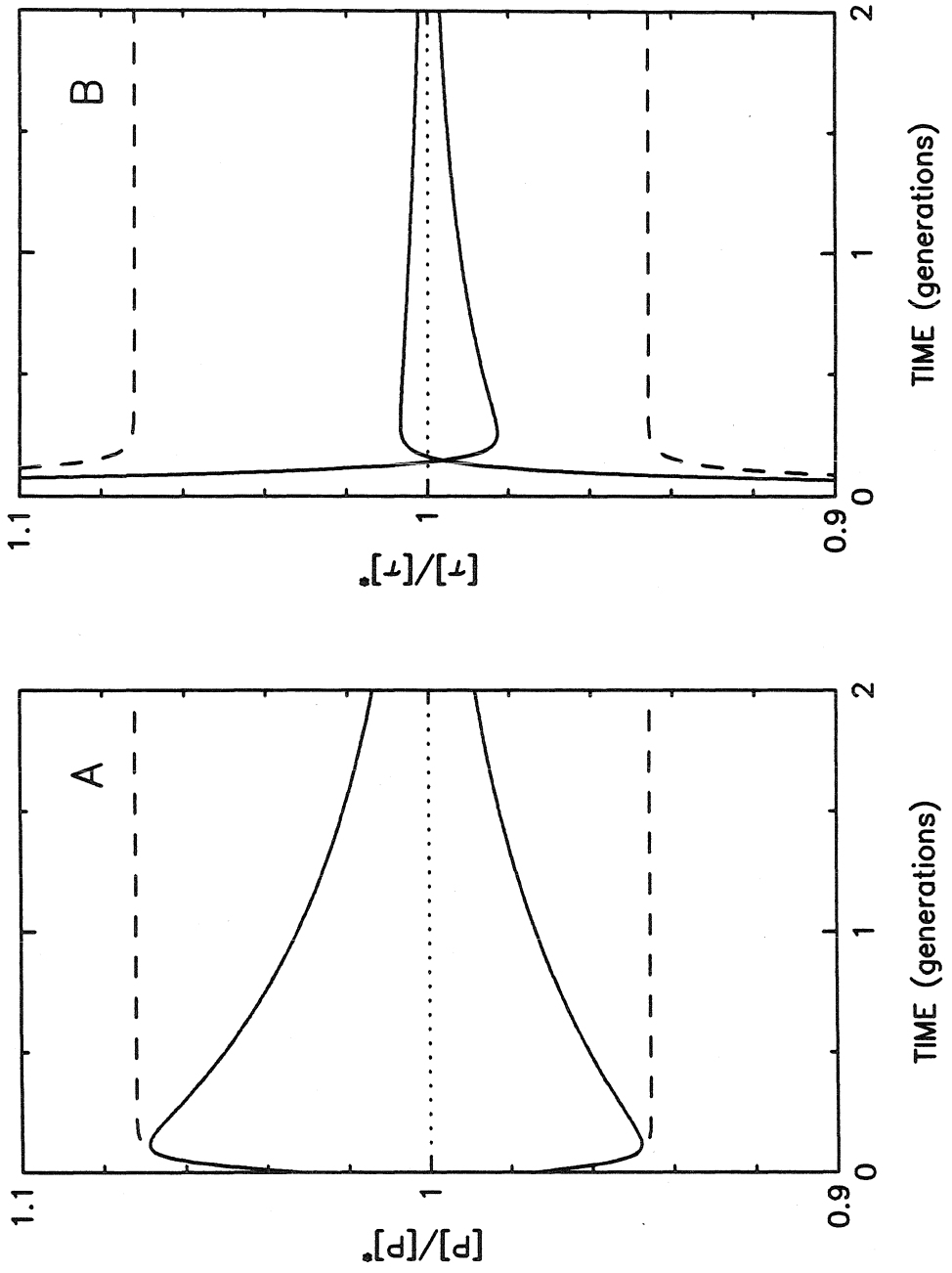

Figure 3

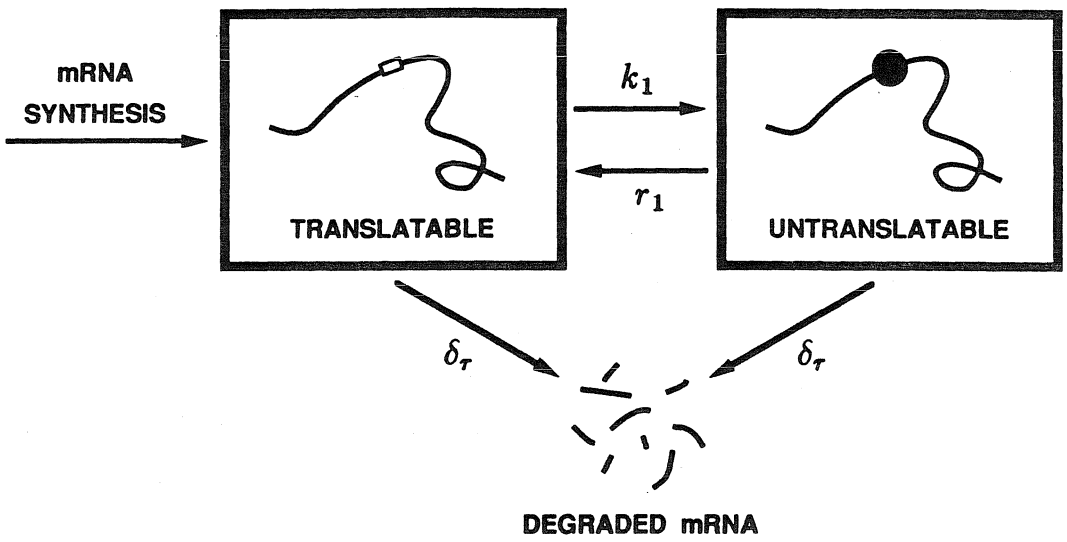
mRNA
SYNTHESIS
TRANSLATABLE
k_1
r_1
UNTRANSLATABLE
δ_r
δ_r
DEGRADED mRNA

Figure 4

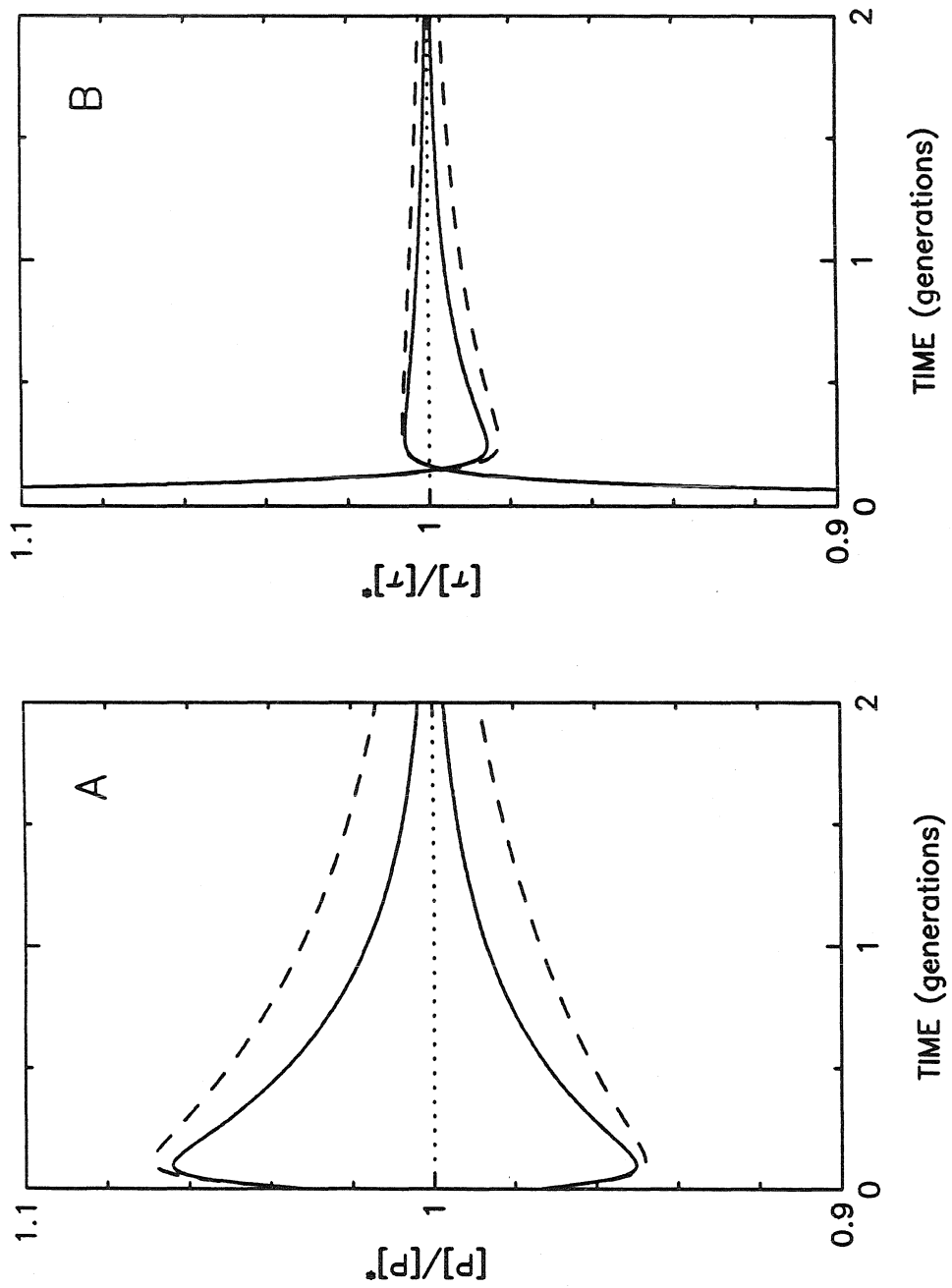

Figure 5

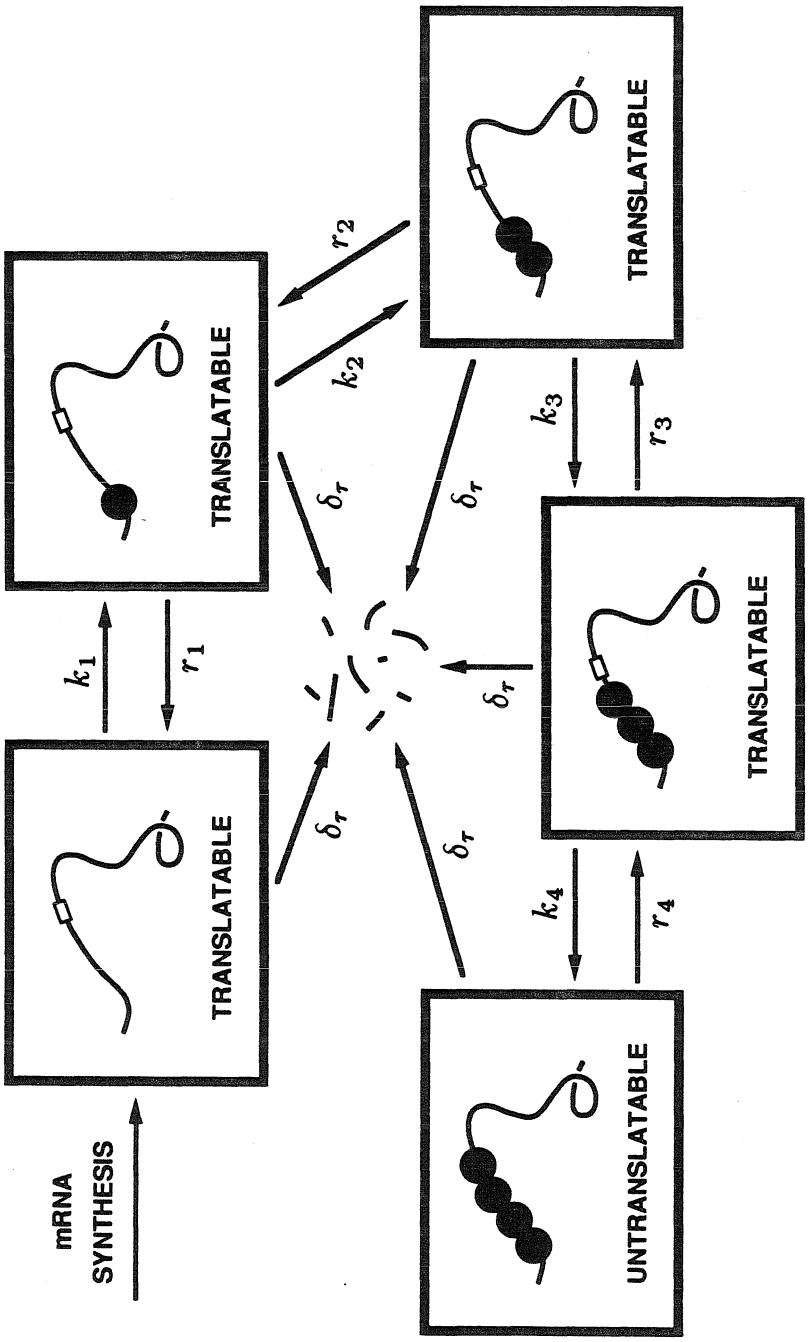
mRNA
SYNTHESIS
TRANSLATABLE
TRANSLATABLE
TRANSLATABLE
TRANSLATABLE
UNTRANSLATABLE
k_1
r_1
k_2
r_2
k_3
r_3
k_4
r_4
δ_r
δ_r
δ_r
δ_r
δ_r

Figure 6

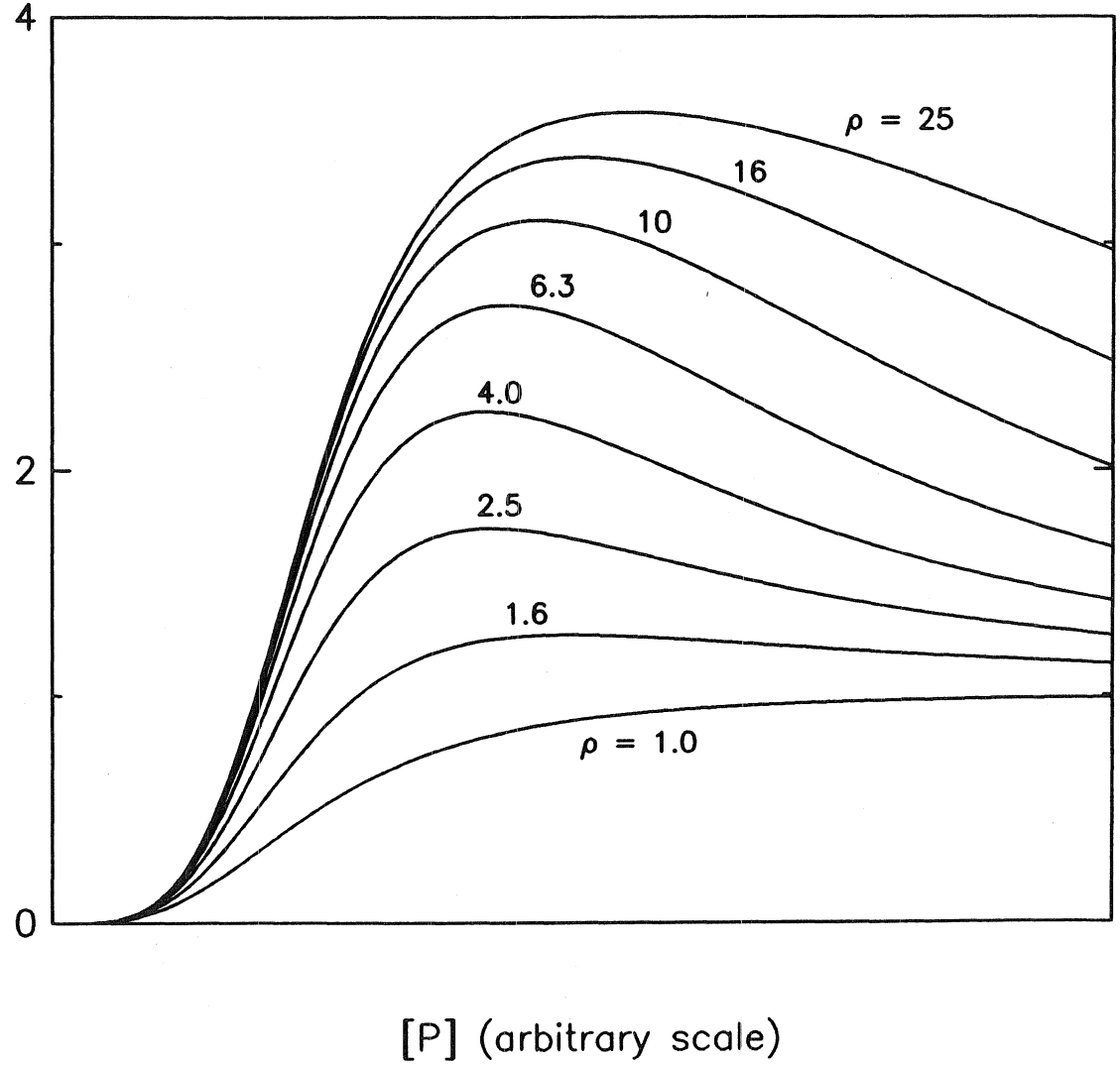

Figure 7

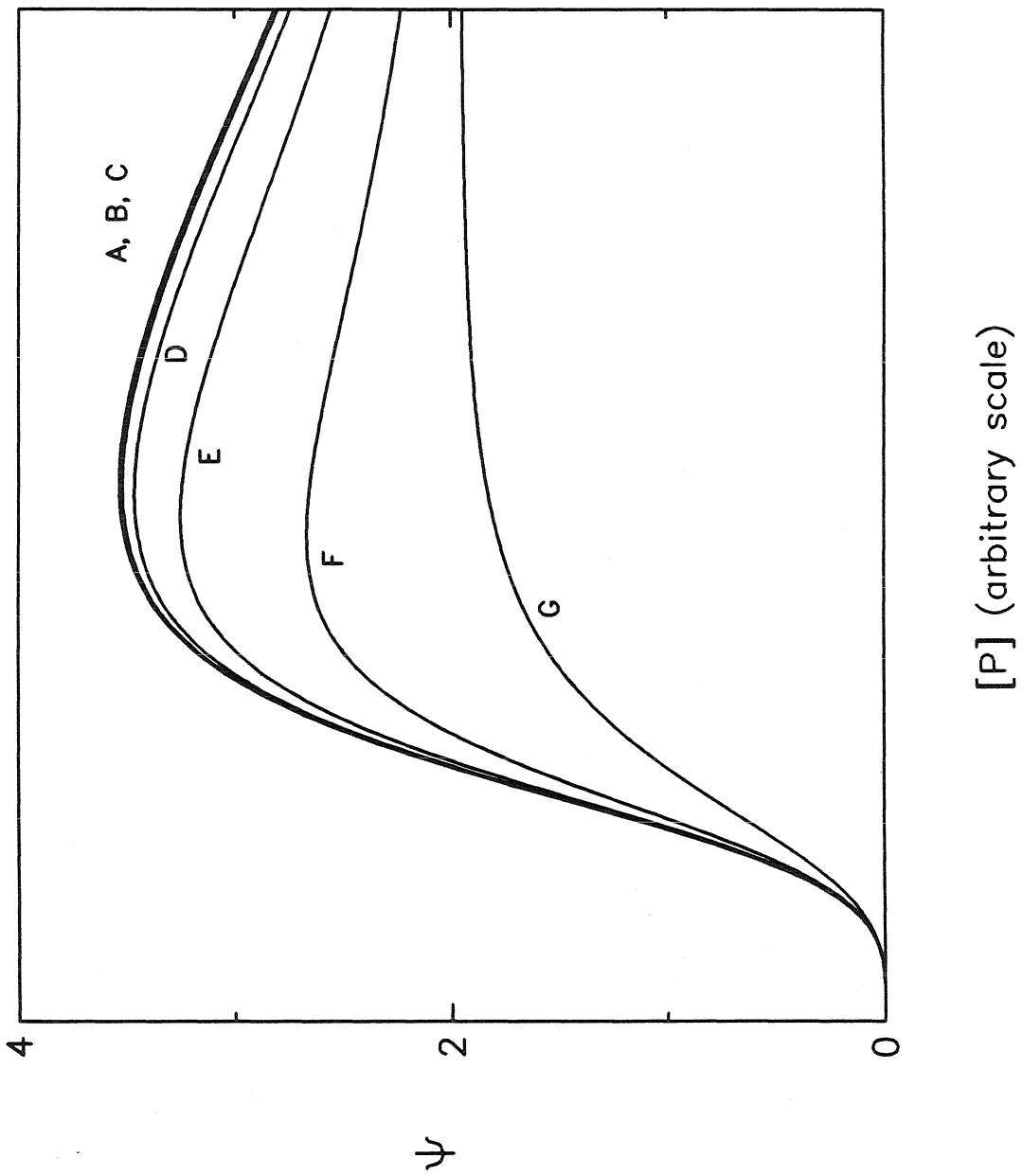

Figure 8

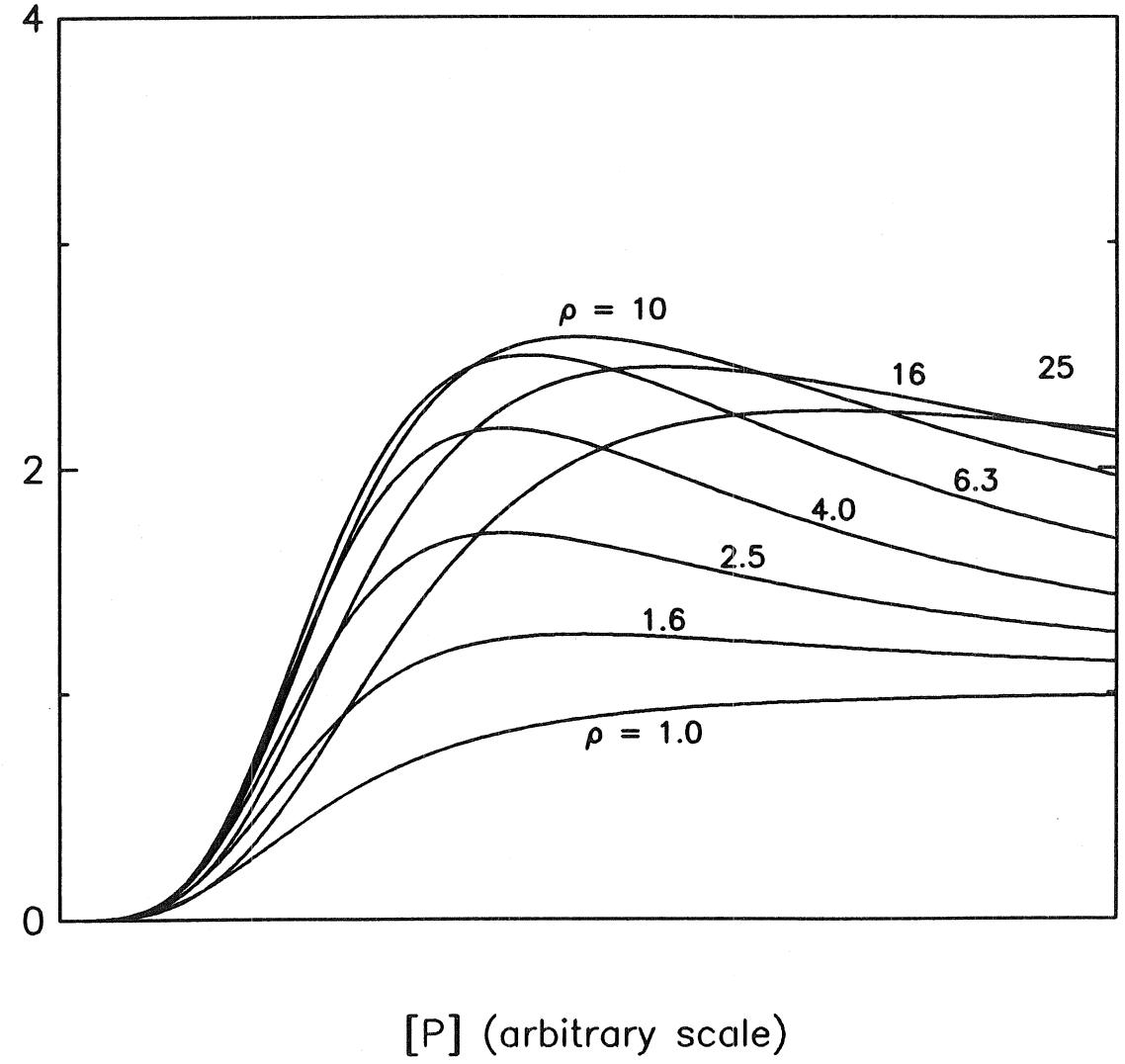

Figure 9

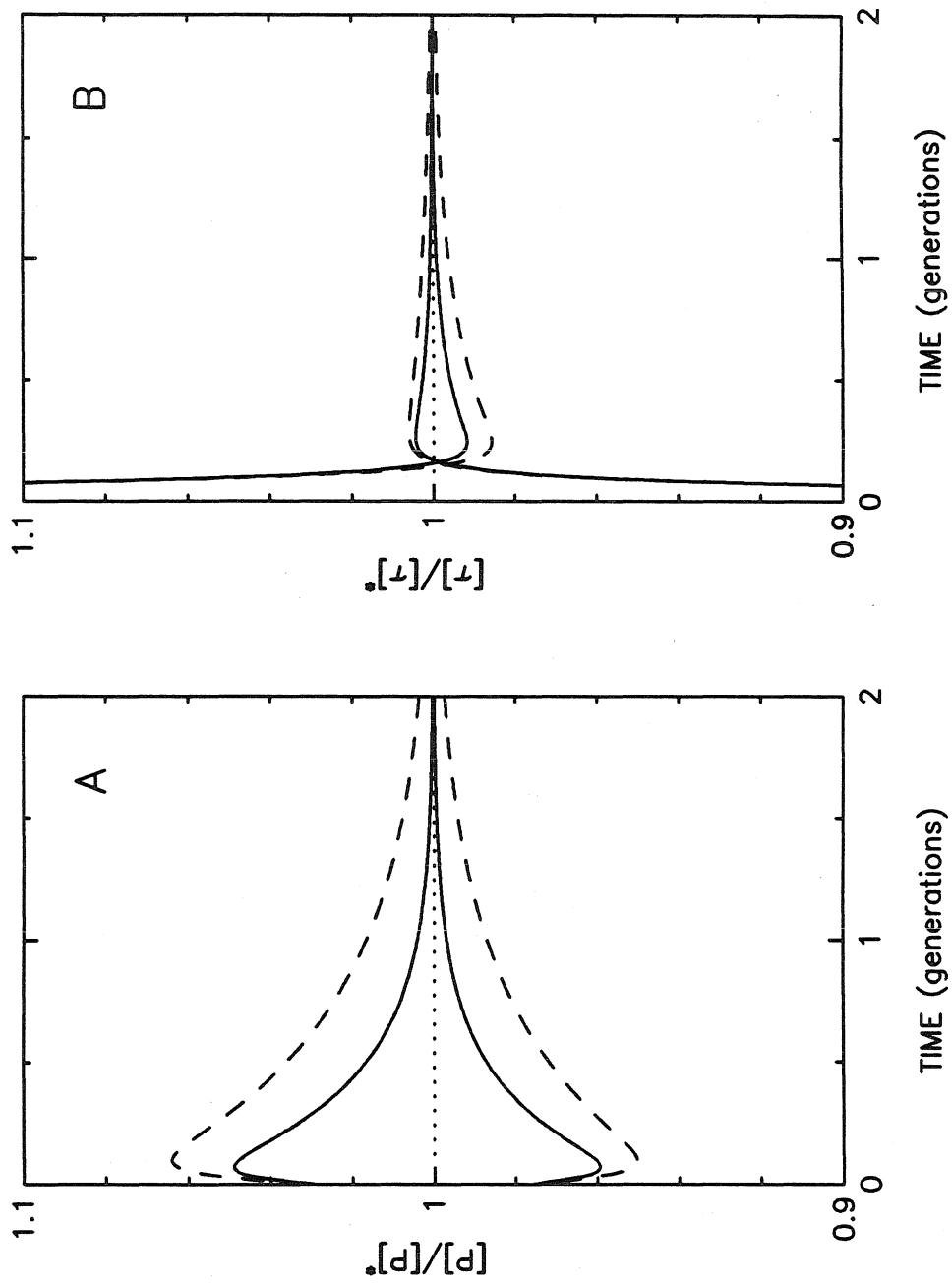

Figure 10

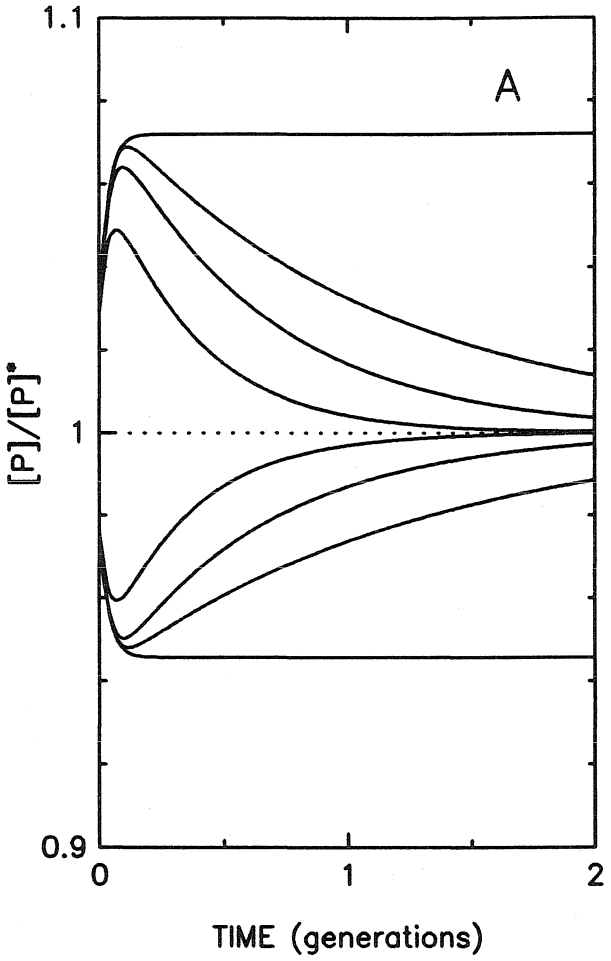

Figure 11

CHAPTER III.

EXPERIMENTAL CHARACTERIZATION OF THE INTERACTION BETWEEN RNA POLYMERASE AND *rpoBC* mRNA

1. Introduction

The theoretical work presented in Chapter II suggests that translational regulation of *rpoB* expression might be accomplished via cooperative binding of RNA polymerase to the *rpoBC* transcript. An experimental investigation of the interaction of RNA polymerase holoenzyme with portions of the *rpoBC* mRNA has been conducted to test this hypothesis. The original objectives of this work were to determine: whether RNA polymerase binds specifically to the *rpoBC* transcript, whether the "pseudo promoter" described in Chapter II is necessary for such binding, and whether multiple sites are bound in a cooperative manner. Though some success was achieved regarding the first and last objectives, questions remain as to the specificity of the observed binding. Further supporting evidence will be needed to definitively demonstrate specificity. The experiments that have been performed are described herein, along with a brief discussion of the results.

A number of approaches have been employed to detect specific interaction of proteins with nucleic acids. Among the most well established are the filter binding assay [1,2,3], the gel mobility-shift assay [4,5,6], and the footprinting technique [7,8]. The filter binding assay relies on a differential affinity of proteins and nucleic acids for nitrocellulose filters. While most proteins will bind to a nitrocellulose filter, duplex DNA (in the presence of Mg^{2+}) does not [9]. This method is probably not suitable for studying binding to RNA because single-stranded nucleic acids interact too strongly with nitrocellulose. Footprinting is based upon the principle of nucleic acid protection: the region on a nucleic acid to which a protein binds is protected from cleavage by an endonuclease and from degradation by an exonuclease. Though footprinting is also most often applied to study DNA-protein interactions, it has been used to study RNA-protein interactions as well [10]. The mobility-shift assay

takes advantage of the fact that the mobility of a protein/nucleic acid complex in a gel matrix is generally lower than the mobility of the unbound nucleic acid. Hence, upon electrophoresis, bound fragments and unbound fragments migrate in distinct bands. In principle, the mobility-shift method is equally applicable to studies of RNA-protein and DNA-protein interaction, though RNA secondary structure can potentially complicate the interpretation of results.

Of the two principal methods available for studying RNA-protein interactions, footprinting and the mobility-shift assay, the latter was deemed preferable primarily for two reasons. First, it is relatively easy to perform, and second, it is very sensitive (i.e., it can be used to detect rather labile complexes). The latter is of importance because one would expect at the outset that a protein that binds to an mRNA for the purpose of regulating translation might do so rather weakly. The half-life of bacterial mRNAs is on the order of a minute, so it seems plausible that the half-life of the regulatory complex would be even smaller (see Chapter II for discussion of this in the context of a specific model of translational control of *rpoB* expression).

2. Materials and Methods

Plasmid constructs:

$\lambda rif^{d}18$ DNA (donated by Dr. D. M. Bedwell) was digested with EcoRI. The digest was run on an agarose gel, and the 1.09 kb fragment (containing the 3′ end of *rplL*, the 5′ end of *rpoB*, and the intergenic region) was excised. This fragment (Figure 1) was ligated with DNA from an EcoRI digest of plasmid pET5 (donated by Dr. F. W. Studier). The ligation mixture was used to transform *E. coli* strain HB101. The 1.09 kb insert was found in several of the resulting clones.

Plasmid pET5 was designed for *in vitro* expression of cloned genes by T7 RNA polymerase [11]. Since our objective was to produce a transcript identical to a portion of the *rplL-rpoB* mRNA, only one of the two possible insert orientations was acceptable. Orientations were ascertained by digestion with SalI and PstI followed by gel electrophoresis. Several clones were identified with the insert in the proper orientation. One such clone was isolated for further work and designated pIRF (for: Intergenic Region, Forward orientation).

RNase precautions:

The proceedures that follow involve synthesis and handling of RNAs. Because RNase activity is ubiquitious in typical laboratory environments, precautions must be taken to preserve the integrity of RNA [12]. Wherever possible, sterile disposable plastic ware was used. All metal spatulas and glassware were baked at 375°F for a minimum of 3 hours. All solutions were made from DEPC-treated distilled water. As noted below, an RNase inhibitor ("Inhibitace," a trade name of 5 Prime → 3 Prime, Inc.) was used as a further precaution.

***In vitro* transcription:**

Due to the highly processive nature of transcription by T7 RNA polymerase, the only way to obtain a population of transcripts of uniform length is to use a linearized template. On the supposition that translational control by binding to mRNA would involve binding in the vicinity of the *rpoB* Shine-Dalgarno sequence, we restricted our investigation to a region within about 500 bases of the SD sequence. Plasmid pIRF was digested with BspHI (Figure 1) and gel purified. As a template for synthesis of the control transcript, plasmid pET5 was digested with PstI (Figure 1) and gel purified. Transcription batches (10 μL total volume) consisted of template

DNA (approximately 200 ng) in 40 mM Tris-HCl (pH 7.9), 6 mM $MgCl_2$, 10 mM dithiothreitol, 2mM spermidine, and 0.5 mM of each NTP. In addition to the above, 0.4 μL of Inhibitace, and 0.4 μL (2 μCi) of ^{3}H ATP were added. Inhibition of any residual RNase activity was allowed to proceed for 10 minutes at room temperature before addition of T7 RNA polymerase (30 units). After addition of polymerase, transcription was allowed to proceed for 40 minutes at 37°C.

Incubation with *E. coli* RNA polymerase:

In order to minimize degradation of transcript, binding experiments were performed immediately after transcription. 1.5 μL of the above transcription batch was added to each of 4 microcentrifuge tubes containing 40 mM Tris-HCl (pH 7.9), 150 mM KCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mg/mL BSA, 0.1 mM dithiothreitol, 0.4 μL of Inhibitase, 0.1 mM ATP, and 0.01 mM each of GTP, CTP, and UTP. Low levels of each nucleic acid were added because of the possibility that RNA polymerase binding is NTP-dependent. After adding different amounts of *E. coli* RNA polymerase to the 4 tubes, they were incubated at 37°C for 5 minutes.

Gel mobility-shift assay:

Immediately after incubation with *E. coli* RNA polymerase, DEPC-treated glycerol was added to the samples, and they were loaded on a 0.8% agarose gel. Samples were electrophoresed at 180 volts (13 v/cm) at a temperature of 0-4°C (the reason for reducing the temperature is discussed below). Gels contained a small amount of ethidium bromide for visual detection of RNA bands. After photographing a gel, the relevant bands were excised, melted with NaI, and mixed with "Scinti Verse" scintilation cocktail (a trade name of Fisher Scientific) for counting in a Beckman LS 5801 liquid scintilation counter.

3. Results and Discussion

Figure 2 illustrates that RNA polymerase does indeed bind to the region of *rplL-rpoB* mRNA under study. By reducing the amount of RNA polymerase added, we were able to resolve intermediate complexes (Figure 3). Shifted bands correspond to complexes between RNA polymerase and RNA. Two shifted bands are clearly visible in lane 2 of Figure 3, and further shifting observed in lanes 3 and 4 suggests that at least one additional complexation state exists. As stated in the introduction, RNA secondary structure can complicate interpretation of band patterns. Note, for example, that in the absence of *E. coli* RNA polymerase, the RNA migrates at two distinct rates (giving rise to two bands in lane 1 of Figure 3). The RNA is presumably present in two conformations that have slightly different mobilities. This raises the question as to whether the different shifted bands in Figure 3 represent different RNA secondary structures or different polymerase/transcript ratios.

The latter appears to be the case for the following reasons. First, the magnitude of the difference in mobilities of secondary structure conformations is not expected to change upon binding of RNA polymerase. This implies that secondary structure differences should cause less band separation as the overall complex mobility decreases. This is not observed in the shifted bands of Figure 3. Secondly, the relative intensities of the shifted bands clearly depend on the amount of RNA polymerase added. Furthermore, the dependence is precisely what we would expect if successive binding sites are being filled: increasing the amount of RNA polymerase causes a shift toward less mobile complexes. We would not expect RNA secondary structures to show this behavior; for a given complex stoichiometry, the relative proportions in different secondary structure states should not depend on

the amount of RNA polymerase added.

RNA polymerase was also observed to bind to the control transcript (Figure 4), though to a lesser extent. Here, three shifted bands are visible. A more quantitative means of analyzing band paterns was needed to determine whether there is a fundamental difference between binding of RNA polymerase to the control transcript and binding to the experimental transcript. In the absence of such a difference, Figures 3 and 4 do not establish that a specific interaction is occuring between *E. coli* RNA polymerase and the region of *rplL-rpoB* mRNA in the vicinity of the *rpoB* SD sequence. A quantitative analysis was performed by radiolabeling the mRNA, excising bands, and counting tritium disintegrations. The results were used to make inferences about binding mechanisms.

Before discussing results, it is informative to consider two simple alternative binding models, both of which could result in three shifted bands. First consider a transcript that has two independent binding sites, A and B. The binding reactions are then:

$$P + R \rightleftharpoons C_A$$

$$P + R \rightleftharpoons C_B$$

$$P + C_A \rightleftharpoons C_{AB}$$

$$P + C_B \rightleftharpoons C_{AB}$$

where P represents RNA polymerase, R represents unbound RNA, C_A and C_B represent the two singly bound complexes, and C_{AB} represents the doubly bound complex. When binding is at equilibrium, the following relationships hold:

$$\frac{[C_A]}{[R]} = K_A[P]$$

$$\frac{[C_B]}{[R]} = K_B[P]$$

$$\frac{[C_{AB}]}{[R]} = K_A K_B[P]^2$$

where brackets denote molar concentrations, and K_A and K_B are the equilibrium constants for binding to the respective sites.

An alternative mechanism that would produce three shifted bands consists of cooperative sequential binding to three sites. Here the binding steps are:

$$P + R \rightleftharpoons C_1$$

$$P + C_1 \rightleftharpoons C_2$$

$$P + C_2 \rightleftharpoons C_3$$

where C_i denotes the complex of RNA with i molecules of RNA polymerase. Note that the second and third sites are formed by binding to the first and second sites, respectively. In this case, binding equilibrium implies the relationships:

$$\frac{[C_1]}{[R]} = K_1[P]$$

$$\frac{[C_2]}{[R]} = K_1 K_2[P]^2$$

$$\frac{[C_3]}{[R]} = K_1 K_2 K_3[P]^3$$

where K_i is the equilibrium constant for binding of the i^{th} polymerase molecule.

Experimental values for $[C_i]/[R]$ are readily obtained by taking the ratio of counts in the appropriate shifted band to counts in the unshifted band. One convenient way to differentiate between the above binding mechanisms is to observe the change in counts for each band after adding a component that competitively

binds RNA polymerase. The effect of this component would be to reduce the concentration of freely diffusing RNA polymerase. Representing the ratio $[C_i]/[R]$ as $\Re_i$, and using $\hat{}$ to denote values obtained in the presence of a competitive binding component, we deduce the following relationships for independent binding to two sites:

$$\frac{\hat{\Re}_A}{\Re_A} = \frac{\hat{\Re}_B}{\Re_B} = f < 1$$

$$\frac{\hat{\Re}_{AB}}{\Re_{AB}} = f^2$$

$$\Re_{AB} = \Re_A \Re_B \qquad \text{and} \qquad \hat{\Re}_{AB} = \hat{\Re}_A \hat{\Re}_B$$

Furthermore, the C_{AB} complex will presumably migrate more slowly in a gel than will either of the singly bound complexes. For cooperative sequential binding at three sites, the corresponding relationships are:

$$\frac{\hat{\Re}_1}{\Re_1} = \rho < 1$$

$$\frac{\hat{\Re}_2}{\Re_2} = \rho^2$$

$$\frac{\hat{\Re}_3}{\Re_3} = \rho^3$$

In this case the distribution of counts among the complexes depends upon the values of the equilibrium constants and needn't follow any particular order. We would expect the mobility of the complexes to decrease in the order C_1, C_2, C_3.

Figure 5(a) shows experimentally determined values of $\Re$ for the experimental and control transcripts. Values on the right axes were obtained in the presence of phage λ DNA as a competitor for RNA polymerase. Qualitatively different behavior was observed for the two RNAs. The experimental RNA exhibited greater binding and greater dependence on effective RNA polymerase concentration. The

above-stated relationships were used to construct corresponding plots (Figure 5(b)) for the ideal cases of binding to two independent sites and three cooperative sites. Comparison of Figures 5(a) and 5(b) strongly suggests that binding to the experimental RNA is cooperative, and that binding to the control RNA is occuring at two independent sites.

References

1. M. Nierenberg and P. Leder *Science* **145**, 1399 (1964).

2. U. Siebenlist, L. Hennighausen, J. Battey, and P. Leder, *Cell* **37**, 381 (1984).

3. L. Hennighausen and B. Fleckenstein, *EMBO J.* **5**, 1367 (1986).

4. M. Fried and D. Crothers, *Nucleic Acids Res.* **9**, 6505 (1981).

5. M. Garner and A. Revzin, *Nucleic Acids Res.* **9**, 3407 (1981).

6. F. Strauss and A. Varshavsky, *Cell* **37**, 889 (1984).

7. D. Galas and A. Schmitz, *Nucleic Acids Res.* **5**, 3157 (1978).

8. D. Salloway, T. Kleinberger, and D. M. Livingston, *Cell* **20**, 411 (1980).

9. L. Hennighausen and H. Lubon. In: *Guide to molecular cloning techniques* (S. Berger and A. Kimmel, eds.). pp. 721-735. Academic Press (1987).

10. D. Moazed and H. F. Noller, *Nature* **327**, 389 (1987).

11. A. H. Rosenberg, B. N. Lade, D. Chui, S. W. Lin, J. J. Dunn, and F. W. Studier, *Gene* **56**, 125 (1987).

12. D. D. Blumberg. In: *Guide to molecular cloning techniques* (S. Berger and A. Kimmel, eds.). pp. 20-24. Academic Press (1987).

Figure Captions

Figure 1.

Construction of the plasmid used for *in vitro* transcription, pIRF. The 1.09 kb EcoRI fragment from $\lambda rif^d 18$ DNA (top) was inserted into the EcoRI site on plasmid pET5 (middle) to form plasmid pIRF (bottom). Run-off transcription from the T7 promoter was accomplished by cutting pIRF at the BspHI site (see Materials and Methods).

Figure 2.

Binding of RNA polymerase to the pIRF transcript. Lane 1: no RNA polymerase added. Lanes 2-4: increasing amounts of RNA polymerase added. All lanes contain the same amount of transcript.

Figure 3.

Same as Figure 2, except smaller amounts of RNA polymerase were added.

Figure 4.

Binding of RNA polymerase to the control transcript (transcript of plasmid pET5). Lane 1: no RNA polymerase. Lane 2: RNA polymerase added.

Figure 5.

(a) Plots of $\Re$ from experimental data. Values on the left axis were obtained without competitor; values on the right were obtained with competitor (phage λ DNA). On the right axes, higher $\Re$ corresponds to higher mobility. Lines are added only as a visual aid; no horizontal axis exists. (b) Plots of $\Re$ obtained from equations describing ideal three-site cooperative binding and two-site independent binding

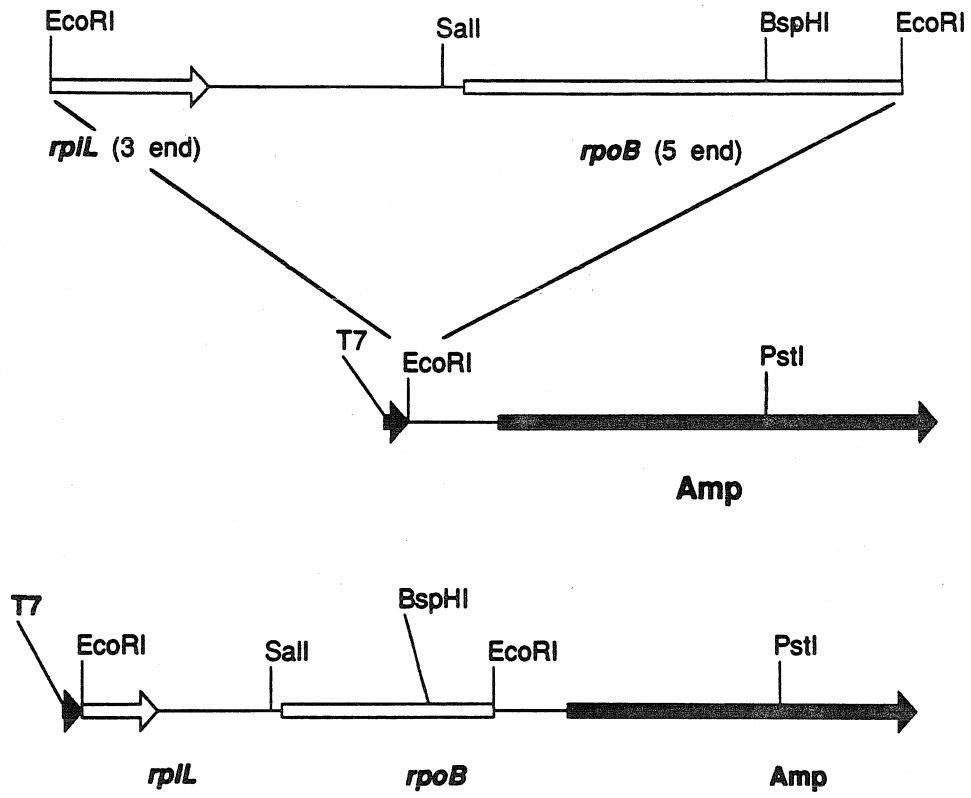
EcoRI
SalI
BspHI
EcoRI
rplL (3 end)
rpoB (5 end)
T7
EcoRI
PstI
Amp
T7
BspHI
EcoRI
SalI
EcoRI
PstI
rplL
rpoB
Amp

Figure 1.

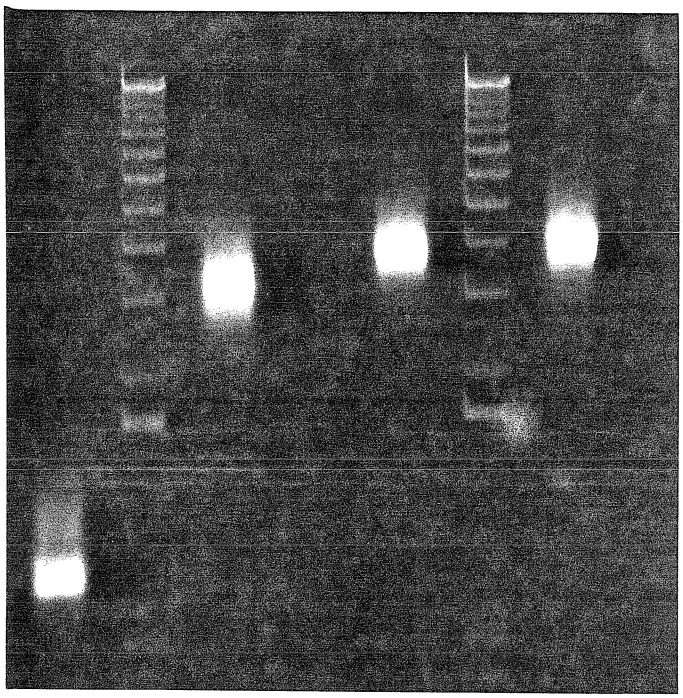

Figure 2.

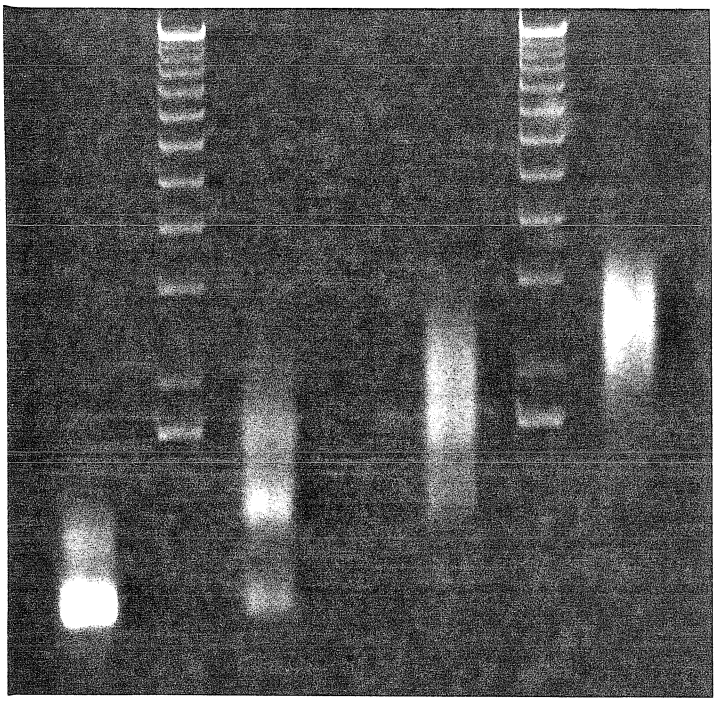

Figure 3.

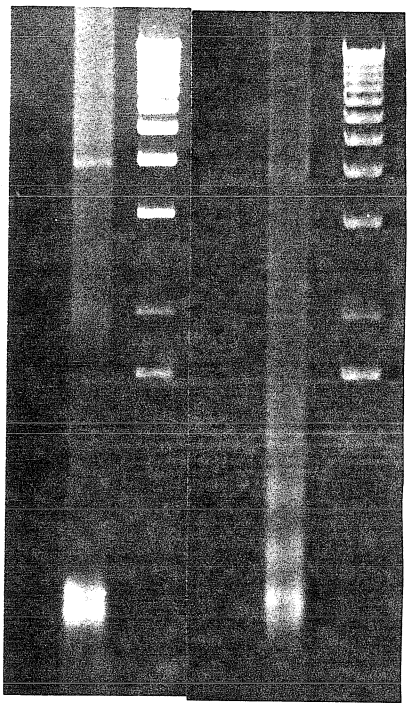

Figure 4.

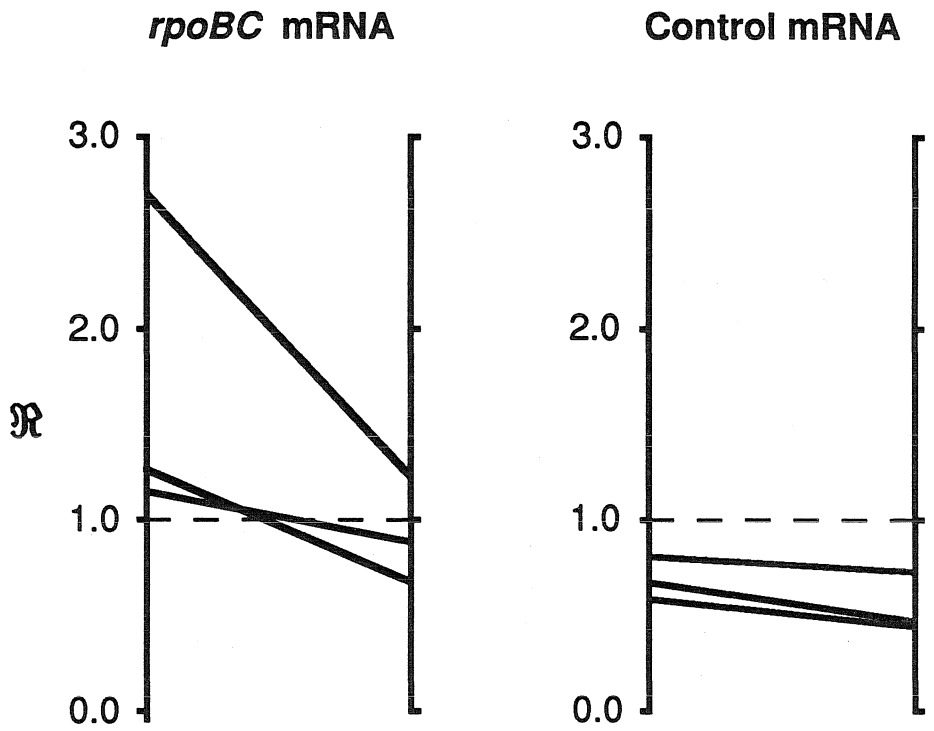

Figure 5a.

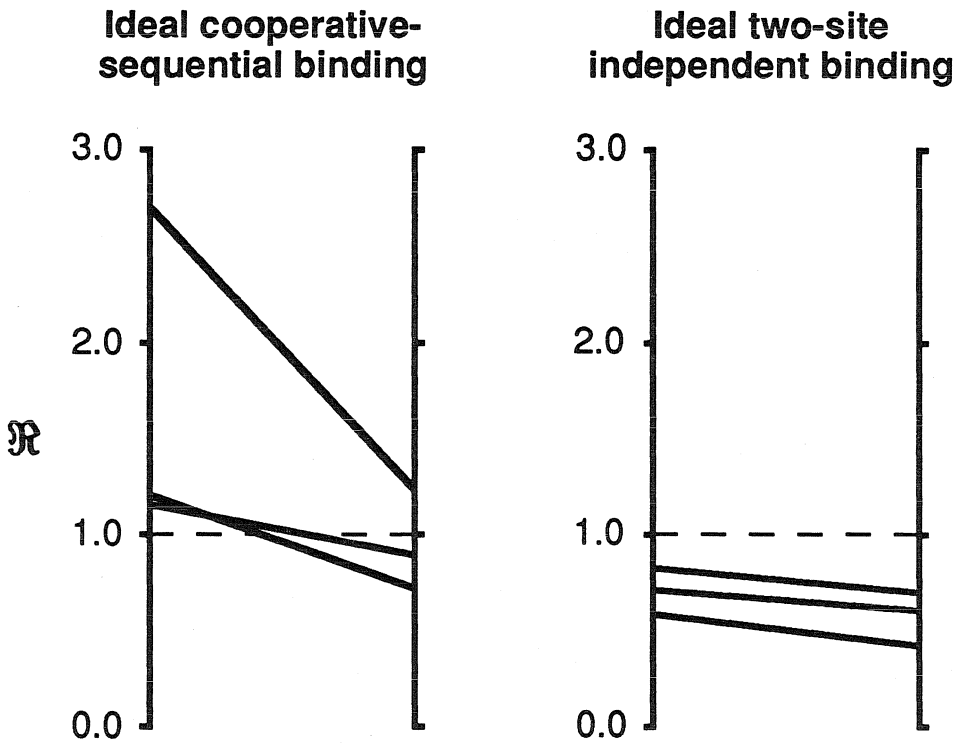
Ideal cooperative-
sequential binding
Ideal two-site
independent binding
3.0
2.0
1.0
0.0
ℜ
3.0
2.0
1.0
0.0

Figure 5b.

CHAPTER IV.

APPLICATION OF ^{31}P NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY TO INVESTIGATE PLASMID EFFECTS ON *Escherichia coli* METABOLISM

The material in this chapter has been published in *Biotechnology Letters.*

Biotechnology Letters Vol 9 No 2 83-88 (1987)
Received January 6

APPLICATION OF ^{31}P NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY TO INVESTIGATE PLASMID EFFECTS ON *Escherichia coli* METABOLISM

Douglas D. Axe and James E. Bailey*

Department of Chemical Engineering
California Institute of Technology
Pasadena, California 91125, USA

SUMMARY:

Glucose metabolism in *E. coli* strain HB101, as a plasmid-free cell and as a host to two plasmids of different copy numbers, has been characterized using ^{31}P NMR. While the low-copy-number strain was found to behave very similarly to the plasmid-free strain, dramatically different behavior was exhibited by the high-copy-number strain. This strain maintained a nearly constant intracellular pH after addition of glucose to a starved suspension while intracellular pH of the other strains dropped considerably. The inorganic phosphate level in the high-copy-number strain was substantially higher than in the other strains, and the NTP level was much lower. Glycolytic rates of all three strains, however, were nearly identical. The trend in glycolytic rate strongly suggests that transport of glucose into the cell is the rate-limiting step under these conditions.

INTRODUCTION:

Plasmid-containing cells may behave very differently from the corresponding plasmid-free cells even if the plasmid does not contain a gene coding for a metabolically active product. This is expected since general plasmid functions (plasmid replication, transcription of plasmid genes, and translation of transcripts) require cell enzymes and consume cell resources. The direct effect of this metabolic load is reduction of precursor pool levels and free enzyme levels. In some cases, the products synthesized as a result of plasmid gene expression influence cell function directly (e.g., toxic effects). If the perturbation is not too extreme, cells eventually adapt to these effects by adjusting enzyme production rates, pool levels, and growth rate until a new balanced-growth configuration is achieved. A clear understanding of the interactions between plasmids and their hosts is needed in order to predict growth and macromolecular accumulation rates for novel host-vector systems and, hence, in order to develop a rational approach for optimizing cloned-gene productivity. Since the entire metabolic network of the cell is influenced either directly or indirectly by plasmid presence, it is highly desirable to investigate host-plasmid interactions through *in vivo* experimentation.

^{31}P NMR is a powerful tool for studying *in vivo* metabolism because the phosphate group plays a central role in cell bioenergetics and carbon catabolic metabolism. Published spectra of glycolyzing *E. coli* cells show peaks caused by phosphorus resonance in nucleoside di- and triphosphates, fructose diphosphate, inorganic phosphate (P_i), and NAD^+ (Urgurbil *et al.*, 1978). Since the chemical shift of the P_i peak is a function of pH, chemical shifts of intra- and extracellular P_i peaks can be used to determine the corresponding pH values (Moon and Richards, 1973).

We have used ^{31}P NMR to investigate the effects of plasmid presence on glucose metabolism in *Escherichia coli* strain HB101. Two closely related β-lactamase-producing plasmids that differ in copy number were chosen for this experiment: plasmids pDM247 (Moser and Campbell, 1983) and RSF1050 (Heffron *et al.*, 1977). These plasmids propagate at approximately 12 and 60 copies per chromosome equivalent, respectively (Moser and Campbell, 1983). The experimental procedure

used is similar to that developed by Ugurbil *et al.* for studying bioenergetics in plasmid-free *E. coli* (Ugurbil *et al.*, 1978). After addition of glucose to a starved cell suspension, spectra were collected at ~4 minute intervals in order to characterize metabolism of the three strains.

MATERIALS AND METHODS:

Experiments were performed with plasmid-free *E. coli* HB101 *pro gal hsdR hsdM recA1* and with two transformed strains consisting of plasmid pDM247 in HB101 and plasmid RSF1050 in HB101. Cells, grown aerobically at 37°C in M9 minimal medium supplemented with L-leucine (41 mg/L), L-proline (164 mg/L), and thiamine HCl (0.166 mg/L), were harvested in exponential-growth phase. Cells were placed on ice before centrifuging. Aerobic conditions were maintained during the cooling process by bubbling O_2 through the cultures. Cells were washed twice with a buffer solution consisting of 100 mM piperazinediethanesulfonic acid (Pipes), 50 mM 4-morpholinoethanesulfonic acid (Mes), 10 mM Na_2HPO_4, 10 mM KH_2PO_4, and 40 mM NaCl (adjusted to pH=7.30 with NaOH), and resuspended in a volume of buffer equal to the pellet volume. Samples (2 mL) were placed in 10 mm diameter NMR tubes and stored on ice until used (less than 2 hours).

In order to minimize sample degradation, the experimental sample was kept on ice while a dummy sample containing 10% D_2O was used for shimming (Shanks and Bailey). Samples were warmed to 20°C and a single spectrum was accumulated. A 100 μL aliquot of 640 mM glucose was then added, and spectra were accumulated at ~4 minute intervals for an hour. Anaerobic conditions were maintained by flushing the tube with N_2 throughout the experiment. Spectra were obtained in the Fourier-transform mode on a Bruker WM-500 spectrometer operating at 202.46 MHz. Spectra represent 3.7 minutes of data acquisition (400 scans at a 0.5s repetition time) using 70° pulses.

Glycerophosphorylcholine (GPC) was used as an internal standard (+0.49 ppm relative to 85% phosphoric acid, upfield shifts given a negative sign). Peak resonance assignments were obtained from the literature (Ugurbil *et al.*, 1978). Intra- and extracellular pH values were determined by using a P_i titration curve (P_i chemical shift vs. pH) obtained by adding acetic acid, in varying amounts, to the standard pH=7.30 buffer solution and collecting a spectrum for each sample. The extracellular volume in the NMR samples was determined by adding a measured quantity of blue dextran (avg. mol. wt. = 2,000,000) to an identical sample, centrifuging, and measuring the absorbance of the supernatant at 620nm. This was used for accurate determination of extracellular P_i concentrations.

RESULTS AND DISCUSSION:

Before attempting to interpret spectral information, it is important to make some general deductions about the state of the cells being investigated. Prior to glucose addition, cells are starved (low energy charge, not glycolyzing), but they are in good condition as evidenced by the speed of their response to glucose addition. Because the cells are suspended in a buffer that lacks both a nitrogen source and the amino acids for which HB101 is auxotrophic, one would expect transcription, translation, and replication rates to be very low. Glycolysis and associated transport processes, including ATP-driven proton translocation, are expected to be the dominant activities following glucose addition. Thus, the pool-depleting effects of plasmid presence should be small. Copy number effects observed under these experimental conditions will be due either to adaptations of the metabolic framework made during exponential growth under the metabolic load of plasmid-directed activities, or to direct effects of plasmid-coded products.

In order to emphasize copy numbers, the three strains: HB101, HB101:pDM247, and HB101: RSF1050, will be denoted by P0, P12, and P60 respectively. Note, however, that these reference copy numbers were determined for cells grown in LB medium (Moser and Campbell, 1983) and therefore represent relative values rather than exact ones for this experiment. The spectra show little difference between P0 and P12. Substantial differences are apparent, however, between P0 and P60 (Figure 1). NTP levels were much lower in P60 throughout the experiment (77% lower at 20 min.). Integrals of the NTP_γ-NDP_β region at -5 ppm indicate that this is due in part to lower nucleotide levels in P60 and not merely to a lower ratio of [ATP] to [ADP]. It is likely, then, that this is one of the adaptative responses to plasmid presence. The most obvious difference between the spectra shown in Figure 1 is in the P_i peak region, ~1 to ~3 ppm. P0 exports inorganic phosphate into the medium to a much greater extent than P60; the extracellular P_i concentrations were determined to be ~43 and ~21 mM, respectively (using 68% (v/v) for the medium fraction as measured by blue dextran exclusion). These remain essentially constant from the time that the two P_i peaks can first be distinguished. Intracellular P_i concentrations, estimated by assuming

that the free cytoplasmic volume constitutes 65% of the gross cell volume, are roughly 41 and 103 mM respectively for P0 and P60. While in P60 essentially all intracellular P_i was retained (buffering medium is prepared with 20 mM P_i), in P0 the intracellular P_i level dropped to the level of the medium. However, since the electrical potential difference across the membrane was not measured, one cannot conclude that P_i achieved chemical equilibrium across the membrane in P0. By finding the ratios of the relevant peak areas from P0 and P60, we estimate that ΔG for ATP hydrolysis is 1 kcal/mole higher in P60. Since the magnitude of ΔG is estimated to be roughly -18 kcal/mol, this difference between strains is not expected to have a significant effect.

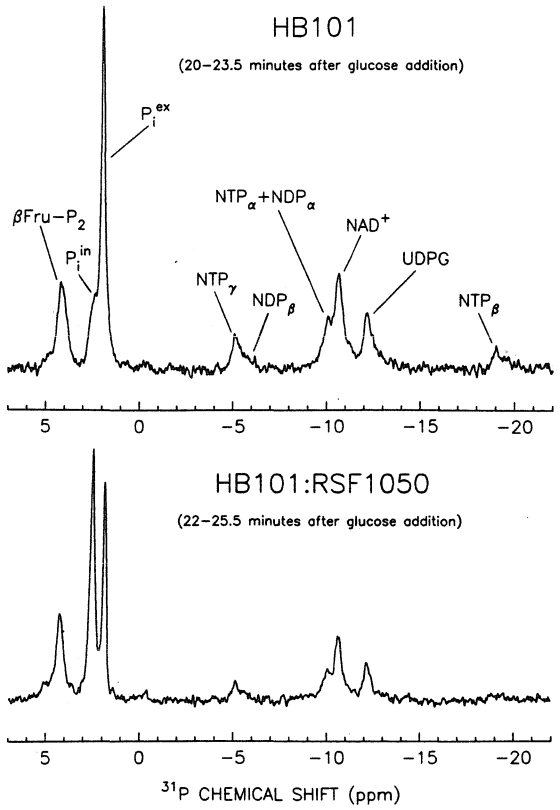

Figure 1: ^{31}P NMR spectra for glycolyzing *E. coli* HB101 (P0) and HB101:RSF1050 (P60); 400 scans, 0.5 sec. repetition time, using 70° pulses at a frequency of 202.46 MHz. Glucose was added to a final concentration of 44 mM. Chemical shifts are referenced to 85% phosphoric acid.

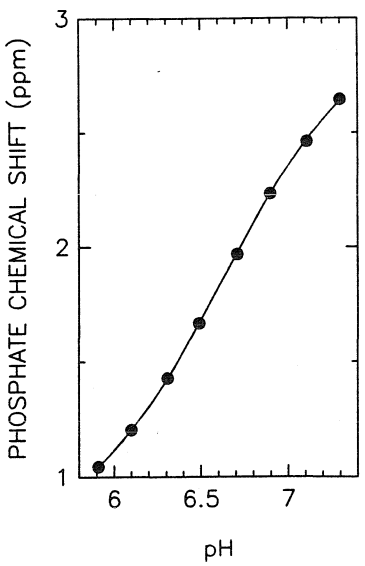

Figure 2: P_i chemical shift titration curve for the buffering medium described in *Materials and Methods* (NMR parameters as stated for Figure 1).

As glycolysis procedes, H^+ is produced and exported (via membrane-bound ATPase) to the medium, causing a steady decrease in the extracellular pH and a consequent upfield shift in P_i^{ex} resonance. A titration curve (Figure 2) was used to determine intra- and extracellular pH values from P_i chemical shifts. The curve was obtained by collecting spectra for samples prepared by adding glacial acetic acid to the same pH=7.30 buffering medium used to suspend cells (lactic acid gave identical results). The curve so obtained may be less accurate for determination of intracellular pH than for extracellular pH because the chemical shift of P_i is a function of ionic strength as well as pH. The dependence on ionic strength is relatively weak, however; at a pH of 6.30, a 160mM decrease in the concentration of singly charged ions was found to have the same effect on the chemical shift of P_i as a 0.05 unit decrease in pH. Since the ionic strength of buffering medium is close to that of the cytoplasm, the curve in Figure 2 is adaquate for intracellular pH determination. Intra- and extracellular pH data are presented in Figure 3. While broad intracellular P_i peaks for P0 and P12 lead to some scatter, extracellular P_i peaks for all strains and the intracellular P_i peak for P60 have well-defined chemical shifts and therefore give well-defined pH trajectories. The

pH of the external medium follows very nearly the same path for each of the three strains. P60 maintains a nearly constant intracellular pH of ~7.1. Intracellular pHs of the other strains jump to ~7.2 at about 15 minutes after glucose addition and then decrease in such a way that a roughly constant ΔpH is maintained. From approximately 20 minutes after glucose addition onward, P60 maintains a larger pH gradient across its cytoplasmic membrane than do P0 and P12.

Typical titration curves (pH vs. acid added) were obtained by adding concentrated lactic acid of known concentration to the pH-adjusted buffering medium. These were used to obtain the net amount of H^+ exported versus time from the pH data. In order to account for differences in extracellular P_i buffering capacity, media were prepared to match the P_i^{ex} of the three strains and separate curves were obtained for each. Time derivatives of the proton export data are presented in Figure 4. From ~10 minutes onward it is evident that the net proton export rate follows the same path in each strain. Since medium acidification requires transport of glucose into the cell, catabolism, and export of end products, one would expect a lag between the times of glucose addition and initial acidification. While the experimental procedure described herein does not afford the time resolution that would be needed to accurately measure this lag time (all three strains export *some* H^+ within 2 minutes of glucose addition), it does allow observation of the start-up time required before the maximum glycolytic rate is achieved. P0 and P12 attained maximum rates within 6 minutes of glucose addition while P60 required ~9 minutes. Assuming that P_i (pK_a=7.2) is the dominant buffering species in the cytoplasm, one can derive an approximate cytoplasmic titration curve, and from that, estimate the extent of acid accumulation in the cell. This calculation indicates that ~95% of the H^+ formed by glycolysis is exported to the medium.

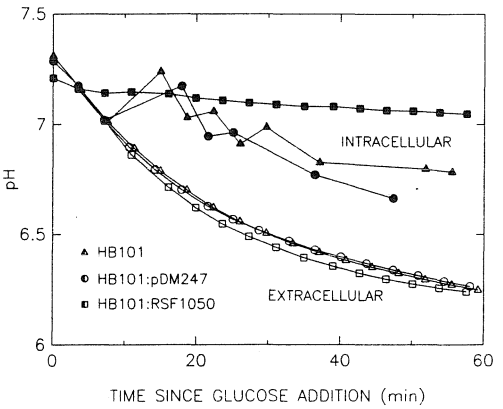

Figure 3: Time evolution of intracellular (solid symbols) and extracellular (open symbols) pH as determined from P_i chemical shifts and Figure 2. Intra- and extracellular values coinside for the first few data points.

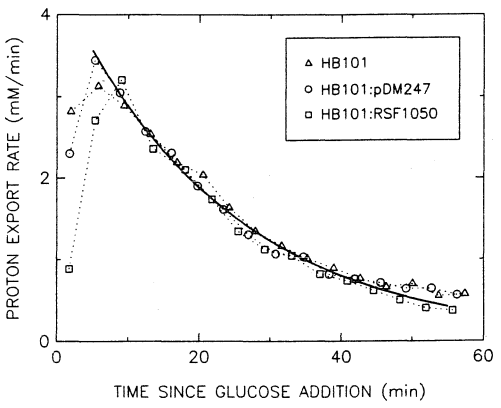

Figure 4: Time evolution of proton export rate (determined as discussed in text). The solid curve is an exponential decay with parameters chosen to fit the data.

Under anaerobic conditions, glycolyzing *E. coli* produce a mixture of ethanol, CO_2, and acetic, lactic, formic, and succinic acids. Lactic acid, the dominant product over a wide pH range (Doelle, 1975), liberates 0.33 protons per carbon atom at any physiological pH. The product proportions reported by Doelle for a medium pH of 6.2 (Doelle, 1975) yield 0.32 protons per carbon atom. At an extracellular pH of 6.6 (more typical of this experiment), ~0.34 protons are liberated per carbon atom metabolized. This suggests, with regard to proton production, that the following is a reasonable representation of the net reaction for glycolysis in this experiment:

$$\text{glucose} + 2HPO_4^{2-} + 2ADP \longrightarrow 2ATP + 2\text{lactate} + 2H_2O \qquad (1)$$

This could be written equally well by replacing $2HPO_4^{2-}$ with $2H_2PO_4^-$ and adding $2H^+$ to the right-hand side. The actual stoichiometry depends on the pH under which the reaction occurs and lies somewhere between these two representations. Thus, in general, H^+ production associated with glycolysis is a function of pH.

Because a representation of the net reaction for glycolysis is the sum of many reaction steps, it reflects the true *in vivo* stoichiometry only during (pseudo-) steady-state glycolysis. If any glycolytic intermediate accumulates at a rate comparable to the rate of its formation, glycolytic end-product formation will occur at a lower rate than is suggested by net reaction stoichiometry. We will assume that glycolysis exhibits pseudo-steady-state behavior under the conditions of this experiment. This assumption will be examined shortly. If we also assume pseudo-steady-state behavior for the ATP level, we conclude that ATP hydrolysis by ATPase occurs at the same rate as ATP production by glycolysis. The net reaction occuring in the cell is then:

$$\begin{aligned} &\text{glucose} + 2HPO_4^{2-} + 2ADP \longrightarrow 2ATP + 2\text{lactate} + 2H_2O \\ &+ \qquad 2(ATP + H_2O \longrightarrow ADP + HPO_4^{2-} + H^+) \\ &\hline \\ &\textbf{overall:} \qquad \text{glucose} \longrightarrow 2\text{lactate} + 2H^+ \end{aligned} \tag{2}$$

Here H^+ is produced at twice the rate that glucose is consumed for any physiologically attainable pH. Since H^+ is almost completely exported, this means that if the pseudo-steady-state assumptions are valid, the curve in Figure 4 also represents 2 × glycolysis rate versus time as well as the rate of ATP production versus time (medium basis; i.e. mmoles per liter of *medium* per minute).

The rate of glycolysis is known to be affected by ATP and AMP levels via their modification of phosphofructokinase (PFK) activity. ATP, when present in abundance, reduces the activity of this enzyme by binding to it, and allosterically modifying the catalytic site. PFK catalyzes phosphorylation of fructose 6-phosphate (F6P) to form fructose 1,6-diphosphate (FDP). Since feedback repression can directly affect the rate of this reaction step, F6P is the most likely glycolytic intermediate to accumulate within the cell. The sugar phosphate peak at ~4 ppm results from the overlap of phosphate resonances in F6P, G6P, and FDP. From ~10 minutes after glucose addition on, we observe only gradual changes in the sugar phosphate region. When the maximum rate of proton export is attained, sugar phosphate peak areas in all three strains are still increasing despite the apparent end of the glycolytic start-up period suggested by Figure 4. Integrals of the sugar phosphate region reach a maximum at ~11 and ~16 minutes for P0 and P60, respectively, and then gradually decrease to ~80% of the peak levels by the end of the experiment. Using the maximum FDP concentration (13 mM) reported by Ugurbil *et al.* (Ugurbil *et al.*, 1978) and making the conservative assumption that all changes in the sugar phosphate region are due to changes in F6P concentration (an equal change in FDP concentration has twice the effect on peak area), we estimate that after proton export rates have peaked, the maximum rates of F6P accumulation are ~0.34 and ~0.11 mM/min (medium basis) for P0 and P60 respectively. If Figure 4 is interpreted as 2 × glycolytic rate versus time, these represent ~23% and ~8% of the glycolytic rate. From ~11 minutes onward for P0 and ~16 minutes onward for P60, F6P levels drop at ~0.03 mM/min (medium basis). Thus, through most of the experiment the assumption of pseudo-steady-state glycolysis seems to be justified.

Spectra from all three strains show that ATP levels reach a maximum within ~5 minutes of glucose addition and then slowly drop throughout the experiment. The ATP accumulation rate, which is equal for nongrowing cells to the rate of ATP production minus the rate of of ATP utilization, is estimated to be ~-0.03 mM/min (medium basis) assuming a maximum [ATP] of 7 mM (Ugurbil *et al.*, 1978). If Figure 4 is interpreted as ATP production rate versus time, the rate of ATP production is seen to be much greater than the rate of accumulation. ATP utilization must approximately balance ATP production, indicating that [ATP] follows pseudo-steady-state behavior as assumed. Since the assumptions leading to (2) above are consistent with the data (start-up period excluded), this appears to be a reasonable representation of the net reaction occuring in the cells in this experiment.

Figure 4 then indicates that the glycolytic rates among the three strains are very similar throughout the experiment. Furthermore, the data indicate that the overall reaction has first-order kinetics. The solid curve in Figure 4 is an exponential decay with parameters chosen to fit the data. If the peak ATP level reached ~5 minutes after glucose addition is high enough to

inhibit phosphofructokinase activity, one would expect the ATP level and the rate of glycolysis to be maintained at nearly constant values until the extracellular glucose concentration drops to a level that makes glucose transport the rate limiting step. At this point the glycolytic rate would decrease rapidly. The fact that the glycolytic rate drops sharply immediately after it peaks indicates that the energy charge is sufficiently low that the enzyme is uninhibited. In other words, the cell's energy demands exceed the energy supplied by glycolysis.

In *E. coli*, glucose is phosphorylated by the phosphoenolpyruvate (PEP):sugar phosphotransferase system as it enters the cell (Romano *et al.*, 1970). If the cytoplasmic level of glucose-6-phosphate in an NMR sample is sufficiently high, cells can presumably glycolyze rapidly enough to meet energy requirements. Since the energy demand is not being met and G6P is always observed to comprise a small fraction of the total sugar phosphate pool (the G6P peak is at the extreme left of the sugar phosphate region), we conclude that transport of glucose into the cell is the rate limiting step. The apparent first-order overall kinetics further evidence this because one would expect glucose-transport limitation to yield such kinetic behavior. Using 1 μm as the characteristic length for diffusion of glucose through the medium, and a typical aqueous diffusivity of 1×10^{-5} cm^2/s, one finds the characteristic timescale for diffusion to be of the order of milliseconds. Since this is many orders of magnitude smaller than the timescale for glycolysis ($\sim 10^3$ seconds), the extracellular medium can be treated as homogeneous. If the PEP:sugar phosphotransferase sytem has first order kinetics with respect to glucose, as would be expected for non-saturating glucose concentrations, then the rate of glucose uptake will decay exponentially with time. Our observations are consistent with this interpretation. We have constructed a cell model that incorporates the relevant features of anaerobic glucose metabolism in *E. coli* as deduced from NMR data (Axe and Bailey). This formulation sheds considerably more light on the metabolic causes of the observed behavior than can be presented here.

Because plasmids pDM247 and RSF1050 are highly homologous, it is reasonable to conclude that the different behavior exhibited by P12 and P60 is a consequence of the difference in copy number. The strong similarity between P0 and P12 suggests that pDM247 is maintained at a sufficiently low copy number that its presence causes only a small perturbation in host metabolism. We are presently unable to explain the unique behavior found for P60 in terms of plasmid activities. It seems unlikely that plasmid presence would typically induce the effects observed in P60. Perhaps the plasmids selected for this experiment are atypical in some sense. We are currently investigating the possibilty that the behavior exhibited in P60 is common to cells that over-produce β-lactamase.

ACKNOWLEDGEMENTS: This work was sponsored by the National Science Foundation. NMR experiments described herein were performed at the facilities of the Southern California Regional Nuclear Magnetic Center (NSF Grant No. CHE-84-40137). The authors wish to express their gratitude for assistance offered by the personnel at this facility and for helpful advice offered by J.L. Galazzo and J.V. Shanks. Software for NMR spectral analysis was provided by the NIH Resource Laboratory at Syracuse University (Grant No. RR-01317).

REFERENCES:

Axe, D.D., and Bailey, J.E. (manuscript in preparation)

Doelle, H.W.(1975). *Bacterial Metabolism* (2nd ed.), New York: Academic Press

Heffron, F., Bedinger, P., Champous, J.J., and Falkow, S.(1977). *PNAS(US)* **74**, 702-706

Moser, D., and Campbell, J.(1983). *J. Bacteriol.* **154**, 809-818

Perlin, D.S., Kasamo, K., Brooker, R.J., and Slayman, C.W.(1984). *J. Biol. Chem.* **259**, 7884-7892

Perlin, D.S., San Francisco, M., Slayman, C.W., and Rosen, B.P.(1986). *Arch. Biochem. Biophys.* **248**, 56-61

Romano, A.H., Eberhard, S.J., Dingle, S.L., and McDowell, T.D.(1970). *J. Bacteriol.* **104**, 808-813

Shanks, J.V., and Bailey, J.E. (manuscript in preparation)

Ugurbil, K., Rottenberg, H., Glynn, P., and Shulman, R.G.(1978). *PNAS(US)* **75**, 2244-2248

Ugurbil, K., Shulman, R.G., and Brown, T.R.(1979). High-resolution ^{31}P and ^{13}C nuclear magnetic resonance studies of *Escherichia coli* cells *in vivo*. In: *Biological Applications of Magnetic Resonance*, R.G. Shulman, ed. pp. 537-589, New York: Academic Press

CHAPTER V.

UNIDIRECTIONAL FLUX CONSTRAINTS FOR PROTEIN-MEDIATED TRANSPORT ACROSS CELL MEMBRANES

1. Introduction

The transport of solutes across membranes is an essential process in all living organisms. Biological transport processes can be divided into two classes: those that involve simple diffusion across a permeable membrane form the first class, and those where transport occurs only at specific structures within the membrane (eg., pores, channels, and carriers) form the second. The first class shares commonality with transport across nonbiological membranes, diffusion across phase boundaries being the characteristic phenomenon. The second class differs greatly in that the action occurs at a complex protein structure that cannot accurately be termed a *phase*. In fact, the lipid bilayer itself (which is entirely responsible for stabilizing the membrane) contributes only passively by holding the protein structure in place and by being essentially impermeable to the solute.

Owing to its similarity to well-studied interfacial transport phenomena, a great deal is known about the first class. In contrast, our understanding of the second class is as rudimentary as our understanding of protein structures and mechanisms. Nonetheless, thermodynamic principles as well as certain kinetic ones can be applied, even when the mechanism of transport is largely unknown. Our objective here is to apply these principles to the case of protein-mediated transport, the specific aim being to discern the relationship between opposing unidirectional fluxes.

2. The Ussing-Teorell Equation

Consider a single solute species, A, the transport of which is independent of other species. If a membrane separates two solutions of A and is permeable to A, the net flux of A from side 1 to side 2 is simply the difference between the two

unidirectional fluxes. Knowledge of the unidirectional fluxes is obviously of greater value from a kinetic standpoint than knowledge of the net flux. For the case of simple diffusion, the ratio of unidirectional fluxes is given by the Ussing-Teorell (or flux ratio) equation (Nobel, 1974; Ussing, 1949; Teorell, 1949):

$$\frac{J_{12}}{J_{21}} = \frac{a_{A1}}{a_{A2}} e^{-\alpha z \Delta \Psi} \tag{1}$$

where J_{ij} is the flux of A from side i to side j, a_{Ai} is the activity of A on side i, $\Delta\Psi$ is the electrical potential of side 2 relative to side 1, z is the charge carried by A, and $\alpha \equiv F/RT$ (F and R being the Faraday and universal gas constants, respectively). This form of the Ussing-Teorell equation is valid when there is no net flux of water through the membrane. The complete form (Ussing, 1952) constitutes a general steady state relationship for free diffusion between well-mixed solutions separated by a membrane that is uniform in directions perpendicular to the direction of flow. Since unidirectional fluxes through biological membranes can readily be measured, this relationship has proven useful in testing transport models (Stein, 1989; Ussing, 1952).

3. Flux Constraints for Protein-Mediated Transport

The Ussing-Teorell equation is derived from a classical continuum transport equation that describes the flux of solute as being proportional to the gradient of its chemical potential (Ussing, 1952). It is assumed that the physical parameters inside the membrane (e.g., electrical potential, activity coefficient, frictional coefficients) depend only upon the distance from a reference plane parallel to the plane of the membrane. This approach is completely valid where diffusion through a membrane phase is concerned, but it does not apply to the protein-mediated transport that

concerns us here. Binding of the solute molecule to the protein is likely to occur in the process of transport. Motion of the molecule will then be highly constrained, bearing little resemblance to free diffusion. We will instead approach the problem from a kinetic viewpoint. Several factors influence the statistical likelihood that a solute molecule will pass from side 1 to side 2. Some of these factors are directly associated with the nature of the membrane-bound protein that catalyzes transport (henceforth referred to as a *transportase*); others are not. Since our discussion should be generally applicable to different transport systems and since the physical details of these systems are largely unknown, it will be useful to lump all factors pertaining to the transportase (including its concentration in the membrane) into a single factor, Φ, which will be referred to as the rate function. The flux of A from side 1 to side 2 will then be given by the product of Φ_{12} and the protein-independent factors. The most obvious protein-independent factor is the solute activity. If all other factors are fixed, the flux J_{12} will be proportional to the activity of A on side 1.

In the case of a charged solute, an additional protein-independent factor is required. The reason is that transport of a charged solute from a solution at one electrical potential to a solution at another is an energetic process. Since we are initially considering the case where transport of A is independent of other species (hence, transport is not driven by transport of another species or by energy input from chemical reaction), we must account for the fact that only a fraction of solute molecules will have sufficient kinetic energy to overcome an energy barrier. For an energy barrier of magnitude E, this fraction is given by the Boltzmann factor, $e^{-E/kT}$, where k is the Boltzmann constant and T is the absolute temperature. Even an uncharged solute will experience a complex change in potential energy on passage through the membrane, but because the initial and final potential energies

are equal, the same factor applies to transport in both directions. Furthermore, since the magnitude of the maximum energy barier is determined by the interaction between the solute molecule and the transportase, this factor is included in Φ. The magnitude of the potential difference across the membrane, however, does not depend on properties of the transportase and, consequently, will require an explicit factor.

Combining these factors, we arrive at the following expressions for the flux from side 1 to side 2:

$$J_{12} = \begin{cases} \Phi_{12} a_{A1} e^{-\alpha z \Delta\Psi}, & \text{if } z\Delta\Psi > 0; \\ \Phi_{12} a_{A1}, & \text{otherwise.} \end{cases} \tag{2}$$

For the flux in the other direction we have:

$$J_{21} = \begin{cases} \Phi_{21} a_{A2} e^{+\alpha z \Delta\Psi}, & \text{if } z\Delta\Psi < 0; \\ \Phi_{21} a_{A2}, & \text{otherwise.} \end{cases} \tag{3}$$

The Boltzmann factor corresponds to an energy barrier of magnitude $z\Delta\Psi$ (α replaces 1/kT so that $\Delta\Psi$ can be expressed in the conventional units, volts). Since a downward step in potential energy does not affect the probability of passage, the Boltzmann factor appears only when passage requires an upward step.

These equations, while valid, will be of little use until we can place some constraints on Φ_{12} and Φ_{21}, or on the relationship between the two. Our intention is that these rate functions encompass everything that influences the ability of the transportase to perform its function. The transportase interacts with A on both sides of the membrane, so we can expect the rate functions to exhibit dependence on both solute activities. The average electrical potential gradient along a transportase molecule can be in excess of 10^5 volts per centimeter, so we can expect significant changes in transportase structure to result from changes in $\Delta\Psi$. Thus,

both rate functions depend on all three parameters (i.e., $\Phi_{ij} \equiv \Phi_{ij}(a_{A1}, a_{A2}, \Delta\Psi)$). Due to its complex dependence on the architecture of the transportase, little can be said about the relationship between Φ_{ij} and $\Delta\Psi$. We will therefore consider $\Delta\Psi$ to be fixed and focus on the dependence of Φ_{ij} on a_{A1} and a_{A2}. For some solute/transport systems, one or both of these dependences may be so subtle as to be negligible, but neither can be eliminated without loss of generality.

Though neither rate function has been constrained, constraints on the relationship between them, if possible, would enable us to constrain the relationship between the two unidirectional fluxes. The case of equilibrium provides an important constraint. When electrical potentials are involved, the appropriate expression for the chemical potential of solute i is:

$$\mu_i = \mu_i^* + RT \ln a_i + z_i F \Psi \tag{4}$$

where μ_i^* is the standard-state reference potential. At equilibrium, the chemical potentials μ_{A1} and μ_{A2} are equal. Thus:

$$RT \ln a_{A1} + zF\Psi_1 = RT \ln a_{A2} + zF\Psi_2$$

and, by rearranging:

$$\frac{a_{A1}}{a_{A2}} = e^{+\alpha z \Delta\Psi} \tag{5}$$

Since the two unidirectional fluxes must also be equal at equilibrium, we have (from equations (2) and (3)):

$$\Phi_{12} a_{A1} e^{-\alpha z \Delta\Psi} = \Phi_{21} a_{A2},$$

and thus:

$$\frac{a_{A1}}{a_{A2}} = \frac{\Phi_{21}}{\Phi_{12}} e^{+\alpha z \Delta\Psi} \tag{6}$$

Comparison of equations (5) and (6) leads immediately to the conclusion that the values of the two rate functions must be equal at equilibrium.

This constraint is illustrated graphically in Figure 1 for a single value of $\Delta\Psi$. If the Φ axis is imagined to be perpendicular to the plane of the figure, Φ_{12} and Φ_{21} would be represented as curved surfaces displaced outward from the figure (lying in the plane of the figure wherever their values are zero). The equilibrium constraint requires that the two surfaces intersect eachother along a line whose projection in the $a_{A1}a_{A2}$ plane is shown. Another constraint becomes evident when we consider the limiting case of infinitesimal solute concentrations. Since the interaction between a transportase molecule and a solute molecule is transient, by considering sufficiently dilute solute (on both sides), one can always find a region in the $a_{A1}a_{A2}$ plane where the dependence of Φ_{12} and Φ_{21} on a_{A1} and a_{A2} vanishes. That is:

$$\lim_{\substack{a_{A1}\to 0 \\ a_{A2}\to 0}} \left(\frac{\partial \Phi_{ik}}{\partial a_{Aj}}\right) = 0 \qquad (i,j,k \in \{1,2\}; \quad i \neq k).$$

As depicted in Figure 2, the vanishing derivative means that the Φ surfaces will be flat and parallel to the $a_{A1}a_{A2}$ plane. Since the line defined by the equilibrium constraint must pass through this flat region, the rate functions must be equal and constant throughout it. It is possible, perhaps even likely, that for some transportases, normal physiological solute activities map into this region of constant Φ.

The two constraints identified thus far are completely general. Additional constraints may apply to specific solute/transportase systems, but it would be difficult or impossible to infer these from the structure and mechanism of a particular system, given our current state of knowledge. Simple conceptual models of transportases typically imply the global constraint $\Phi_{12} \equiv \Phi_{21}$ (see Stein, 1986). These models are of significant value in the interpretation of kinetic data from real systems, but

they cannot be used to deduce the unidirectional flux relationship for real systems. While the final constraint to be discussed is not universal, the condition for its applicability does not place explicit restrictions on the mode of transport. That, in combination with the significance of its implications for Φ_{12} and Φ_{21}, makes it worthy of consideration here.

As illustrated in Figure 3, the distance from a plane that parallels the membrane can be used as the reaction coordinate in an analysis that draws upon transition state theory. If we fix a_{A1}, a_{A2}, and $\Delta\Psi$, we have fixed the average state of a transportase. Individual transportase molecules continue to experience rapid fluctuations, but collectively, they give rise to well-defined averages. Solute concentration (molecules per transportase per unit length on reaction coordinate), potential energy, statistical entropy, and chemical potential all have well-defined (though not readily measurable) profiles along the reaction coordinate. These profiles characterize every aspect of the membrane that is relevent to solute transport, and consequently, they determine the values of Φ_{12} and Φ_{21}.

Suppose, now, that solute rapidly equilibrates within the transportase. This means that conditions inside the transportase at one end adjust very rapidly to a change in conditions at the other end. The dimensions of typical membrane proteins are on the order of 10 nm, so it is easy to immagine this being the case. The chemical potential profile is then essentially flat across the length of the transportase. If solute is in equilibrium in the bulk solutions separated by the membrane, μ_A has the same value throughout the system (Figure 4). By decreasing either a_{A1} or a_{A2} we will displace the system from equilibrium, but by increasing the activity on the opposite side by the appropriate amount, we can restore the chemical potential within the transportase to its initial value. All profiles within the transportase will then be virtually identical to the initial profiles. Thus, Φ_{12} and Φ_{21} will be unchanged, and

because the initial equilibrium required that they be equal, they must remain equal after the displacement. The locus of all such displacements starting from a single equilibrium state defines a line of constant Φ.

The rate function constraints in the case of rapid equilibrium are illustrated in Figure 5. Here the equality $\Phi_{12} = \Phi_{21}$ holds for all values of a_{A1} and a_{A2}, so the single variable, Φ, can be used to represent the rate function. Note that while the same rate function applies to fluxes in both directions (equations 2 and 3), the membrane need not be symmetric. If the lines of constant Φ are not symmetric with respect to the line defined by $a_{A1} = a_{A2}$, membranes with opposite orientation will yield different fluxes under identical solution conditions.

One point about the concentration profile should be addressed before proceeding. We have treated the activities a_{A1} and a_{A2} as if they apply to the entire solution volume on their respective sides of the membrane. Actually, if net transport is occurring, there is generally a region in the immediate vicinity of the membrane (the Nernst diffusion layer; see Lakshminarayanaiah, 1969) where a significant concentration gradient exists. In fact, if the rapid equilibrium condition holds, gradients in μ_A exist *only* in the Nernst diffusion layers (as in Figure 4). Diffusion through these layers is the rate-limiting step in the transport process. Mixing reduces the thickness of this transition region but does not eliminate it. At equilibrium this is not an issue because gradients do not exist in the aqueous phases. When significant departures from equilibrium are considered, it should be remembered that a_{A1} and a_{A2} refer to solution activities in the immediate vicinity of the transportase.

4. Discussion

We have seen that when an independently transported solute is in equilibrium across a membrane, in rapid equilibrium within a transportase, or present at sufficiently low concentrations, the rate functions for passage in the two directions are equal. Under such conditions, we conclude from equations (2) and (3) that:

$$\frac{J_{12}}{J_{21}} = \frac{a_{A1}}{a_{A2}} e^{-\alpha z \Delta \Psi}, \tag{7}$$

which is identical to equation (1), the flux ratio equation. Thus, the equation derived for transport by simple diffusion through a uniform membrane applies to protein-catalyzed transport under the conditions discussed.

The net rate of transport is also of interest. Under the same conditions, equations (2) and (3) lead to:

$$J_{12} - J_{21} = \begin{cases} \Phi(a_{A1} e^{-\alpha z \Delta \Psi} - a_{A2}), & \text{if } z\Delta\Psi > 0; \\ \Phi(a_{A1} - a_{A2} e^{+\alpha z \Delta \Psi}), & \text{otherwise.} \end{cases} \tag{8}$$

Thus, the *net* flux can be expressed as the product of a parameter that depends on the physical state of the transportase, Φ, and an expression that depends only on measurable properties of the two aqueous compartments. While the dependence of Φ on these measurables will vary from one transportase to the next, the expressions in parentheses are the same for *any* transportase catalyzing transport of A independent of other species. These expressions are thus intrinsic to the transport reaction. Equations (8) imply the following proportionality:

$$J_{12} - J_{21} \propto a_{A1} e^{\alpha z \Psi_1} - a_{A2} e^{\alpha z \Psi_2} \tag{9}$$

which enables us to consolidate the expressions. The right-hand side of this relationship can be termed the "driving force" for transport of A from side 1 to side 2.

If Φ remains constant, the net flux depends on the electrical potentials and activities according this driving force expression.

A general expression for the driving force of processes involving multiple species would be of considerable value, because it would enable us to compare alternative transport models under specified conditions. The extension of our single-species analysis is actually straightforward if we note that relation (9) can be rewritten as:

$$\Omega = \exp\left(\frac{\mu_{A1}}{RT}\right) - \exp\left(\frac{\mu_{A2}}{RT}\right) \tag{10}$$

where Ω is used to represent the driving force function, an intrinsic property of the transport reaction. We can view a transport process in the way we would a chemical process; reactants are consumed to form products. If a transport reaction is written in such a way that the stoichiometric coefficients, ν_i (defined as positive on the right-hand side, negative on the left), reflect an actual molecular process, equation (10) can be generalized as follows:

$$\Omega = \exp\left(\sum_i (|\nu_i| - \nu_i)\frac{\mu_i}{2RT}\right) - \exp\left(\sum_i (|\nu_i| + \nu_i)\frac{\mu_i}{2RT}\right) \tag{11}$$

Terms involving products cancel out in the first summation, and terms involving reactants cancel out in the second. Equation (11) gives the intrinsic driving force for any transport process that does not involve chemical reactions (such as ATP hydrolysis). The net rate of transport is directly proportional to Ω only if Φ is constant. However, Φ is apt to depend much less strongly on solute activities than Ω in situations where it is not truely constant. Thus, Ω can give useful information regarding net transport rates.

References

Lakshminarayanaiah, N. (1969). *Transport Phenomena in Membranes.* New York: Academic Press.

Nobel, P. S. (1974). *Introduction to Biophysical Plant Physiology.* San Francisco: Freeman.

Stein, W. D. (1986). *Transport and Diffusion Across Cell Membranes.* Orlando: Academic Press.

Stein, W. D. (1989). Kinetics of transport: Analyzing, testing, and characterizing models using kinetic approaches. *Meth. Enzymol.* **171**, 23.

Teorell, T. (1949). Membrane electrophoresis in relation to bio-electrical polarization effects. *Arch. Sci. Physiol.* **3**, 205.

Ussing, H. H. (1949). The distinction by means of tracers between active transport and diffusion. *Acta Physiol. Scand.* **19**, 43.

Ussing, H. H. (1952). Some aspects of the application of tracers in permeability studies. *Advan. Enzymol.* **13**, 21.

Figure Captions

Figure 1.

The equilibrium constraint. The Φ axis extends perpendicularly from the plane of the figure. The dashed line is the locus of equilibrium points in the a_{A1}, a_{A2} plane.

Figure 2.

The low-concentration constraint. $\Phi_{12} = \Phi_{21}$ =constant in the shaded region.

Figure 3.

Profiles of potential energy, E, chemical potential, μ, and concentration, $[A]$, across the transportase. The profiles shown correspond to equilibrium.

Figure 4.

Profiles of μ_A under the assumption of rapid equilibrium within the transportase. An infinite number of such profiles exist for a given value of μ_A within the transportase.

Figure 5.

The rapid equilibrium constraint. $\Phi_{12} = \Phi_{21}$ throughout the a_{A1}, a_{A2} plane.

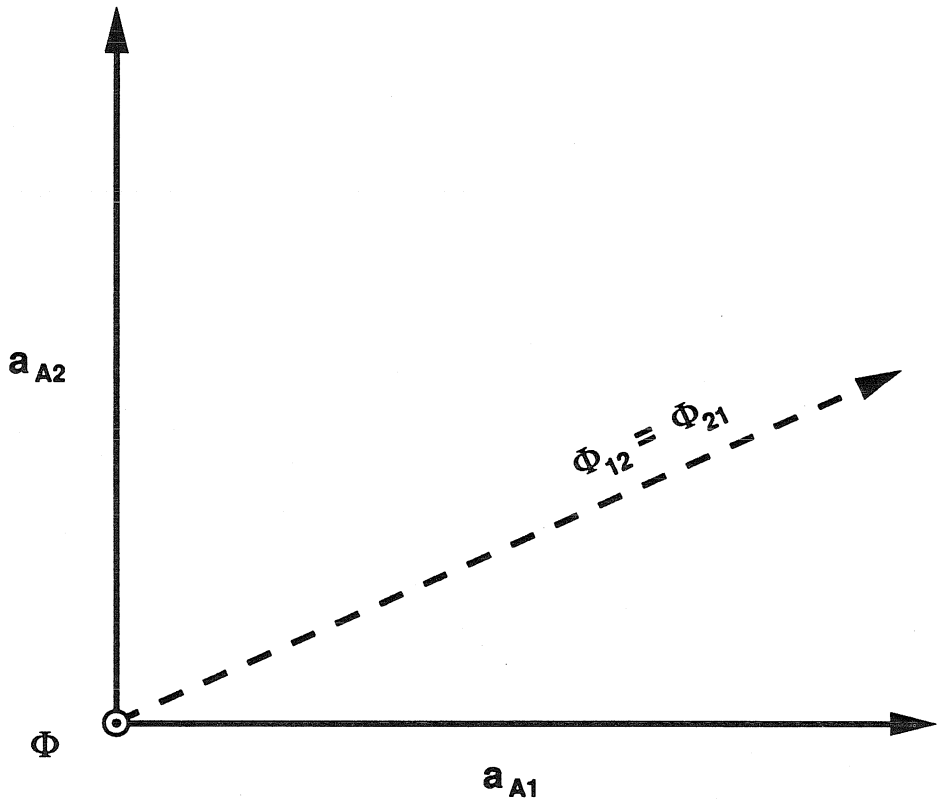

Figure 1.

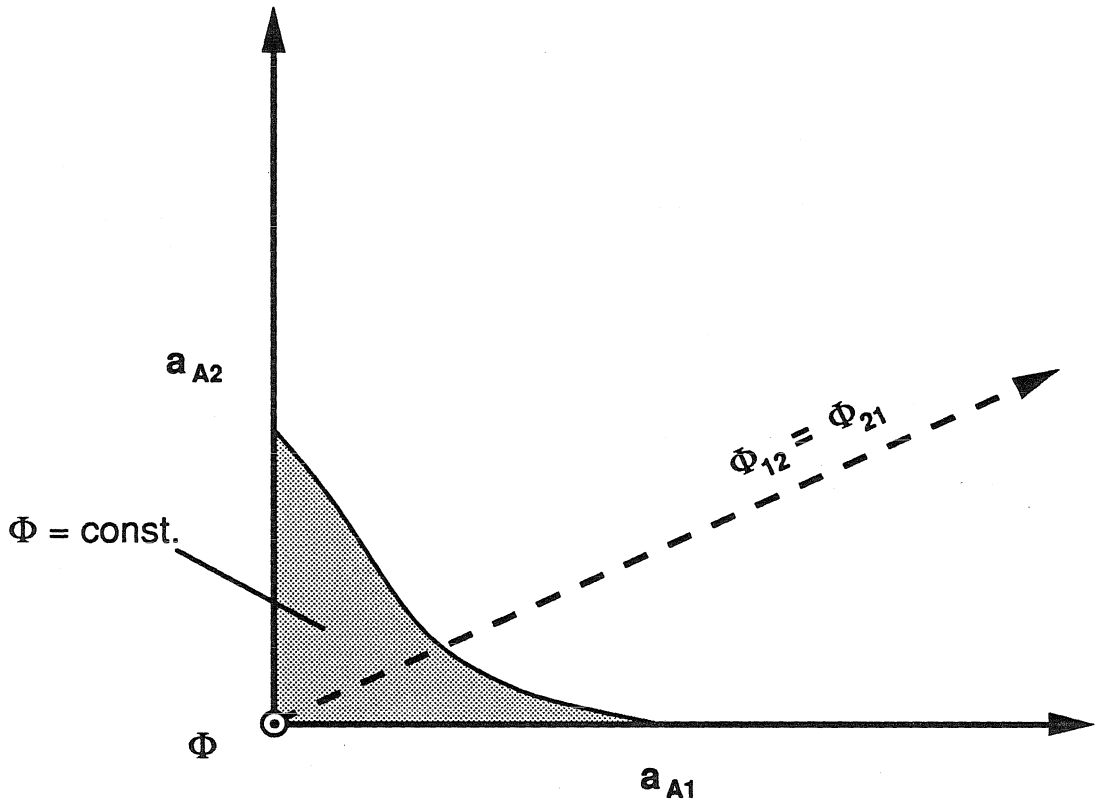

Figure 2.

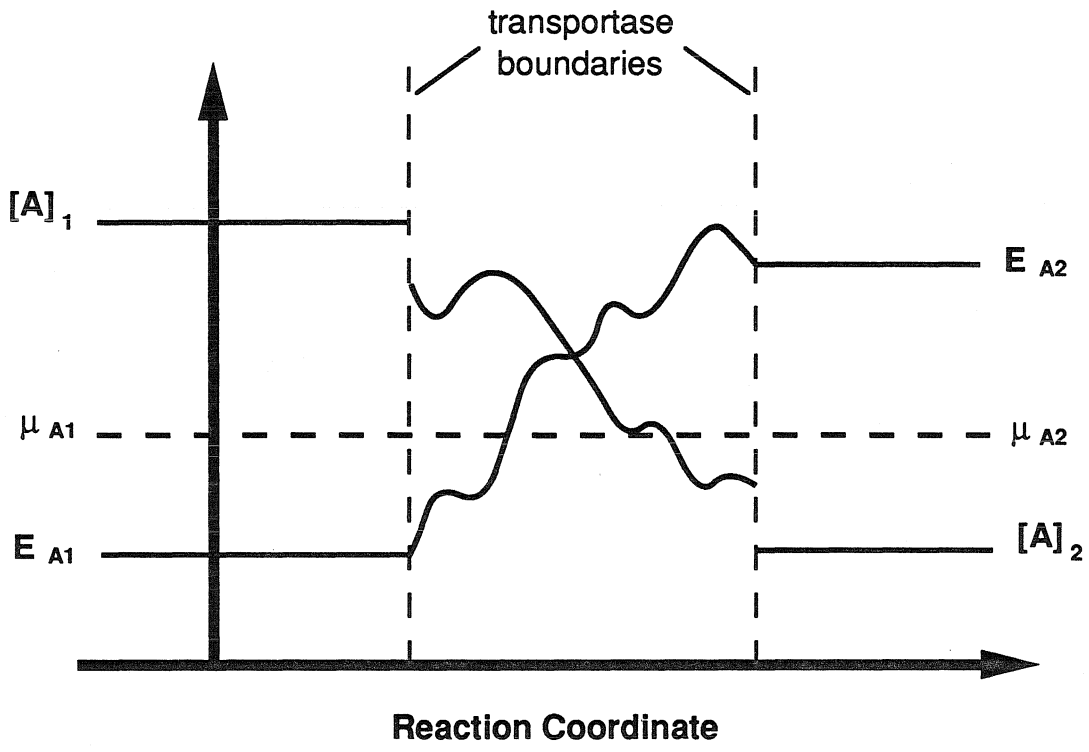
transportase
boundaries
[A]1
E A2
μ A1
μ A2
E A1
[A]2
Reaction Coordinate

Figure 3.

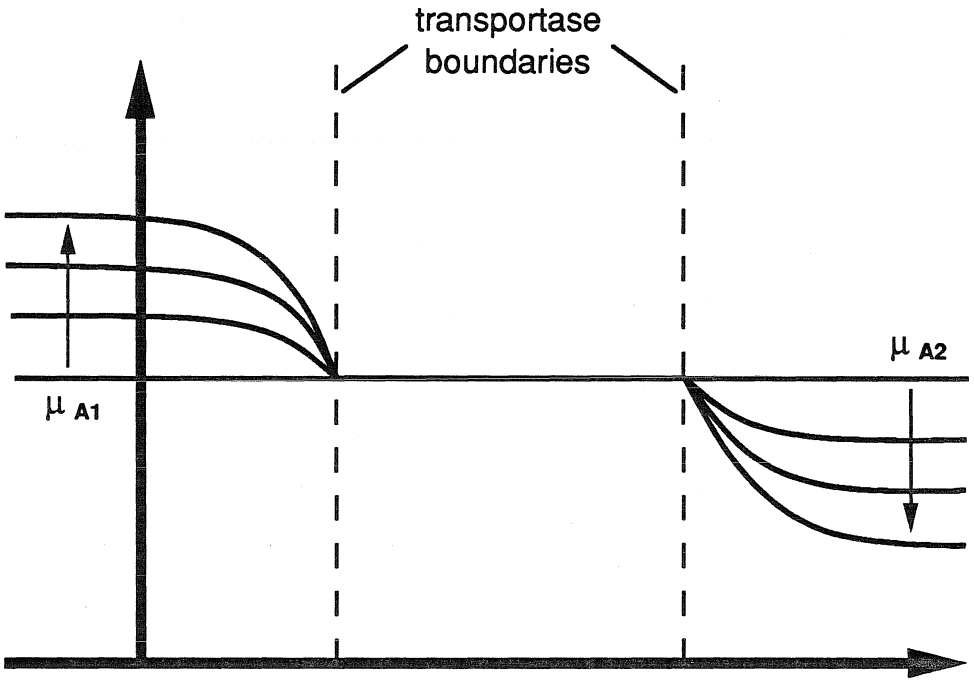

Figure 4.

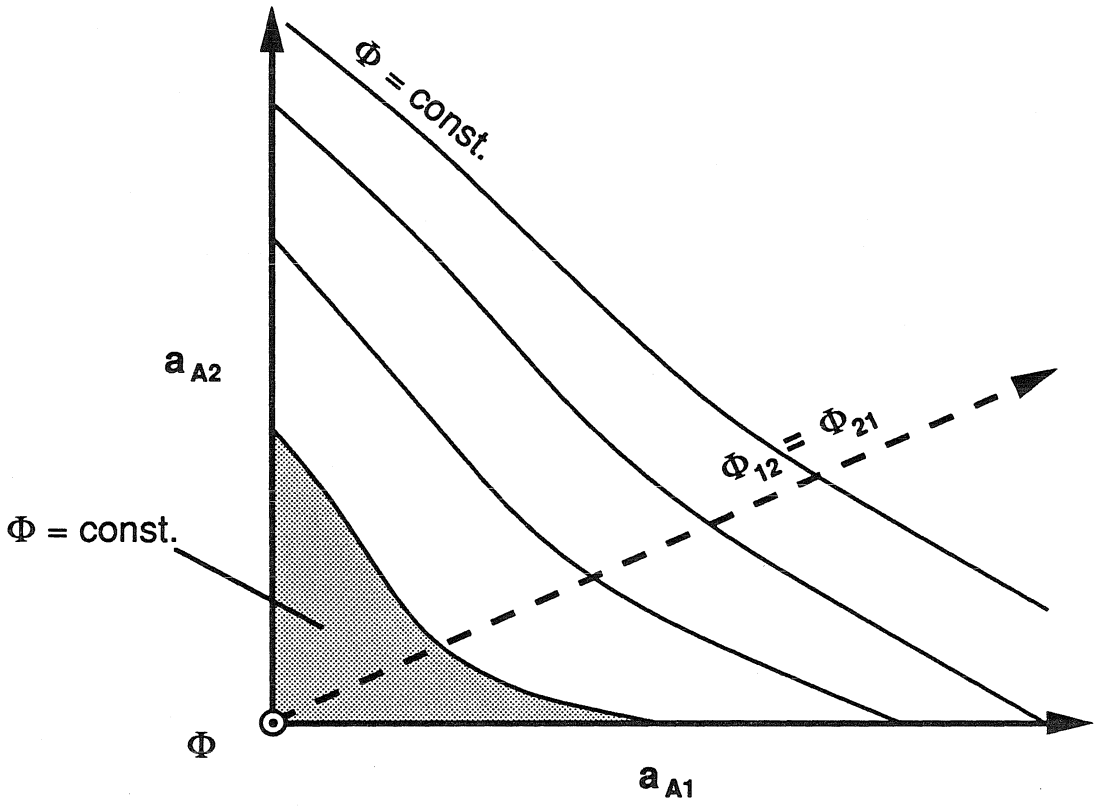

Figure 5.

CHAPTER VI.

TRANSPORT OF LACTATE AND ACETATE THROUGH THE ENERGIZED CYTOPLASMIC MEMBRANE OF *Escherichia coli*

Summary

Lactic and acetic acids are produced by *Escherichia coli* during both anaerobic and aerobic glycolysis. While it is generally accepted that acetate exits the cell by simple diffusion through the cytoplasmic membrane, the forms in which it diffuses have not been identified. Evidence for a specific lactate transport system in *E. coli* has been presented, but the stoichiometry of transport has not been identified. A model proposed in 1979 has been the topic of many papers since that time and appears to be the only published hypothesis regarding the mechanism of lactate efflux in bacteria. In this work, the model is shown to be seriously flawed on theoretical grounds.

Experiments performed on anaerobically glycolyzing *E. coli* to determine the identity of the species involved in lactate and acetate transport are described. The results indicate that acetate is a classical uncoupling agent, permeating the membrane at comparable rates in the dissociated and undissociated forms. Clear evidence was found for the activity of a lactate transport system. It appears as though this system responds to some indicator of glycolytic activity, possibly cAMP or ATP. The transport system appears to catalyze the passage of two lactate molecules with one proton, though the passage of lactate alone could not be definitively ruled out. The data further indicate that lactate permeates the membrane in the dissociated and undissociated forms, though perhaps not as readily as acetate.

1. Introduction

During anaerobic growth on glucose, *Escherichia coli* produces and excretes acetic, lactic, formic, and succinic acids. Appreciable amounts of acetic acid and lactic acid are also produced under aerobic conditions when glucose is present above growth-limiting levels (Tempest and Neijssel, 1987; Doelle *et al.*, 1983; Reiling *et al.*, 1985). The mechanisms of acidic endproduct efflux have remained somewhat elusive. Acetate is known to have a stong inhibitory effect on growth, and this effect has been linked to collapse of the transmembrane pH gradient (Baronofsky, J. J. *et al.*, 1984). Presumably, then, acetic acid is one of a group of weak acids, called uncoupling agents, that inhibits generation of a protonmotive force, Δp. The generally accepted mechanism of uncoupler action (Figure 1) involves permeation of the cytoplasmic membrane in both the dissociated and undissociated forms (Cobley and Cox, 1983). Experiments with *E. coli* membrane vesicles have provided evidence that specific transport systems for lactate and succinate exist in *Escherichia coli* (Matin and Konings, 1973). No such system has been implicated in the transport of acetate.

A model for lactate transport proposed by Michels and coworkers (1979) suggests that a lactate transport system with variable stoichiometry can efficiently transform the free energy from lactate efflux into a protonmotive force (see also: ten Brink and Konings, 1980, 1982, 1986; and Konings *et al.*, 1984). Their lactate model is based upon the general model of carrier-mediated proton-substrate cotransport proposed by Rottenberg (1976). A schematic of the Rottenberg carrier model for transport of a singly charged anion is given in Figure 2. The novel feature of this model is that it allows the transport stoichiometry to vary according to the external pH (dependence on external pH alone was *posulated* rather than inferred).

The carrier protein (an integral part of the membrane) is envisaged as having an acidic proton that dissociates with a pK_a in the physiological range. At low pH, the carrier is uncharged, and the mode of transport is represented by the stoichiometry H^+X^-. That is, the transport of a single X^- molecule is tightly coupled to the transport of a single proton. At higher pH, the carrier proton dissociates, and the transport stoichiometry becomes $H_2^+X^-$.

The rationale behind the two-mode model is that it makes efficient use of free energy derived from the protonmotive force (Rottenberg, 1976): at low external pH, when Δp is large, uptake of a metabolite with a net charge of -1 requires transport of a single proton down its potential gradient, whereas at neutral external pH, when Δp is smaller, two protons are transported per metabolite molecule. However, an important point was overlooked in the formulation of this model. Over a significant pH range, the proportions of carrier molecules in the two charge states are comparable. Thus, *both* transport modes are in operation, and *both* must be taken into account for the model evaluation to be meaningful.

The rate of an ATP-independent transport reaction can be viewed as the product of a kinetic parameter that pertains to the state of the membrane-bound transport system (e.g., carrier or channel) and a "driving force" expression that describes the intrinsic dependence of the transport rate on the concentrations of the transported species (see Chapter V). Under conditions where the kinetic parameter is roughly constant, the net flux through the membrane is directly proportional to the driving force. In Chapter V it was shown that the driving force, Ω, is given by:

$$\Omega = \exp\left(\sum_j (|\nu_j| - \nu_j)\frac{\mu_j}{2RT}\right) - \exp\left(\sum_j (|\nu_j| + \nu_j)\frac{\mu_j}{2RT}\right) \qquad (1)$$

where ν_j is the stoichiometric coefficient of the j^{th} species (negative for species appearing on the left side of the transport equation), and μ_j is the chemical potential

of the j^{th} species. When the reaction consists of transport across an energized membrane, the appropriate expression for μ_j is:

$$\mu_j = \mu_j^* + RT \ln a_j + z_j F \Psi \tag{2}$$

where a_j, z_j, and Ψ are, respectively, the activity of, the charge on, and the electrical potential experienced by species j, and F is the Faraday constant.

Considering again the uptake of a metabolite with unit negative charge, X^-, the two transport reactions in the Rottenberg model are:

$$X_o^- + H_o^+ \rightleftharpoons X_i^- + H_i^+, \quad \text{and}$$

$$X_o^- + 2H_o^+ \rightleftharpoons X_i^- + 2H_i^+$$

where i and o denote values pertaining to the inside and outside of the cytoplasm. We will adopt the convention that inside values are referenced to outside values. Thus: $\Psi_o \equiv 0$, and $\Delta n \equiv n_i - n_o$ for any parameter n (consequently, $\Delta\Psi = \Psi_i$). Then, by writing expressions for the chemical potential for each species we obtain the following for the inward driving force of the first reaction:

$$\Omega_1 = [H^+]_o[X^-]_o - [H^+]_i[X^-]_i \tag{3}$$

The inward driving force for the second transport reaction:

$$\Omega_2 = [H^+]_o^2[X^-]_o - [H^+]_i^2[X^-]_i e^{\alpha\Psi_i} \tag{4}$$

($\alpha \equiv F/RT$) involves Ψ_i because there is net transport of charge. Note that concentrations (denoted by brackets) have been substituted for activities. This approximation, which will be used henceforth, is valid for a particular solute if the activity coefficients for that solute are nearly equal on the two sides of the membrane (they needn't, however, be close to unity).

Equations (3) and (4) represent the intrinsic concentration dependence of the net transport rates for the one-proton and two-proton transport modes. If the kinetic parameters for these two modes are comparatively weak functions of the concentrations (see Chapter 5 for discussion), the net transport rates are nearly proportional to Ω_1 and Ω_2. Note that neither mode alone represents the Rottenberg model. The net rate of the latter is determined by considering the proportion of carriers operating in each of the two transport modes. Normalized driving forces for four values of pH_o have been computed (Table I) using typical values for Ψ_i, pH_i, $[X^-]_i$, and $[X^-]_o$. We have taken pH_i to be constant, as has been observed to be approximately true in *E. coli* (Booth, 1985). Values listed in the second and third columns are proportional to the net transport rate by the one- and two-proton modes, respectively, but the two constants of proportionality probably differ.

If we assume a pK_a of 6.7 for the acidic carrier proton, then 91% of the carriers operate in the one-proton mode at $pH_o = 5.7$, and the same proportion operate in the two-proton mode at $pH_o = 7.7$. It is apparent from these values and those in Table I that if influx by the one-proton mode dominates at low pH as postulated, *efflux* by the same mode must dominate at high pH. If, on the other hand, influx by the two-proton mode dominates at high pH, it must do so to a much greater extent at low pH. Furthermore, the net transport rate must increase by nearly four orders of magnitude as the external pH drops from 7.7 to 5.7. Regardless of the values of the proportionality constants, the one-proton mode operates in reverse when pH_o exceeds 6.1. The two-proton mode simultaneously operates in the forward direction. Since the combined effect of one forward translocation event by the two-proton mode and one reverse translocation event by the one-proton mode is equivalent to the passage of a proton from outside to inside, an energy-dissipating futile cycle operates when $pH_o \geq 6.1$. The Rottenberg model is actually an example

of the classical uncoupling mechanism illustrated in Figure 1. The only addition being that the ratio of effective permeabilities of the two species is a function of pH. All that is required for energy consumption by futile cycling is that the transport modes operate at nonnegligible rates in opposite directions.

The above mentioned model for lactate efflux (termed the "energy-recycling" model) is an example of the Rottenberg model operating in reverse: efflux of lactate at a high intracellular concentration is supposed to drive the efflux of protons, thus contributing to the protonmotive force. The theoretical analysis of this model (Michels *et al.*, 1979) is seriously flawed, however. The authors represent the "transport stoichiometry" $(H^+ : L^-)$ as:

$$n = 1 + \frac{K_a}{[H^+]_o + K_a}$$

where K_a is the acid dissociation constant of the carrier. They then proceed to analyze the model as if the following (single) transport reaction is occurring:

$$L_i^- + nH_i^+ \rightleftharpoons L_o^- + nH_o^+$$

where n is permitted to take on any real value from 1.0 to 2.0. This eliminates the problem of futile cycling from their analysis, but only at the high cost of invoking a fictitious reaction.

The "energy-recycling" model would seem to be an unlikely candidate for lactate efflux. The objective of this work is to obtain information about the mechanisms of transport of both acetate and lactate from intact *E. coli* cells. Our approach has been to measure intracellular and extracellular concentrations of both, along with ΔpH and $\Delta\Psi$, during and after transient anaerobic glycolysis. Since the ratio of inside and outside concentrations at equilibrium depends very strongly on the stoichiometry of the transport reaction, this method can reliably distinguish different

modes of transport. An additional advantage to this method is that it avoids artifacts that might arise in studies with membrane vesicles.

2. Materials and Methods

Growth and harvest of cells:

Escherichia coli strain K12 was grown aerobically in 3.5 liter batches in a modified M9 medium (24 mM glucose, 1.0 mM $MgSO_4$, 0.1 mM $CaCl_2$, 50 mM sucrose, and (per liter): 13.2 g Na_2HPO_4, 3.0 g KH_2PO_4, 0.5 g NaCl, 1.0 g NH_4Cl and 200 ml filtered, autoclaved tap H_2O) at 37°C in a Chemap fermenter without pH control. Sucrose was added to the medium to increase its osmolarity. Without sucrose, the medium osmolarity is significantly lower than the osmolarity of the buffer into which the cells are resuspended. This could result in loss of cell turgor. When the culture reached $A_{590} = 0.33$, the fermenter contents were rapidly cooled *in situ* to 8°C, and 3 l were dispensed into 500 ml centrifuge bottles (on ice). Cells were centrifuged in a Beckman JA-10 rotor at 4000 rpm for 10 minutes at 4°C, washed in 1 l of chilled buffer (100 mM 1,4-Piperazinediethanesulfonic acid (PIPES), 20 mM NaH_2PO_4, 20 mM NaCl, adjusted to pH 7.30 by addition of KOH) centrifuged as above, resuspended in about 20 ml of the PIPES buffer, and centrifuged in two equal aliquots in graduated conical-bottom polycarbonate tubes in a JA 13.1 (swinging bucket) rotor at 7500 rpm for 10 minutes at 4°C. Use of a swinging bucket rotor and transparant graduated tubes is necessary for accurate determination of the volume of the cell pellets. Pellets were resuspended to a final volume equal to twice the pellet volume in the same PIPES buffer. These suspensions were kept on ice for up to an hour before use.

Anaerobic glycolysis:

A 2.3 ml aliqout of the final cell suspension was warmed to 20°C in a capped tube. The head space in the tube was then flushed with nitrogen, 10.3 mg of D-glucose-1-^{13}C (Sigma Chemical) dissolved in 50 μl H_2O was added, and the tube was immediately re-capped and shaken to thoroughly distribute the glucose. After mixing, the head space was again quickly flushed with nitrogen and capped tightly. The glycolyzing cell suspension was kept at 20°C with occasional mixing. One of several experiments (described below) was performed with the cell suspension.

Measurement of pH_i and pH_o:

Measurement of intracellular and extracellular pH values was accomplished via ^{31}P NMR spectroscopy (Slonczewski *et al.*, 1980; Axe and Bailey, 1987) with a Bruker AM 500 spectrometer. Using the intact cell suspension as the sample, data from 60 70° pulses at 0.50 s intervals were acquired for each spectrum, giving excellent time resolution of inorganic phosphate chemical shifts. Data were acquired for one spectrum prior to glucose addition (described above) and for other spectra at 35 second intervals beginning approximately one minute after glucose addition. A calibration curve was generated by measuring the chemical shift of inorganic phosphate in samples of PIPES buffer (see above) to which various quantities of acetic acid were added.

Measurement of $\Delta\Psi$:

A cell suspension was prepared as described above except that the PIPES buffer contained EDTA added to a final concentration of 20 mM. This is necessary to render the cell wall permeable to the probe used in the measurement

(Kashket, 1982). Approximately 45 minutes prior to glucose addition, 1.0 μl (1.0 μCi) Tetra[^{3}H]phenylphosphonium bromide, TPP^+Br^- (Amersham), was added per ml of suspension. After mixing, the suspension was placed back on ice until use. Once before glucose addition, and at various times afterward, a 500 μl aliquot of the cell suspension was placed in a 1.5 ml microcentrifuge tube and spun at 15,850 g for 1 min in a Beckman Microfuge E centrifuge. The supernatant was removed and placed in a fresh tube immediately after spinning. After all samples were collected, 230 μl from each were added to 800 μl H_2O in separate 6 ml polyethylene scintillation vials. 5 ml of Scintiverse II (Fisher) were added to each vial, and counts for each were recorded with a Beckman LS 5801 liquid scintilation counter. To correct for binding of TPP^+ to cell components (Kashket, 1985), a cell suspension was prepared as described except that *n*-butanol was added prior to the final resuspension (7% v/v), and the resulting suspension was incubated at 34°C (Kashket *et al.*, 1980) for 25 minutes. This treatment disrupts the membrane (Kashket *et al.*, 1980), enabling TPP^+ *binding* to be measured without interference from TPP^+ *uptake.* Samples for scintillation counting were prepared as just described.

Preparation of lysate and supernatant samples:

The procedures described below involve measurements performed on lysate and supernatant samples prepared as described here. Two 700 μl aliquots were drawn from the cell suspension, placed in 1.5 ml microcentrifuge tubes, and centrifuged at 15,850 g for 1 min in a Beckman Microfuge E centrifuge. Approximately 30 seconds after starting centrifugation, another 700 μl aliquot of cell suspension was injected into 1.60 ml of lysis solution (3% (w/v) SDS, 0.20 N NaOH, 40% (v/v) D_2O). This mixture was immediately capped and shaken vigorously. Supernatants were collected from the first two aliqouts immediately after completing centrifugation.

570 μl of the combined supernatant (600 μl in the case of compartment volume determination) were added to a 1.6 ml solution consisting of 500 μl of the above lysis solution, 400 μl of D_2O, and 700 μl of distilled H_2O. This was mixed gently to avoid foaming.

Measurement of cytoplasmic and extracytoplasmic volumes:

Cytoplasmic and extracytoplasmic solution volumes were measured by preparing the usual cell suspension and adding ethyl-1-^{13}C alcohol (1.3 μl per ml suspension) in addition to D-glucose-1-^{13}C (both obtained from Sigma Chemical). After allowing glycolysis to proceed for a few minutes, lysate and supernatant samples were prepared as described above. ^{13}C NMR spectroscopy was used to quantify the amounts of ethanol and glucose in the two samples. For each spectrum, data from 3000 pulses (pulse width = 9.5 μs) at 2.0 s intervals were acquired with inverse-gated proton decoupling. Some distortion of peak areas is expected in these spectra because of saturation and the nuclear Overhauser effect (NOE). Correction factors were obtained with similarly prepared samples that had somewhat higher levels of labeled glucose and ethanol. Two spectra were obtained for each of these samples: one with pulsing at 2.0 s intervals, the other with pulsing at 30.0 s intervals. That latter were considered to be free from saturation/NOE distortion. Corrected glucose and ethanol peak areas were used to compute the extracytoplasmic and total solution volumes (see Results), the difference being equal to the cytoplasmic solution volume.

Measurement of cytoplasmic and extracytoplasmic concentrations:

At three time points after addition of glucose to the cell suspension, lysate and supernatant samples were prepared as described above. Lactate and acetate concentrations were determined by the method used to determine glucose and ethanol concentrations except that a pulse interval of 1.0 s was used. Factors for correcting saturation/NOE distortion were again obtained with similar samples using a 30.0 s pulse interval.

3. Results

The cell suspensions used in this work are dense (roughly 100 mg dry weight per ml), so a significant fraction of the sample volume is not accessible to aqueous solutes. The fraction that *is* accessible is divided by cell inner membranes into numerous cytoplasmic solution volumes and a single extracytoplasmic solution volume. Since we are using measurements that cannot distinguish individual cells, the volumes of relevance are the combined cytoplasmic solution volume and the extracytoplasmic solution volume. All measurements will correspond to mean values for the entire population of cells in the suspension. ^{31}P NMR measurements (see below) provide direct evidence that the distribution of individual cell behavior about the mean is quite narrow, so it is reasonable to use these mean quantities to infer actual cellular values.

Our method for measuring the two volumes relies upon two well-established principles. First, ethanol readily permeates lipid bylayers (Brahm, 1983). Since it is uncharged, we expect it to distribute itself uniformly throughout the solution space. Second, glucose is phosphorylated upon entry into the cytoplasm (Saier,

1977) and the resulting glucose 6-phosphate is rapidly metabolized (Ugurbil *et al.*, 1978). Hence, glucose is expected to be uniformly distributed throughout the extracytoplasmic solution space but nonexistent elsewhere. By comparing the ^{13}C NMR peak areas for glucose and ethanol (see Figure 3) in supernatant and lysate samples, one can determine the volume fractions in the cell suspension that are accessible to either molecule. The volume fraction accessible to ethanol corresponds to the total volume fraction that is occupied by aqueous solution, and the volume fraction accessible to glucose corresponds to the volume fraction that is occupied by extracytoplasmic solution. The *E. coli* suspensions used here were found to have the following volume composition: 62% extracytoplasmic solution space, 20% cytoplasmic solution space, and 18% excluded space.

^{31}P NMR can provide a wealth of information on the state of glycolyzing cells without disturbing their behavior (Ugurbil *et al.*, 1979). This technique is used here primarily to determine cytoplasmic and extracytoplasmic pH values. At physiological pH, inorganic phosphate exists in two forms, $H_2PO_4^-$ and HPO_4^{2-}, that resonate at different frequencies. In aqueous solution, the exchange between these forms is sufficiently rapid that a single resonance is observed, the frequency of which depends on the relative amounts of the two forms. The chemical shift of the phosphate peak can therefore be used to measure pH (Gadian, 1982). Since the cytoplasmic pH typically differs from the extracytoplasmic pH, two inorganic phosphate resonances are observed (Figure 4). Using an appropriate calibration curve to determine pH values from chemical shifts, the cytoplasmic and extracytoplasmic pH trajectories in a glycolyzing cell suspension were found to be as shown in Figure 5. The extracytoplasmic pH, pH_o, decreases steadily throughout the experiment. pH_i, on the other hand, increases slightly after decreasing in the first two minutes of glycolysis. After about five minutes, pH_i also decreases with time. ΔpH increases rapidly in the first

five minutes following glucose addition (Figure 5) and then remains at about 0.55 until twenty minutes, at which point it increases significantly.

Two other important pieces of information can be obtained from these spectra. The β phosphate of ATP resonates at approximately -19 ppm. As shown in Figure 6, ATP is not detectable prior to glucose addition. Within about a minute of glucose addition, ATP is present at detectable levels. It remains at a nearly constant level for about 15 minutes and then quickly falls to barely detectable levels. As mentioned previously, it is also possible to obtain information regarding the heterogeneity of the cell population from ^{31}P NMR spectra. Cytoplasmic pH heterogeneity will broaden the intracellular inorganic phosphate peak. It should be noted, however, that a number of other factors (including molecular mobility and anisotropy of local environment) broaden intracellular resonances relative to the extracellular phosphate resonance (Gadian, 1982). As seen in figure 4, the intracellular inorganic phosphate resonance is noticeably more broad than the extracellular resonance. However, even perfect uniformity in pH_i throughout the population would not result in an intracellular peak as narrow as the extracellular peak. When this is taken into consideration, the population heterogeneity appears to be quite small.

The transmembrane electrical potential, $\Delta\Psi$, was found to be quite stable during the course of the experiment (Figure 7). Prior to glucose addition, a significant protonmotive force exists with contributions from both ΔpH and $\Delta\Psi$ (Figures 5 and 7). After glycolysis begins, $\Delta\Psi$ increases slowly but measurably for roughly 15 minutes. Following a brief drop, it increases at a higher rate than previously for the duration of the experiment. All measurements determined the value of $\Delta\Psi$ to be between -36 and -42 mV, substantially lower in magnitude than reported values (-93 to -122 mV) for *E. coli* growing anaerobically in mid-exponential phase (Kashket, 1983).

Thus far, three independent measurements have indicated that a transition occurs fifteen to twenty minutes after the addition of glucose: $\Delta\Psi$ undergoes a transition at about 15 minutes, ATP drops from a steady level at roughly 17 minutes, and ΔpH starts to increase from a stable value at about 20 minutes. It seems likely that these phenomena result from glucose exhaustion. Measurements of lactate and acetate were performed before and after this transition period so that effects on their transport, if any, might be detected. 10.1 minutes after glucose addition, the cytoplasmic-to-extracytoplasmic concentration ratios were found to be 1.1 ± 0.1 and 2.1 ± 0.1 for lactate and acetate, respectively. 24.7 minutes after glucose addition, the ratios were 2.4 ± 0.1 and 1.9 ± 0.1, and at 29.1 minutes, they were 2.4 ± 0.1 and 2.0 ± 0.1. The significance of these numbers is discussed below.

4. Discussion

We will consider the four transport modes of the following stoichiometries for both lactate and acetate: X^-, $H^+X_2^-$, H^+X^-, and $H_2^+X^-$, where X^- may be lactate or acetate. These will be termed the zero-, half-, one-, and two-proton modes, referring to the number of protons transported per conjugate base molecule. Equations for the *inward* driving force of the latter two have been written (equations (3) and (4)). Outward driving forces are equal in magnitude but opposite in sign. The outward driving force for the zero- and half-proton modes are:

$$\Omega_0 = [X^-]_i e^{-\alpha\Psi_i} - [X^-]_o \qquad \text{and} \tag{5}$$

$$\Omega_{1/2} = [H^+]_i [X^-]_i^2 e^{-\alpha\Psi_i} - [H^+]_o [X^-]_o^2 \tag{6}$$

We will not limit our analysis to a single transport mode for each species. Multiple modes are likely to operate; if acetate is indeed a classical uncoupling agent, it is

transported via two modes simultaneously.

A simple but informative approach is to compare measured concentration ratios ($[X^-]_i/[X^-]_o$) with the equilibrium ratios for the different transport modes. The equilibrium condition for a transport mode is determined by setting the driving force equal to zero. Straightforward algebraic manipulation then yields these expressions for the equilibrium ratios of the zero-, half-, one-, and two-proton modes, respectively:

$$\frac{[X^-]_i}{[X^-]_o} = e^{\alpha\Psi_i}$$

$$\frac{[X^-]_i}{[X^-]_o} = e^{\frac{1}{2}\alpha\Psi_i}\sqrt{\frac{[H^+]_o}{[H^+]_i}}$$

$$\frac{[X^-]_i}{[X^-]_o} = \frac{[H^+]_o}{[H^+]_i}$$

$$\frac{[X^-]_i}{[X^-]_o} = \left(\frac{[H^+]_o}{[H^+]_i}\right)^2 e^{-\alpha\Psi_i}$$

Table II lists the equilibrium ratios for all four modes evaluated with Ψ_i and ΔpH values measured 10.1, 24.7, and 29.1 minutes after glucose addition (the times at which lactate and acetate measurements were made).

Note that the equilibrium ratios for the different transport modes are well separated. If the differences were small, it would be very difficult to differentiate between these modes experimentally. The experimentally determined concentration ratio for acetate is approximately 2 at all three time points. This is sufficiently low compared to the ratios for the two-proton mode that this mode can safely be ruled out with regard to acetate transport. One cannot conclude that *no* acetate is transported by this mode, but a mode with lower equilibrium ratios must operate much more rapidly. It is not such an easy matter to reject a mode with equilibrium ratios that are *lower* than the observed ratios because we cannot assume *a priori*

that the actual transport is in equilibrium. Transport limitation would result in observed ratios that are higher than the equilibrium ratios.

If the transient behavior observed at 15 to 20 minutes after glucose addition is due to glucose exhaustion, one would expect the rates of lactate and acetate production at 24.7 and 29.1 minutes to be much lower than the rates at 10.1 minutes. This was checked by comparing the peak areas in the ^{13}C NMR spectra of the lysate samples prepared at 24.7 and 29.1 minutes. The comparison indicates that the rate of acetate production roughly 25 minutes after glucose addition is about 20% of the average rate in the first ten minutes of glycolysis. The rate of lactate production at 25 minutes is only 11% of the average rate in the first ten minutes. Thus, if there is transport limitation for either acid at 10.1 minutes, it should be greatly reduced at 24.7 minutes. A substantial decrease in the observed concentration ratio from the first time point to the second would then imply transport limitation. This is not observed for either acid (Table II).

In the absence of transport limitation, the observed concentration ratio should be nearly equal to the appropriate equilibrium ratio if a single transport mode is operating. The observed ratios for acetate deviate significantly from all equilibrium ratios, so it appears that acetate is transported by competing modes. The half-proton and two-proton transport modes are untenable unless a carrier or channel is involved because the implied transport species do not exist at appreciable levels in solution. A membrane protein, however, can bind solutes (not necessarily in close proximity to each other) to form pseudo-compounds that do not exist in solution. Acetate transport is generally thought to occur by simple diffusion through the bilayer because transport deficient mutants have not been isolated and because the saturation behavior typical of protein-dependent transport has not been observed (Baronofsky, 1984). Assuming this to be correct, we can eliminate the half-proton

mode from consideration. The two remaining transport modes (zero- and one-proton) are the ones believed to be responsible for uncoupler action (Figure 1). The observed acetate concentration ratios fall between the equilibrium ratios for these two modes, indicating that they operate at comparable rates but in opposite directions. Our results therefore support the idea that acetate acts as a classical uncoupling agent.

The transport of lactate appears to be more complex. A remarkable feature of the concentration ratio data for lactate is that the ratio increases more than twofold between 10.1 and 24.7 minutes after glucose addition. A surge in lactate production could cause the ratio to increase, but this is clearly not the case. As mentioned above, the rate of lactate production was determined to be very low beyond 24.7 minutes. The only other explanation is that a major shift in the mode of lactate transport occurs between 10.1 and 24.7 minutes. Changes in a number of physical parameters might alter the permeability properties of the bilayer, but the effect is apt to be much more subtle than that seen here. Furthermore, the parameter most likely to influence the structure of the bilayer, $\Delta\Psi$, undergoes a very modest (-4%) change during this period. This strongly suggests that a transport system is involved in lactate efflux. As already mentioned, previous work (Matin and Konings, 1973) has produced evidence for the existence of a specific lactate transport system in *E. coli*.

The "energy recycling" model proposes that the number of protons transported per lactate molecule decreases with decreasing pH. The ratios in Table II show the opposite trend. As previously discussed, when two or more passive transport modes compete, futile cycling results. It is difficult to conceive of a mechanism whereby an abrupt and synchronized transition from one mode to another could occur with a large number of transporters distributed over the membrane, yet this is the only

way that multiple modes could operate without competing.

A more plausible explanation is that the shift in tranport modes is from protein-mediated efflux to simple diffusion across the bilayer. Simple diffusion is presumably occurring throughout the experiment, but protein-mediated transport ceases sometime between 10.1 and 24.7 minutes. This suggests that the transport system is responding to some cellular indicator of glycolytic activity, perhaps cyclic AMP or ATP. If it has ATPase activity, none of the modes we are considering apply; the free energy of ATP hydrolysis would have to be included in the Ω equations. Efflux of L^- is energetically favorable throughout the experiment, however. Unless proton export is a major function of the lactate transporter, it is unlikely that it would hydrolyze ATP. Assuming lactate transport to be a passive process, protein-mediated transport in the early part of the experiment would either be via the zero-proton mode (in which case simple diffusion is dominant) or via the half-proton mode (in which case protein-mediated transport is dominant). Simple diffusion, then, seems to be occurring by the same modes identified for acetate. The lactate transport system does not eliminate futile cycling, and hence, lactate still acts as an energy dissipating uncoupler. The role of the transport system is presumably to improve the net rate of efflux. If that is the case, one would expect protein-mediated transport to dominate when it is in operation. The data presented here would then indicate that the lactate transport system operates via the half-proton mode.

References

Axe, D. D. and Bailey, J. E. (1987). Application of ^{31}P nuclear magnetic resonance spectroscopy to investigate plasmid effects on *Escherichia coli* metabolism. *Biotech. Lett.* **9**, 83.

Baronofsky, J. J., Schreurs, W. J. A., and Kashket, E. R. (1984). Uncoupling by acetic acid limits growth of and acetogenesis by *Clostridium thermoaceticum*. *Appl. Env. Microbiol.* **48**, 1134.

Booth, I. R. (1985). Regulation of cytoplasmic pH in bacteria. *Microbiol. Rev.* **49**, 359.

Brahm, J. (1983). Permeability of human red cells to a homologous series of aliphatic alcohols. *J. Gen. Physiol.* **81**, 283.

Cobley, J. G. and Cox, J. C. (1983). Energy conservation in acidophilic bacteria. *Microbiol. Rev.* 47, 579.

Doelle, H. W., Ewings, K. N., and Hollywood, N. W. (1982). Regulation of glucose metabolism in bacterial systems. *Adv. Bioch. Eng.* **23**, 1.

Gadian, D. G. (1982). *Nuclear magnetic resonance and its applications to living systems.* New York: Oxford University Press.

Kashket, E. R., Blanchard, A. G. and Metzger, W. C. (1980). Proton motive force during growth of *Streptococcus lactis* cells. *J. Bacteriol.* **143**, 128.

Kashket, E. R. (1982). Stoichiometry of the H^+-ATPase of growing and resting, Aerobic *Escherichia coli*. *Biochem.* **21**, 5534.

Kashket, E. R. (1983). Stoichiometry of the H^+-ATPase of *Escherichia coli* cells during anaerobic growth. *FEBS Lett.* **154**, 343.

Kashket, E. R. (1985). The proton motive force in bacteria: A critical assessment of methods. *Ann. Rev. Microbiol.* **39**, 219.

Konings, W. N., Hellingwerf, K. J., and Elferink, M. G. L. (1984). The interaction between electron transfer, proton motive force and solute transport in bacteria. *Antonie van Leeuwenhoek* **50**, 545.

Matin, A. and Konings, W. N. (1973). Transport of lactate and succinate by membrane vesicles of *Escherichia coli, Bacillus subtilis* and a *Pseudomonas* species. *Eur. J. Biochem.* **34**, 58.

Michels, P. A. M., Michels, J. P. J., Boonstra, J., and Konings, W. N. (1979). Generation of an electrochemical proton gradient in bacteria by the excretion of metabolic end products. *FEMS Microbiol. Lett.* **5**, 357.

Reiling, H. E., Laurila, H., and Fliechter, A. (1985). Mass culture of *Escherichia coli:* Medium development for low and high density cultivation of *Escherichia coli* B/r in minimal and complex media. *J. Biotechnol.* **2**, 191.

Rodriguez, R. L. and Tait, R. C. (1983). *Recombinant DNA Techniques: An Introduction.* Reading, MA: Addison-Wesley.

Rottenberg, H. (1976). The driving force for proton(s) metabolites cotransport in bacterial cells. *FEBS Lett.* **66**, 159.

Saier, M. H. (1977). Bacterial phosphoenolpyruvate:sugar phosphotransferase systems: Structural, functional, and evolutionary interrelationships. *Bacteriol. Rev.* **41**, 856.

Slonczewski, J. L., Rosen, B. P., Alger, J. R., and Macnab, R. M. (1980). pH homeostasis in *Escherichia coli:* measurements by nuclear magnetic resonance of methylphosphonate and phosphate. *Proc. Nat. Acad. Sci. USA* **78**, 6271.

Tempest, D. W. and Neijssel, O. M. (1987). Growth yield and energy distribution. In: *Escherichia coli and Salmonella typhimurium.* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanic, B., Schaechter, M., & Umbarger, H. E., eds.) pp. 797-806.

ten Brink, B. and Konings, W. N. (1980). Generation of an electrochemical proton gradient by lactate efflux in *Escherichia coli* membrane vesicles. *Eur. J. Biochem.* **111**, 59.

ten Brink, B. and Konings, W. N., (1986). Generation of a protonmotive force in anaerobic bacteria by end-product efflux. *Meth. Enzymol.* **125**, 492.

Ugurbil, K., Rottenberg, H., Glynn, P., and Shulman, R. G. (1978). ^{31}P nuclear magnetic resonance studies of bioenergetics and glycolysis in anaerobic *Escherichia coli* cells. *Proc. Nat. Acad. Sci. USA* **75**, 2244.

Ugurbil, K., Shulman, R. G. and Brown, T. R. (1979). High resolution ^{31}P and ^{13}C nuclear magnetic resonance studies of *Escherichia coli* cells *in vivo*. In: *Biological Applications of Magnetic Resonance*. New York: Academic Press.

TABLE I

Driving forces (Ω) for both transport modes in the Rottenberg model *

pH_o	one-proton mode	two-proton mode
5.7	1	1
6.4	-0.33	4.0×10^{-2}
7.0	-0.58	2.4×10^{-3}
7.7	-0.65	2.0×10^{-5}

* The transmembrane eletrical potential is taken to be 120 mV (inside negative), pH_i is taken to be 7.4, and the intracellular concentration of metabolite (X^-) is taken to be twentyfold higher than the extracellular concentration. Driving forces are normalized to the values at $pH_o = 5.7$.

TABLE II

Theoretical and experimental concentration ratios

mode	10.1 minutes	24.7 minutes	29.1 minutes
zero-proton:	0.22	0.21	0.20
half-proton:	0.87	0.93	0.93
one-proton:	3.4	4.2	4.4
two-proton:	54.	87.	97.
acetate:	2.1	2.1	2.2
lactate:	1.1	2.4	2.5

Figure Captions

Figure 1.

The generally accepted mechanism of uncoupling by weak acids. HX represents the undissociated acid.

Figure 2.

The carrier-mediated proton-substrate cotransport model (Rottenberg, 1976). C represents the carrier, which is a membrane-bound protein that catalyzes transport.

Figure 3.

Representative ^{13}C NMR spectra (top: supernatant sample; bottom: lysate sample). βG and αG are the resonances of the glucose-C1. E1 indicates the position of the resonance of ethanol-C1 (not present in these spectra). L and A indicate the resonances of lactate and acetate, respectively. Acquisition parameters are given in Materials and Methods.

Figure 4.

Representative ^{31}P NMR spectrum. SP indicates the sugar phosphate resonances. P_{in} and P_{out} indicate the intracellular and extracellular inorganic phosphate resonances, and A indicates the resonance of the β phosphate of ATP. The acquisition time for this spectrum was 30 s. Acquisition parameters are given in Materials and Methods.

Figure 5.

pH values versus time, as determined by ^{31}P NMR spectroscopy. The upper dotted curve is the intracellular pH trajectory; the lower dotted curve is the

extracellular pH trajectory (left scale). The solid curve is the ΔpH trajectory (right scale).

Figure 6.

ATP β phosphate resonance at various times relative to glucose addition. From the top: pre-glucose, 1.4 min, 6.1 min, 10.2 min, 13.7 min, 17.9 min, and 20.3 min after glucose addition.

Figure 7.

The transmembrane electrical potential, $\Delta\Psi$, versus time after glucose addition. See Materials and Methods for a description of the procedure used.

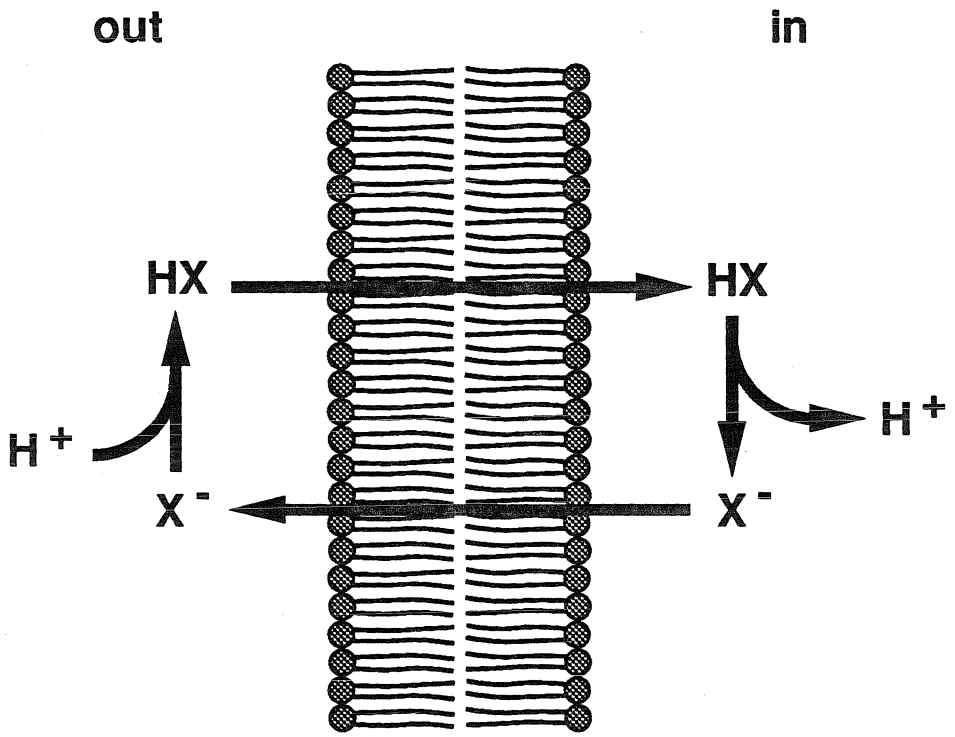

Figure 1.

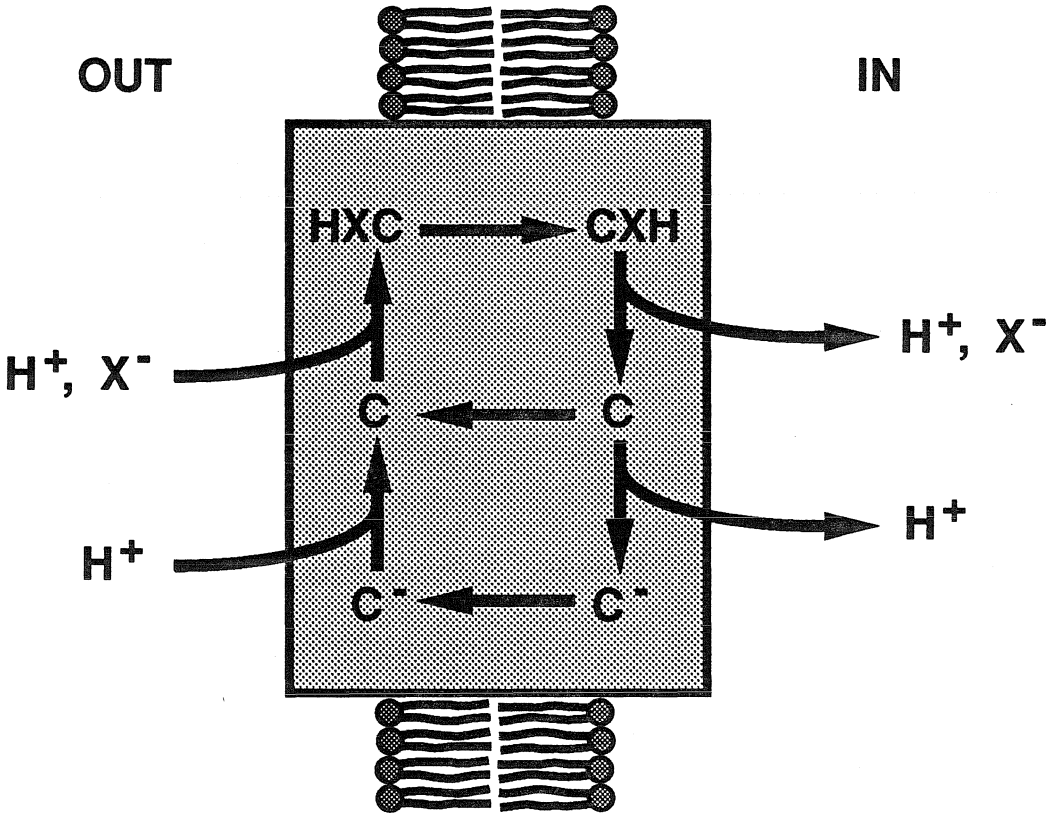

Figure 2.

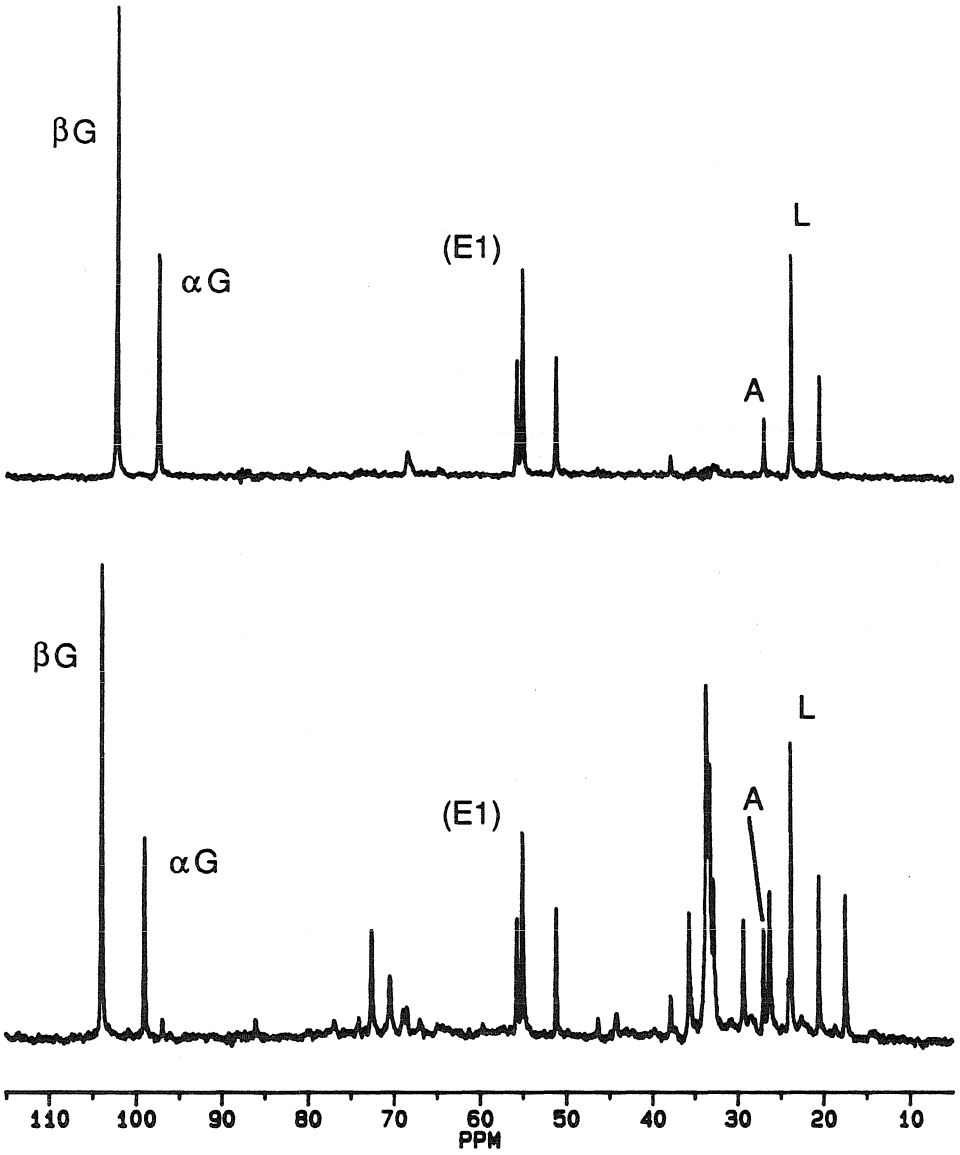

Figure 3.

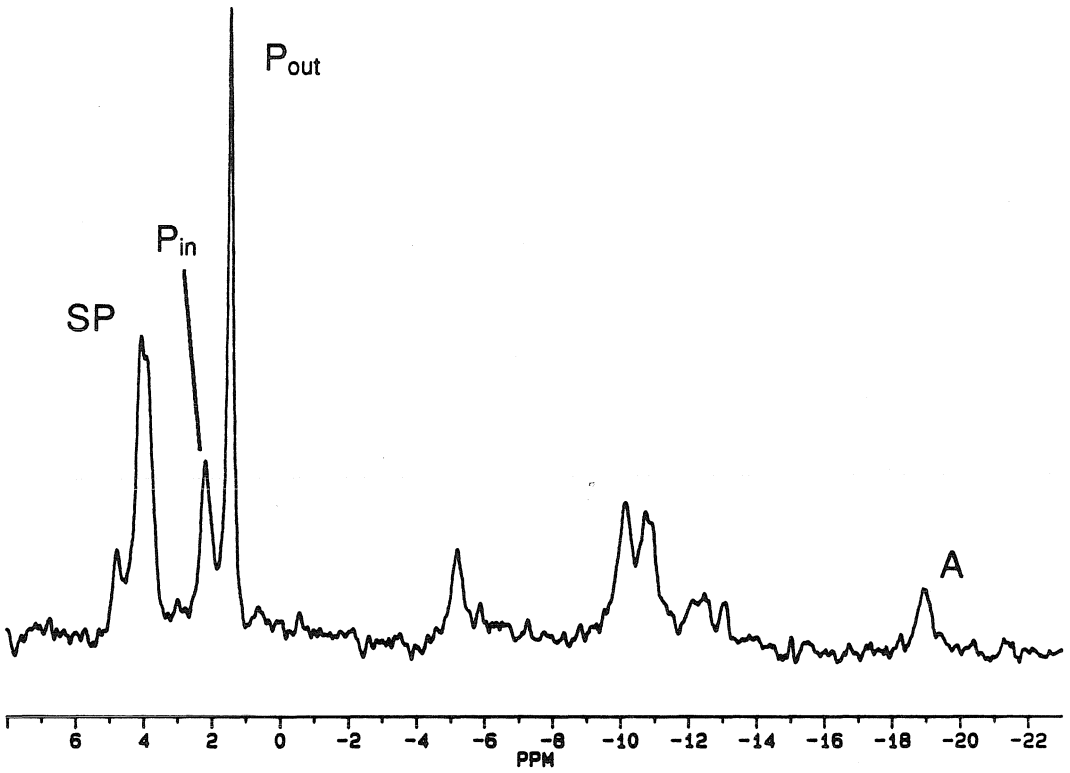

Figure 4.

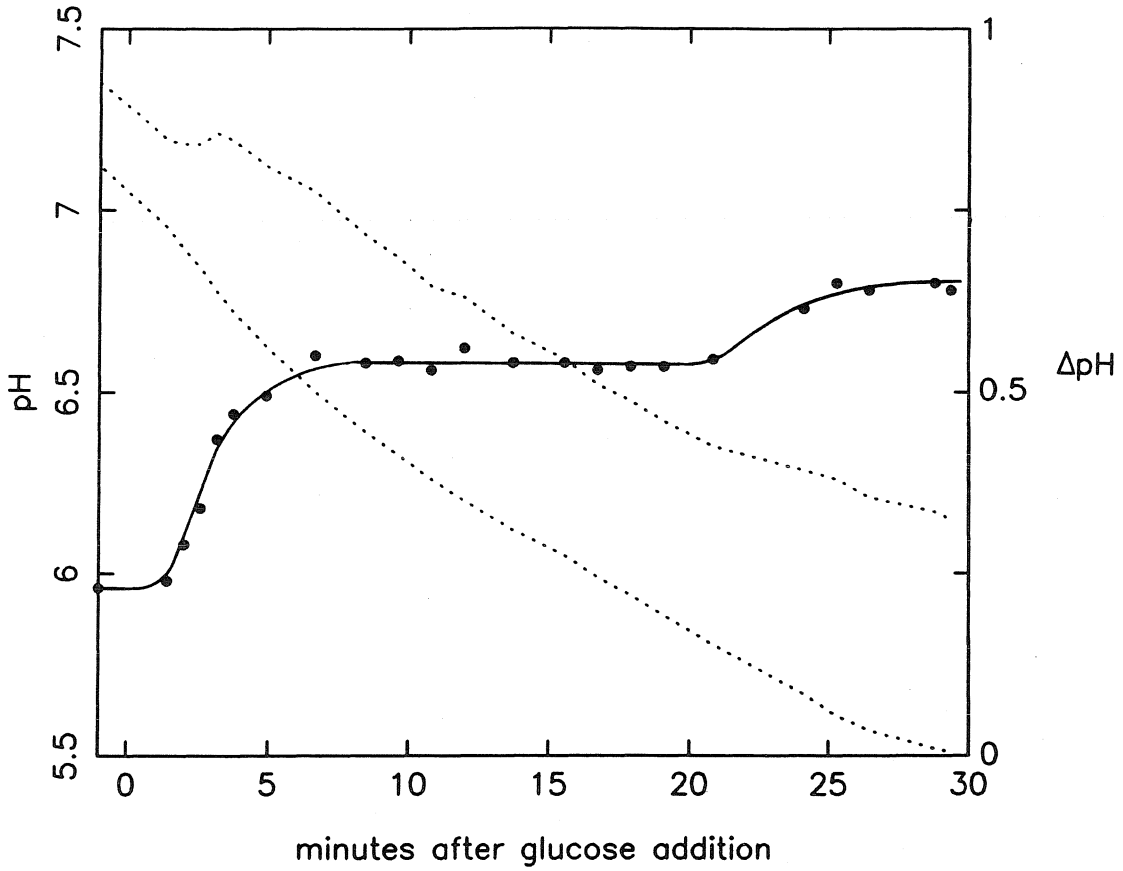
7.5
7
6.5
6
5.5
pH
1
0.5
0
ΔpH
0
5
10
15
20
25
30
minutes after glucose addition

Figure 5.

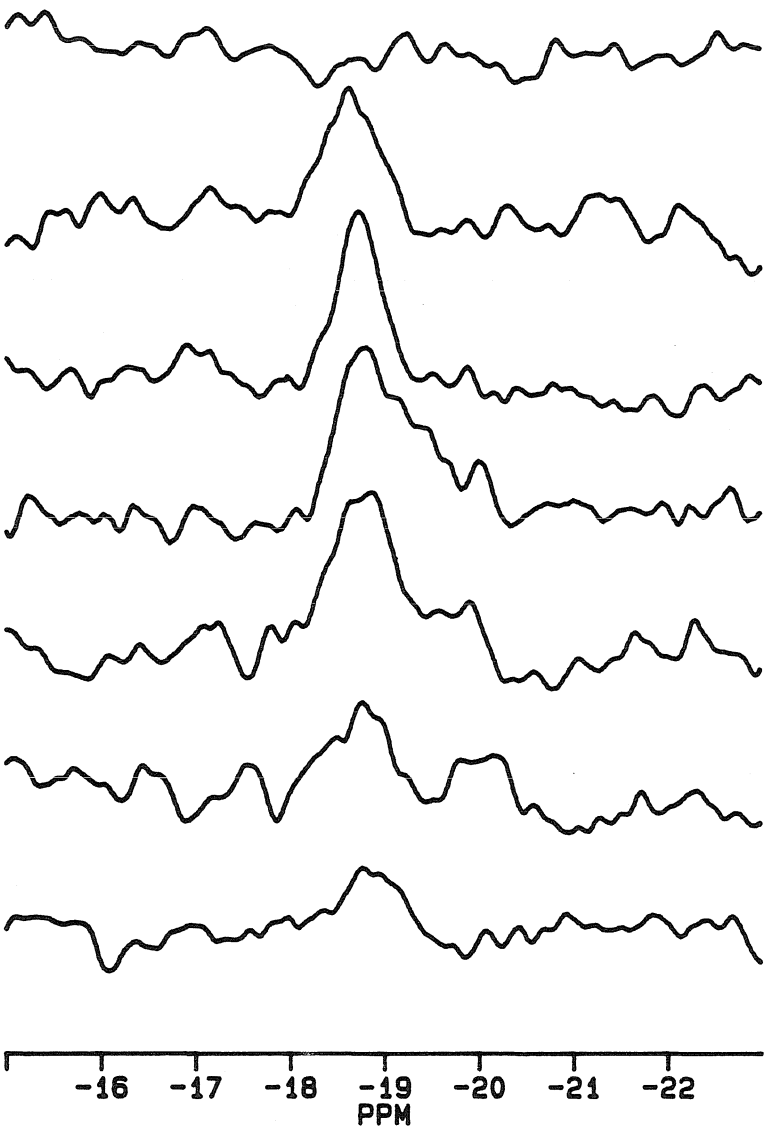

Figure 6.

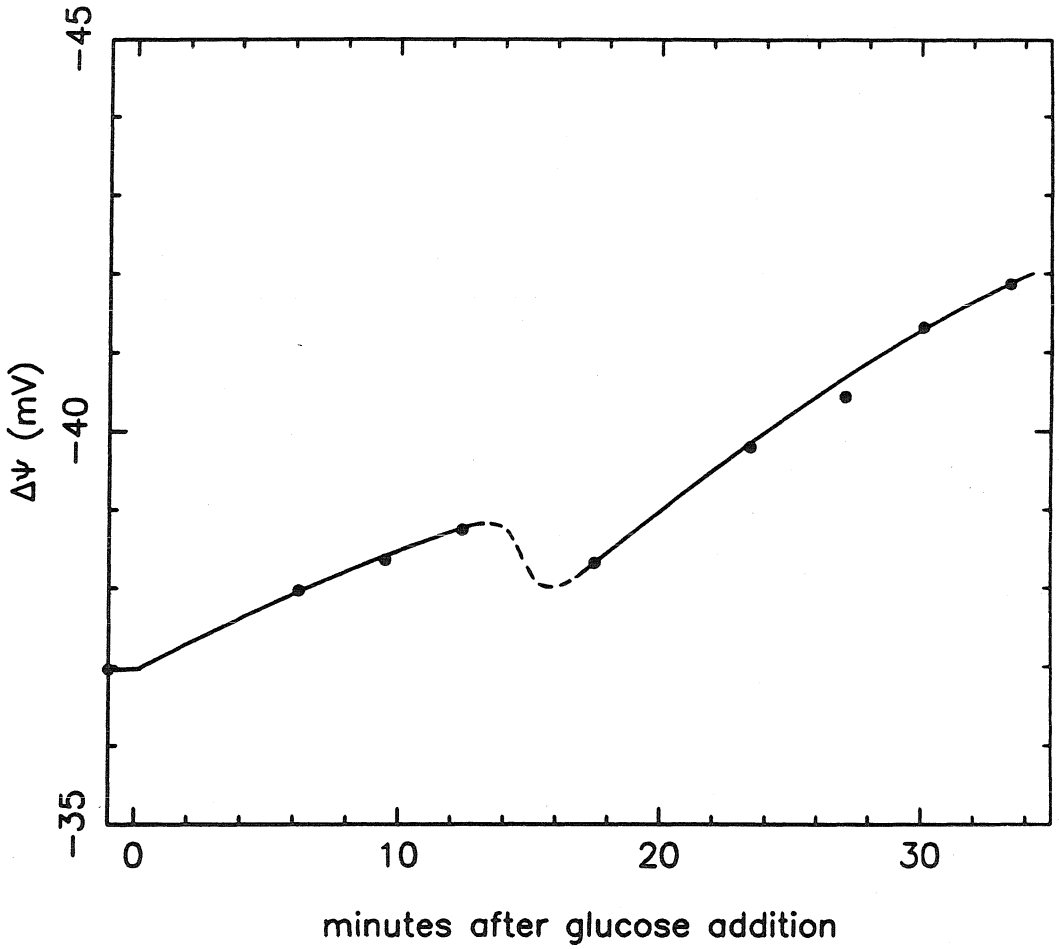

Figure 7.

APPENDIX A.

SOURCE CODE

FOR MODEL OF *rpoB* EXPRESSION

```
C     Program RPOB.FOR (Vax fortran)
C
C     Driver program for modeling regulation of synthesis of the
C     RNA polymerase beta subunit in Escherichia coli.  Differential
C     equations are solved by the Gear's backward difference method
C     using the IMSL DIVPAG routine (version 1.0, April 1987).
C
      INTEGER DDEV,COUNT,IDO,I,J,K,N
      REAL*8 M(0:5),MARCH(1:21,0:4,-20:0),DVDR(0:279),TOL,DISUM(0:23)
      REAL*8 DARCH(-20:0),TIME,TSOLVE,STEP,PARAM(50),A(1,1),SSRTX
      REAL*8 PHI1,PHI2,KF(1:4),KR(1:4),TMP1,TMP2,TMP3,DVTR(0:279)
      REAL*8 SSTDR,ACTDR,EXTDR,T,RARCH(-2:0),DI(20),P
      COMMON DARCH,RARCH,TIME,KF,KR,PHI1,PHI2,STEP,
     *       SSRTX,DDEV,DVTR,DISUM,J
      EXTERNAL DEQS,FCNJ
      OPEN(UNIT=1,FILE='CPS.PL',STATUS='OLD')
      OPEN(UNIT=2,FILE='KVAL.DAT',STATUS='OLD')
      OPEN(UNIT=3,FILE='SSLD.DAT',STATUS='OLD')
      OPEN(UNIT=4,FILE='DVTR.TAB',STATUS='OLD')
      OPEN(UNIT=5,FILE='DVDR.TAB',STATUS='OLD')
      READ(2,10) PHI1,PHI2,DDEV
      READ(2,20) (KF(I),I=1,4)
      READ(2,20) (KR(I),I=1,4)
      DO 4 I=0,4
        DO 2 J=1,19,3
    2     READ(3,30) (MARCH(N,I,0),N=J,J+2)
    4 CONTINUE
      DO 5 J=1,17,4
    5   READ(2,20) (DI(N),N=J,J+3)
      DO 6 J=0,275,5
    6   READ(4,40) (DVTR(K),K=J,J+4)
      DO 8 J=0,275,5
    8   READ(5,40) (DVDR(K),K=J,J+4)
   10 FORMAT(2(D13.6),I3)
   20 FORMAT(4(D13.6))
   30 FORMAT(3(D14.7))
   40 FORMAT(5(D14.7))
C ******** Set remaining physical parameters.
      P=1.33D-5                              !!P = protein  conc. at TIME.
      SSRTX=.25*PHI1*(P/100)*1.8066D+12      !!txpts per 5000 cells per min.
      SSTDR=SSRTX-2.398472D-10*1.8066D+12 !!txpts per 5000 cells per min.
      DO 42 K=-20,0
   42   DARCH(K)=1.00
      DO 43 K=-2,0
   43   RARCH(K)=2.304890D-7
      DISUM(0)=0.00
      DISUM(1)=0.00
      DO 46 J=2,21                           !!DISUM(J) holds the total degr.
   46   DISUM(J)=DISUM(J-1)+DI(J-1)           !!rate constant for a txpt. in
      DISUM(22)=DISUM(21)                    !!length class J (used in DEQS).
      DISUM(23)=DISUM(21)
      DO 49 I=0,4
        DO 48 J=1,21
          DO 47 K=-20,-1
   47       MARCH(J,I,K)=MARCH(J,I,0)
   48   CONTINUE
   49 CONTINUE
C ******** Prepare for initial call to routine DIVPAG.
      STEP=6.666666666666666D-2
      DO 50 K=1,50
   50   PARAM(K)=0.00
      PARAM(1)=5.00D-3
      PARAM(3)=STEP
      PARAM(10)=2.00
```

```
      PARAM(11)=1.00D-13
      PARAM(12)=2.00
      PARAM(13)=2.00
      TOL=1.00D-6
      TIME=-4.00
      T=-4.00        !!T, not TIME, is passed as an argument.  Thus,
      COUNT=5        !!TIME corresponds to the last completed step.
C ******** TOP OF MAIN LOOP ********************************************
  100 CONTINUE
C ******** Write protein and message concentrations every 20 seconds.
      IF (COUNT.NE.5) GO TO 118
      COUNT=0
      TMP1=0.00
      DO 114 I=0,4
        DO 110 J=1,21
  110     TMP1=TMP1+MARCH(J,I,0)
  114 CONTINUE
      WRITE(1,30) TIME,TMP1/1.3840D-8,P/1.33D-5
  118 CONTINUE
C ******** Solve equations over a 4 second interval.
      DO 124 K=-20,-1                          !!Start by aging archive by
        DO 122 I=0,4                           !!one STEP.
          DO 120 J=1,21
  120       MARCH(J,I,K)=MARCH(J,I,K+1)
  122   CONTINUE
  124 CONTINUE
      DO 200 J=1,22                            !!(top of DIVPAG loop)
        IF(J.NE.1) GO TO 128                   !!See note below on J=22.
        DO 125 I=0,4                           !!Set M(I) elements to pre-DIVPAG
  125     M(I)=0.00                            !!values.
        GO TO 132
  128   CONTINUE
        DO 132 I=0,4
          M(I)=MARCH(J-1,I,-1)
  132   CONTINUE
        M(5)=P                                 !!Set M(5) and T to current values.
        T=TIME
        TSOLVE=TIME+STEP
        IDO=1
        CALL DIVPAG(IDO,6,DEQS,FCNJ,A,T,TSOLVE,TOL,PARAM,M)
        IF(J.EQ.22) GO TO 140
        DO 134 I=0,4
  134     MARCH(J,I,0)=M(I)                    !!Archive results.
        GO TO 150
  140   CONTINUE                               !!J=22 is used for computational
        DO 150 I=0,4                           !!purposes to distinguish between
          MARCH(21,I,0)=MARCH(21,I,0)+M(I)     !!previous J=20 and J=21.
  150   CONTINUE
        IDO=3                                  !!Clear IMSL workspace.
        CALL DIVPAG(IDO,6,DEQS,FCNJ,A,T,TSOLVE,TOL,PARAM,M)
  200 CONTINUE
      TIME=TIME+STEP
      COUNT=COUNT+1
      P=M(5)
C ******** Update archive containing the ratio of actual degradaion rate
C ******** to expected degradation rate, DARCH(-20:0).
      DO 210 K=-20,-1
  210   DARCH(K)=DARCH(K+1)
C ******** Compute DARCH(0) from current message distribution.
      DARCH(0)=1.00
      IF (TIME.GE.0.00) GO TO 300
      TMP1=0.00
      DO 250 J=2,21
        DO 220 I=0,4
```

```
  220      TMP1=TMP1+MARCH(J,I,0)*DISUM(J)
  250 CONTINUE
      EXTDR=TMP1*1.8066D+12                    !!txpts per 5000 cells per min.
C ******** Verify that (EXTDR-SSTDR) has the expected sign.
      IF ((EXTDR/SSTDR)**DDEV.LT.0.999) STOP
     *  ' (EXTDR/SSTDR)**DDEV.LT.1'
      K=IIDINT(DABS(DLOG(EXTDR/SSTDR)/4.987541511D-3))
      TMP1=SSTDR*1.005**(DDEV*K)
      TMP2=TMP1*1.005**(DDEV)
      ACTDR=DVDR(K)+(DVDR(K+1)-DVDR(K))*(EXTDR-TMP1)/
     *        (TMP2-TMP1)                      !!txpts per 5000 cells per min.
      DARCH(0)=ACTDR/EXTDR
  300 CONTINUE
C ******** Update protein synthesis rate archive, RARCH(-2:0).
      RARCH(-2)=RARCH(-1)
      RARCH(-1)=RARCH(0)
      RARCH(0)=0.00
C ******** Compute RARCH(0) from the archived message length distributions
C ******** and archived ACTDR/EXTDR values (in DARCH vector).
      DO 400 J=1,21
        TMP2=0.00
        TMP3=0.00
        DO 304 I=0,3
  304     TMP3=TMP3+MARCH(J,I,-20)             !!Total avlbl msg in class J.
        IF(J.EQ.21) GO TO 310
        DO 310 N=1,21-J                        !!Index N corresponds to time.
          TMP1=(DARCH(N-21)+DARCH(N-20))/2.0   !!TMP1 is ACT/EXP ratio.
          TMP2=TMP2+DI(J+N-1)*TMP1/2.00        !!TMP2 holds Xi summation.
          DO 306 K=N,J+N-2
  306       TMP2=TMP2+DI(K)*TMP1
  310   CONTINUE
        IF(J.EQ.1) GO TO 314
        DO 314 N=22-J,20
          TMP1=(DARCH(N-21)+DARCH(N-20))/2.0
          DO 312 K=N,20
  312       TMP2=TMP2+DI(K)*TMP1
  314   CONTINUE
        RARCH(0)=RARCH(0)+TMP3*DEXP(-STEP*TMP2)
  400 CONTINUE
      RARCH(0)=9.8851315D-1*PHI2*RARCH(0)      !!Final value (.98...=dil. fact.).
C ******** End if TIME is greater than 80 minutes.
      IF (TIME.LT.80.0) GO TO 100
      END
C
C
      SUBROUTINE DEQS(NEQ,T,M,DMDT)
C
C ******** Given length distribution of viable message (in terms
C ******** of concentrations of the different length classes), the
C ******** protein concentration, the time, and a record of message
C ******** length distribution and specific degradation rate, compute
C ******** time derivatives.
C
      INTEGER I,J,K,N,NEQ,DDEV
      REAL*8 M(0:5),DMDT(0:5),KF(1:4),KR(1:4)
      REAL*8 PHI1,PHI2,THETA1,TIME,SSRTX,DRAT,DISUM(0:23)
      REAL*8 RPSYN,RARCH(-2:0),MU,P,STEP,T,INTRP,DEG
      REAL*8 DVTR(0:279),DEX1,DEX2,RPEX1,RPEX2,DARCH(-20:0)
      REAL*8 EXRTX,ACRTX,TMP1,TMP2
      COMMON DARCH,RARCH,TIME,KF,KR,PHI1,PHI2,STEP,
     *         SSRTX,DDEV,DVTR,DISUM,J
C
      IF (T.LT.TIME) STOP ' T .LT. TIME!'
      IF (T.GT.TIME+2*STEP) STOP ' T .GT. TIME+2*STEP!'
```

```
      MU=1.733D-2
      P=M(5)/100.00                          !!P=free protein conc.
      THETA1=0.25+0.30*(1.00-P/1.33D-7)
C ******** Since evaluation of RPSYN and DRAT is somewhat costly, it
C ******** is done only once per step (in the main program), and DEQS
C ******** obtains values by extrapolation (over less than 4 seconds).
C ******** Evaluation of ACRTX is inexpensive, so it is done in DEQS.
C
C ******** Compute actual tx rate ACRTX using Poisson/Sterling formula.
      ACRTX=PHI1*THETA1*P                    !!During recovery ACRTX=EXRTX
      IF (T.GE.0.00) GO TO 200
C ******** Verify that (EXRTX-SSRTX) has the expected sign.
      EXRTX=PHI1*THETA1*P*1.8066D+12         !!txpts. per 5000 cells per min.
      IF ((SSRTX/EXRTX)**DDEV.LT.0.999) STOP
     *  ' (SSRTX/EXRTX)**DDEV.LT.1'
      K=IIDINT(DABS(DLOG(EXRTX/SSRTX)/4.987541511D-3))
      TMP1=SSRTX*1.005**(-DDEV*K)
      TMP2=TMP1*1.005**(-DDEV)
      ACRTX=DVTR(K)+(DVTR(K+1)-DVTR(K))*(EXRTX-TMP1)/
     *      (TMP2-TMP1)                      !!txpts. per 5000 cells per min.
      ACRTX=ACRTX/1.8066D+12                 !!Standard units (M/min).
  200 CONTINUE
C ******** Estimate the rate of protein synthesis, RPSYN, by
C ******** second order extrapolation.
      N=IIDINT((T-TIME)/STEP)                !!N is either 0 or 1.
      INTRP=(T-TIME)/STEP-N
      RPEX1=RARCH(-2)+3*(RARCH(0)-RARCH(-1))    !!(EXtrapolated value)
      RPEX2=RARCH(-1)+3*(RPEX1-RARCH(0))        !!(EXtrapolated value)
      IF (N.EQ.0) RPSYN=RARCH(0)+INTRP*(RPEX1-RARCH(0))
      IF (N.EQ.1) RPSYN=RPEX1+INTRP*(RPEX2-RPEX1)
C ******** Estimate the actual/expected degradation ratio, DRAT, from
C ******** DARCH values by second order extrapolation.
      DEX1=DARCH(-2)+3*(DARCH(0)-DARCH(-1))     !!(EXtrapolated value)
      DEX2=DARCH(-1)+3*(DEX1-DARCH(0))          !!(EXtrapolated value)
      IF (N.EQ.0) DRAT=DARCH(0)+INTRP*(DEX1-DARCH(0))
      IF (N.EQ.1) DRAT=DEX1+INTRP*(DEX2-DEX1)
      IF (T.GE.0.00) DRAT=1.00               !!ACSDR is fixed during recovery.
C ******** Differential equations to be solved:
C
      DEG=DISUM(J-1+N)+INTRP*(DISUM(J+N)-DISUM(J-1+N))
      DEG=DEG*DRAT                           !!Actual degradative rate const.
      DMDT(0)=KR(1)*M(1)-(KF(1)*P+MU+DEG)*M(0)
      DO 240 I=1,3
  240   DMDT(I)=KF(I)*P*M(I-1)+KR(I+1)*M(I+1)-(KF(I+1)*P+KR(I)
     *     +MU+DEG)*M(I)
      DMDT(4)=KF(4)*P*M(3)-(KR(4)+MU+DEG)*M(4)
      IF(J.EQ.1) DMDT(0)=DMDT(0)+ACRTX
      DMDT(5)=RPSYN-MU*M(5)                  !!DMDT(5)=DPDT
      RETURN
      END
C
C
      SUBROUTINE FCNJ(NEQ,X,Y,DYPDY)
      INTEGER NEQ
      REAL*8 X,Y(NEQ),DYPDY(*)
C ******** This subroutine is never called.
      RETURN
      END
```

```
C     Program RPOBSS.FOR (Vax fortran)
C
C     This program is used to generate values for the SSLD matrix
C     (steady-state length distribution) used in the main program.
C     Given kinetic constants, this program solves for the transcript
C     length distribution at steady state.
C
      INTEGER DDEV,COUNT,IDO,I,J,K,N
      REAL*8 M(0:5),MARCH(1:21,0:4,-20:0),DYDR(0:279),TOL,DISUM(0:23)
      REAL*8 DARCH(-20:0),TIME,TSOLVE,STEP,PARAM(50),A(1,1),SSRTX
      REAL*8 PHI1,PHI2,KF(1:4),KR(1:4),TMP1,TMP2,TMP3,DVTR(0:279)
      REAL*8 SSTDR,ACTDR,EXTDR,T,RARCH(-2:0),DI(20),P
      COMMON DARCH,RARCH,TIME,KF,KR,PHI1,PHI2,STEP,
     *          SSRTX,DDEV,DVTR,DISUM,J
      EXTERNAL DEQS,FCNJ
      OPEN(UNIT=1,FILE='CPSS.OUT',STATUS='OLD')
      OPEN(UNIT=2,FILE='KVAL.DAT',STATUS='OLD')
      OPEN(UNIT=3,FILE='SSLD.DAT',STATUS='OLD')
      OPEN(UNIT=4,FILE='SSLD.PRE',STATUS='OLD')
      READ(2,10) PHI1,PHI2,DDEV
      READ(2,20)(KF(I),I=1,4)
      READ(2,20)(KR(I),I=1,4)
      DO 4 I=0,4
        DO 2 J=1,19,3
    2     READ(4,30)(MARCH(N,I,0),N=J,J+2)
    4 CONTINUE
      DO 5 J=1,17,4
    5   READ(2,20)(DI(N),N=J,J+3)
   10 FORMAT(2(D13.6),I3)
   20 FORMAT(4(D13.6))
   30 FORMAT(3(D14.7))
C ******** Set remaining physical parameters.
      P=1.33D-5                                !!P = protein  conc. at TIME.
      SSRTX=.25*PHI1*(P/100)*1.8066D+12        !!txpts per 5000 cells per min.
      SSTDR=SSRTX-2.398472D-10*1.8066D+12      !!txpts per 5000 cells per min.
      DO 42 K=-20,0
   42   DARCH(K)=1.00
      DO 43 K=-2,0
   43   RARCH(K)=2.304890D-7
      DISUM(0)=0.00
      DISUM(1)=0.00
      DO 46 J=2,21                             !!DISUM(J) holds the total degr.
   46   DISUM(J)=DISUM(J-1)+DI(J-1)             !!rate constant for a txpt. in
      DISUM(22)=DISUM(21)                      !!length class J (used in DEQS).
      DISUM(23)=DISUM(21)
      DO 49 I=0,4
        DO 48 J=1,21
          DO 47 K=-20,-1
   47       MARCH(J,I,K)=MARCH(J,I,0)
   48   CONTINUE
   49 CONTINUE
C ******** Prepare for initial call to routine DIVPAG.
      STEP=6.666666666666666D-2
      DO 50 K=1,50
   50   PARAM(K)=0.00
      PARAM(1)=5.00D-3
      PARAM(3)=STEP
      PARAM(10)=2.00
      PARAM(11)=1.00D-19
      PARAM(12)=2.00
      PARAM(13)=2.00
      TOL=1.00D-6
      TIME=0.00
      T=0.00          !!T, not TIME, is passed as an argument.  Thus,
```

```
      COUNT=30       !!TIME corresponds to the last completed step.
C ******** TOP OF MAIN LOOP ***************************************
  100 CONTINUE
C ******** Write protein and message concentrations every 120 seconds.
      IF (COUNT.NE.30) GO TO 118
      COUNT=0
      TMP1=0.00
      TMP2=0.00
      DO 114 I=0,4
        DO 110 J=1,21
          IF(I.NE.1) TMP2=TMP2+MARCH(J,I,0)
  110     TMP1=TMP1+MARCH(J,I,0)
  114 CONTINUE
      WRITE(1,30) TMP1/1.3840D-8,TMP2/1.3840D-8,RARCH(0)
      DO 118 I=0,4
        DO 116 J=1,19,3
  116     WRITE(3,30)(MARCH(N,I,0),N=J,J+2)
  118 CONTINUE
C ******** Solve equations over a 4 second interval.
      DO 124 K=-20,-1                          !!Start by aging archive by
        DO 122 I=0,4                           !!one STEP.
          DO 120 J=1,21
  120       MARCH(J,I,K)=MARCH(J,I,K+1)
  122   CONTINUE
  124 CONTINUE
      DO 200 J=1,22                            !!(top of DIVPAG loop)
        IF(J.NE.1) GO TO 128                   !!See note below on J=22.
        DO 125 I=0,4                           !!Set M(I) elements to pre-DIVPAG
  125     M(I)=0.00                            !!values.
        GO TO 132
  128   CONTINUE
        DO 132 I=0,4
          M(I)=MARCH(J-1,I,-1)
  132   CONTINUE
        M(5)=P                                 !!Set M(5) and T to current values.
        T=TIME
        TSOLVE=TIME+STEP
        IDO=1
        CALL DIVPAG(IDO,8,DEQS,FCNJ,A,T,TSOLVE,TOL,PARAM,M)
        IF(J.EQ.22) GO TO 140
        DO 134 I=0,4
  134     MARCH(J,I,0)=M(I)                    !!Archive results.
        GO TO 150
  140   CONTINUE                               !!J=22 is used for computational
        DO 150 I=0,4                           !!purposes to distinguish between
          MARCH(21,I,0)=MARCH(21,I,0)+M(I)     !!previous J=20 and J=21.
  150   CONTINUE
        IDO=3                                  !!Clear IMSL workspace.
        CALL DIVPAG(IDO,8,DEQS,FCNJ,A,T,TSOLVE,TOL,PARAM,M)
  200 CONTINUE
      TIME=TIME+STEP
      COUNT=COUNT+1
      P=M(5)
C ******** Update archive containing the ratio of actual degradaion rate
C ******** to expected degradation rate, DARCH(-20:0).
      DO 210 K=-20,-1
  210   DARCH(K)=DARCH(K+1)
C ******** Compute DARCH(0) from current message distribution.
      DARCH(0)=1.00
      IF (TIME.GE.0.00) GO TO 300
      TMP1=0.00
      DO 250 J=2,21
        DO 220 I=0,4
  220     TMP1=TMP1+MARCH(J,I,0)*DISUM(J)
```

```
  250 CONTINUE
      EXTDR=TMP1*1.8066D+12                    !!txpts per 5000 cells per min.
C ******** Verify that (EXTDR-SSTDR) has the expected sign.
      IF ((EXTDR/SSTDR)**DDEV.LT.0.999) STOP
     *  ' (EXTDR/SSTDR)**DDEV.LT.1'
      K=IIDINT(DABS(DLOG(EXTDR/SSTDR)/4.987541511D-3))
      TMP1=SSTDR*1.005**(DDEV*K)
      TMP2=TMP1*1.005**(DDEV)
      ACTDR=DVDR(K)+(DVDR(K+1)-DVDR(K))*(EXTDR-TMP1)/
     *      (TMP2-TMP1)                        !!txpts per 5000 cells per min.
      DARCH(0)=ACTDR/EXTDR
  300 CONTINUE
C ******** Update protein synthesis rate archive, RARCH(-2:0).
      RARCH(-2)=RARCH(-1)
      RARCH(-1)=RARCH(0)
      RARCH(0)=0.00
C ******** Compute RARCH(0) from the archived message length distributions
C ******** and archived ACTDR/EXTDR values (in DARCH vector).
      DO 400 J=1,21
        TMP2=0.00
        TMP3=0.00
        DO 304 I=0,3
  304     TMP3=TMP3+MARCH(J,I,-20)             !!Total avlbl msg in class J.
        IF(J.EQ.21) GO TO 310
        DO 310 N=1,21-J                        !!Index N corresponds to time.
          TMP1=(DARCH(N-21)+DARCH(N-20))/2.0   !!TMP1 is ACT/EXP ratio.
          TMP2=TMP2+DI(J+N-1)*TMP1/2.00        !!TMP2 holds Xi summation.
          DO 306 K=N,J+N-2
  306       TMP2=TMP2+DI(K)*TMP1
  310   CONTINUE
        IF(J.EQ.1) GO TO 314
        DO 314 N=22-J,20
          TMP1=(DARCH(N-21)+DARCH(N-20))/2.0
          DO 312 K=N,20
  312       TMP2=TMP2+DI(K)*TMP1
  314   CONTINUE
        RARCH(0)=RARCH(0)+TMP3*DEXP(-STEP*TMP2)
  400 CONTINUE
      RARCH(0)=9.8851315D-1*PHI2*RARCH(0)      !!Final value (.98...=dil. fact.).
C ******** End if TIME is greater than 4 minutes.
      IF (TIME.LT.4.01) GO TO 100
      DO 420 I=0,4
        TMP1=0.00
        TMP2=0.00
        DO 410 J=1,21
          TMP1=TMP1+MARCH(J,I,0)
  410     TMP2=TMP2+DISUM(J)*MARCH(J,I,0)
  420   WRITE(1,30)TMP1,TMP2,TMP2/TMP1
      END
C
C
      SUBROUTINE DEQS(NEQ,T,M,DMDT)
C
C ******** Given length distribution of viable message (in terms
C ******** of concentrations of the different length classes), the
C ******** protein concentration, the time, and a record of message
C ******** length distribution and specific degradation rate, compute
C ******** time derivatives.
C
      INTEGER I,J,K,N,NEQ,DDEV
      REAL*8 M(0:5),DMDT(0:5),KF(1:4),KR(1:4)
      REAL*8 PHI1,PHI2,THETA1,TIME,SSRTX,DRAT,DISUM(0:23)
      REAL*8 RPSYN,RARCH(-2:0),MU,P,STEP,T,INTRP,DEG
      REAL*8 DVTR(0:279),DEX1,DEX2,RPEX1,RPEX2,DARCH(-20:0)
```

```
      REAL*8 EXRTX,ACRTX,TMP1,TMP2
      COMMON DARCH,RARCH,TIME,KF,KR,PHI1,PHI2,STEP,
     *         SSRTX,DDEV,DVTR,DISUM,J
C
      IF (T.LT.TIME) STOP ' T .LT. TIME!'
      IF (T.GT.TIME+2*STEP) STOP ' T .GT. TIME+2*STEP!'
      MU=1.733D-2
      P=1.33D-7                          !!P=free pol. conc.
      THETA1=0.25
C ******** Since evaluation of RPSYN and DRAT is somewhat costly, it
C ******** is done only once per step (in the main program), and DEQS
C ******** obtains values by extrapolation (over less than 4 seconds).
C ******** Evaluation of ACRTX is inexpensive, so it is done in DEQS.
C
C ******** Compute actual tx rate ACRTX using Poisson/Sterling formula.
      ACRTX=PHI1*THETA1*P                !!During recovery ACRTX=EXRTX
      IF (T.GE.0.00) GO TO 200
C ******** Verify that (EXRTX-SSRTX) has the expected sign.
      EXRTX=PHI1*THETA1*P*1.8066D+12     !!txpts. per 5000 cells per min.
      IF ((SSRTX/EXRTX)**DDEV.LT.0.999) STOP
     *  ' (SSRTX/EXRTX)**DDEV.LT.1'
      K=IIDINT(DABS(DLOG(EXRTX/SSRTX)/4.987541511D-3))
      TMP1=SSRTX*1.005**(-DDEV*K)
      TMP2=TMP1*1.005**(-DDEV)
      ACRTX=DVTR(K)+(DVTR(K+1)-DVTR(K))*(EXRTX-TMP1)/
     *       (TMP2-TMP1)                 !!txpts. per 5000 cells per min.
      ACRTX=ACRTX/1.8066D+12             !!Standard units (M/min).
  200 CONTINUE
C ******** Estimate the rate of protein synthesis, RPSYN, by
C ******** second order extrapolation.
      N=IIDINT((T-TIME)/STEP)            !!N is either 0 or 1.
      INTRP=(T-TIME)/STEP-N
      RPEX1=RARCH(-2)+3*(RARCH(0)-RARCH(-1))   !!(EXtrapolated value)
      RPEX2=RARCH(-1)+3*(RPEX1-RARCH(0))       !!(EXtrapolated value)
      IF (N.EQ.0) RPSYN=RARCH(0)+INTRP*(RPEX1-RARCH(0))
      IF (N.EQ.1) RPSYN=RPEX1+INTRP*(RPEX2-RPEX1)
C ******** Estimate the actual/expected degradation ratio, DRAT, from
C ******** DARCH values by second order extrapolation.
      DEX1=DARCH(-2)+3*(DARCH(0)-DARCH(-1))    !!(EXtrapolated value)
      DEX2=DARCH(-1)+3*(DEX1-DARCH(0))         !!(EXtrapolated value)
      IF (N.EQ.0) DRAT=DARCH(0)+INTRP*(DEX1-DARCH(0))
      IF (N.EQ.1) DRAT=DEX1+INTRP*(DEX2-DEX1)
      IF (T.GE.0.00) DRAT=1.00           !!ACSDR is fixed during recovery.
C
      DEG=DISUM(J-1+N)+INTRP*(DISUM(J+N)-DISUM(J-1+N))
      DEG=DEG*DRAT                       !!Actual degradative rate const.
C ******** Differential equations to be solved:
C
      DMDT(0)=KR(1)*M(1)-(KF(1)*P+MU+DEG)*M(0)
      DO 240 I=1,3
  240   DMDT(I)=KF(I)*P*M(I-1)+KR(I+1)*M(I+1)-(KF(I+1)*P+KR(I)
     *    +MU+DEG)*M(I)
      DMDT(4)=KF(4)*P*M(3)-(KR(4)+MU+DEG)*M(4)
      IF(J.EQ.1) DMDT(0)=DMDT(0)+ACRTX
      DMDT(5)=0.00                       !!DMDT(5)=DPDT
      RETURN
      END
C
C
      SUBROUTINE FCNJ(NEQ,X,Y,DYPDY)
      INTEGER NEQ
      REAL*8 X,Y(NEQ),DYPDY(*)
C ******** This subroutine is never called.
      RETURN
      END
```

```
c     STRL.FOR (vax fortran) computes and tabulates values of deviated
c     rate according to expected rate, for a given probability value.
c     The Stirling approximation is used in conjunction with the Poisson
c     formula.
c
      INTEGER I,J
      REAL OUT(0:279)
      REAL*8 LP,RSS,REX,RDQ,RD1,RD2,LPQ,LP1,LP2,K
      OPEN(UNIT=1,FILE='STRL.INP',STATUS='OLD')
      OPEN(UNIT=2,FILE='STRL1.OUT',STATUS='OLD')
      OPEN(UNIT=3,FILE='STRL2.OUT',STATUS='OLD')
      READ (1,40) LP,RSS
   40 FORMAT(2(10X,D13.6,/))
C ******************** EXR .LE. SSR  .AND.  ACR .GE. EXR ************
      DO 200 K=0,1
        REX=RSS*(1.005)**(-K)
        RD2=REX
        LP2=-0.500000000000*DLOG(6.283185308*REX)
        RD1=REX*4
        LP1=RD1-REX+RD1*(DLOG(REX)-DLOG(RD1))-
     *       0.500000000000000*DLOG(RD1)-0.9189385335
   60   CONTINUE
          RDQ=RD1+(RD2-RD1)*(LP-LP1)/(LP2-LP1)
          LPQ=RDQ-REX+RDQ*(DLOG(REX)-DLOG(RDQ))-
     *         0.500000000000000*DLOG(RDQ)-0.9189385335
          IF (DABS(LPQ-LP).LT.0.001) GO TO 100
          RD2=RD1
          LP2=LP1
          RD1=RDQ
          LP1=LPQ
        GO TO 60
  100   CONTINUE
        OUT(K)=RDQ
  200 CONTINUE
      DO 400 K=2,279
        REX=RSS*(1.005)**(-K)
        RD2=OUT(K-2)
        LP2=RD2-REX+RD2*(DLOG(REX)-DLOG(RD2))-
     *       0.500000000000000*DLOG(RD2)-0.9189385335
        RD1=OUT(K-1)
        LP1=RD1-REX+RD1*(DLOG(REX)-DLOG(RD1))-
     *       0.500000000000000*DLOG(RD1)-0.9189385335
  260   CONTINUE
          RDQ=RD1+(RD2-RD1)*(LP-LP1)/(LP2-LP1)
          LPQ=RDQ-REX+RDQ*(DLOG(REX)-DLOG(RDQ))-
     *         0.500000000000000*DLOG(RDQ)-0.9189385335
          IF (DABS(LPQ-LP).LT.0.001) GO TO 300
          RD2=RD1
          LP2=LP1
          RD1=RDQ
          LP1=LPQ
        GO TO 260
  300   CONTINUE
        OUT(K)=RDQ
  400 CONTINUE
      DO 401 J=0,275,5
  401   WRITE (2,804)(OUT(I),I=J,J+4)
  804 FORMAT(5(D14.7))
C ******************** EXR .GE. SSR  .AND.  ACR .LE. EXR ************
      DO 2000 K=0,1
        REX=RSS*(1.005)**(K)
        RD2=REX
        LP2=-0.500000000000*DLOG(6.283185308*REX)
        RD1=REX/4.000
```

```
      LP1=RD1-REX+RD1*(DLOG(REX)-DLOG(RD1))-
     *       0.5000000000000*DLOG(RD1)-0.9189385335
  600   CONTINUE
        RDQ=RD1+(RD2-RD1)*(LP-LP1)/(LP2-LP1)
        LPQ=RDQ-REX+RDQ*(DLOG(REX)-DLOG(RDQ))-
     *       0.5000000000000*DLOG(RDQ)-0.9189385335
        IF (DABS(LPQ-LP).LT.0.001) GO TO 1000
        RD2=RD1
        LP2=LP1
        RD1=RDQ
        LP1=LPQ
      GO TO 600
 1000   CONTINUE
      OUT(K)=RDQ
 2000 CONTINUE
     DO 4000 K=2,279
      REX=RSS*(1.005)**(K)
      RD2=OUT(K-2)
      LP2=RD2-REX+RD2*(DLOG(REX)-DLOG(RD2))-
     *       0.5000000000000*DLOG(RD2)-0.9189385335
      RD1=OUT(K-1)
      LP1=RD1-REX+RD1*(DLOG(REX)-DLOG(RD1))-
     *       0.5000000000000*DLOG(RD1)-0.9189385335
 2600   CONTINUE
        RDQ=RD1+(RD2-RD1)*(LP-LP1)/(LP2-LP1)
        LPQ=RDQ-REX+RDQ*(DLOG(REX)-DLOG(RDQ))-
     *       0.5000000000000*DLOG(RDQ)-0.9189385335
        IF (DABS(LPQ-LP).LT.0.001) GO TO 3000
        RD2=RD1
        LP2=LP1
        RD1=RDQ
        LP1=LPQ
      GO TO 2600
 3000   CONTINUE
      OUT(K)=RDQ
 4000 CONTINUE
     DO 4001 J=0,275,5
 4001   WRITE (3,804)(OUT(I),I=J,J+4)
     END
```