J. theor. Biol. (1993) **162,** 153–186

Variation in Concentrations of RNAs and Proteins Involved in Gene Expression of *Escherichia coli*

JOSEPH M. MAHAFFY†

Department of Mathematical Sciences, San Diego State University, San Diego, CA 92182, *U.S.A.*

(*Received on* 18 *May* 1991, *Accepted on* 26 *August* 1992)

A growing cell is not a steady-state situation. At some point in the cell cycle each gene doubles its concentration, which may result in a doubling of the production rate of its mRNA. The variations in concentration of mRNAs and proteins in an individual exponentially growing cell are studied for a gene which is constitutive, autorepressed, or autoactivated. Analysis of the mathematical model for a constitutive gene shows a fixed variation in concentration of a stable RNA between its minimum and maximum concentration of approximately 6%, independent of cell cycle time or gene location. The variation in the concentrations of the mRNA and protein for a constitutive gene are studied as gene position, molecular stability, and cell cycle time are varied. One result shows that doubling the growth rate can more than double the percentage of product from a gene near the origin of replication compared to one near the terminus. Results from these theoretical studies compare favorably to experimental results for rRNAs and the fully induced *lac* gene.

Additional studies were performed to determine the effects of autorepression and autoactivation on the variation of concentration of mRNAs and proteins. The studies show that RNA and protein products of an autorepressed gene have almost the same variation in concentration as products of a constitutively expressed gene though the absolute concentrations are decreased. It is shown that the stability of an mRNA or protein affects variation in its concentration throughout the cell cycle, much more than the type of genetic control. The strength of repression has no effect on the variation in concentration of RNA and protein gene products through a cell cycle. Studies of an autoactivated gene show that it is significantly more responsive to a shift up or down, such as those caused by nutritional changes. An example is provided where the autoactivated gene is not expressed at one growth rate, but turns on at a higher growth rate. Furthermore, it is shown that gene position may determine whether or not the autoactivated gene is expressed.

1. Introduction

From a biochemical point of view, the growing cell is never in a true steady state, as many of the biochemical species have constantly varying concentrations. As the chromosome is replicated, the various genes are copied at different times, which results in a punctuated production of the genes. This in turn affects the concentration of the RNAs transcribed from these genes, which after translation affects the concentration of the resulting proteins. The concentrations also vary as a result of the expanding volume from the growing cell.

† The work of this author was supported in part by NSF grants DMS-8807360 and DMS-9007718.

0022-5193/93/010153+34 $08.00

© 1993 Academic Press Limited

This is a theoretical study of the concentrations of RNAs and proteins in exponentially growing cultures of *Escherichia coli*. In particular, the focus is on an individual idealized cell as it progresses through its cell cycle. During the cell cycle, each gene doubles at some specific time. The mathematical models examine several simplified production schemes that are based on deterministic biochemical kinetics and use the gene doubling when examining the rate of RNA transcription. Our models consider RNA and protein synthesis with very basic positive, negative, and neutral feedback mechanisms. These models are simplified versions of what is actually occurring within the bacteria, but hopefully they provide an adequate description for the purposes of our study. To obtain some qualitative properties for the basic features of gene expression for a specific gene in *E. coli*, the models necessarily oversimplify some of the biochemical processes used by the cell in transcription and translation. However, we believe the assumptions represent a reasonable first approximation to what is occurring. Our study examines the variations in concentrations of the transcribed RNAs and translated proteins with respect to the kinetic parameters associated with their production and degradation, generation time, and gene position. Though this is a theoretical study, examples from experiments on the *rrn* genes and a fully induced *lac* gene for *E. coli* are used to illustrate the findings and implications.

A *constitutively* produced RNA or protein is the simplest example for observing the fluctuations in concentrations during the cell cycle, where we define constitutive production to mean a production whose rate of synthesis is solely dependent on its transcription and translation apparatus, i.e. the production has no positive or negative intracellular controls. In our model, the RNA molecule is produced at a rate directly proportional to the concentration of the genes which code for the specific RNA. Similarly, a constitutive protein has a constitutive RNA and in our model is produced at a rate directly proportional to the concentration of the RNA from which it is translated. These assumptions allow one to formulate elementary differential equations that can be readily solved. The solutions allow one to compute analytically the amount of fluctuation in concentration of a constitutively produced RNA or protein depending on the basic kinetic parameters for the model of a growing cell, such as gene position, molecular stability, and cell cycle time.

As a first approximation which allows detailed mathematical analysis, the models assume that the rates of transcription and translation are directly proportional to the concentrations of RNA polymerase and ribosomes, respectively, and that these concentrations are maintained at a constant level by the cell. Some experiments indicate that RNA polymerase must be present in excess at all growth rates except possibly the highest growth rate (Matzura *et al.*, 1973; Gausing, 1980; Shepherd *et al.*, 1980). However, Churchward *et al.* (1982) and Jensen & Pedersen (1990) suggest that the rates of transcription and translation may be the key rate-limiting steps. The latter authors have experimental evidence that elongation rates during transcription may vary resulting in competition among the promoters for a limited supply of free RNA polymerase and that some promoters may actually sequester the RNA polymerase. The stringently controlled promoters may lose activity owing to competition with other promoters for RNA polymerase (Cashel & Rudd, 1987; Jensen & Pedersen, 1990).

Similarly, the process of translation may be affected by competition among mRNAs for free ribosomes. Experimental evidence suggests that cells control their intracellular concentration of ribosomes by a negative feedback of rRNA synthesis from an excess of functioning ribosomes or some product of them (Jinks-Robertson *et al.*, 1983; Gourse *et al.*, 1985; Jinks-Robertson & Nomura, 1987). This information supports our assumption that the ribosomes maintain an almost constant concentration over the cell cycle. Because controversy remains on these issues about availability of free RNA polymerase and free ribosomes, we believe that our approximation of a constant concentration is the best modeling assumption that can be used to derive information about gene expression in *E. coli.* Competition for RNA polymerase molecules may affect our calculations for the amount of variation in concentration of RNA and protein over the cell cycle, but should not change the qualitative observations comparing types of intracellular control of RNA and protein synthesis, gene position, and molecular stability.

The results show that a constitutively produced stable RNA in an exponentially growing cell has a constant concentration fluctuation that is independent of the specific gene location or the time of the cell cycle. As expected, the concentration of RNA varies more as the stability of the molecule decreases. In addition, variation in concentration of a less stable RNA increases with an increase in the cell cycle time. The variation in concentration of the constitutive protein parallels that of its mRNA template's concentration though its variation is less pronounced. When comparing two identical genes that are located at different positions on the chromosome, the one closest to the origin of replication produces more product. Furthermore, when growth rate increases, the gene closer to the origin gives rise to even more of its gene product when compared to a gene further away.

The effects of repression of a gene by its own endproduct or *autorepression* in an exponentially growing cell are examined. Biologists have long assumed that the negative feedback mechanism of repression is crucial for maintaining a relatively constant concentration of a particular protein (Ingraham *et al.*, 1983: 339; Ptashne, 1986: 30). The analysis below shows that this assumption is false. The mathematical model assumes that the protein, which is produced from some hypothetical gene, serves as a repressor and binds to the operator region of its gene to prevent further transcription of its mRNA when there are sufficient quantities of the protein. It is assumed that this is the only control in the production of this protein. The non-linear differential equations which describe this simple biochemical production scheme are solved numerically. As expected, the strength of repression affects the total concentration of the gene products, but, surprisingly, it has no effect on the amount of variation through the cell cycle. In fact, it is the stability of the RNAs and proteins that most affects the ratio between the maximum and minimum concentrations. Autorepression is effective in minimizing the change in mean concentration of a protein over a variety of growth conditions.

Another mathematical model is developed to examine *autoactivation*, where the endproduct binds to the operator region of its gene and facilitates the expression of its gene. Similar positive control schemes have been shown to possess multiple equilibria, including one state where the gene is "off" and another where the gene is "on"

(Othmer, 1976). The behavior of this model exhibits many of the same sensitivities to the kinetic parameters of the other models. However, the autoactivated gene can be shown to be significantly more responsive to changes in production rates. Also, the position of this gene may determine whether or not it is expressed at a particular growth rate. An example is provided where the autoactivated gene is inactive at one growth rate, but turns on at a higher growth rate. This information suggests that autoactivation may be important in a cell's response to various growth conditions.

A series of calculations are performed to determine how the different controls mentioned above affect specific protein concentrations when a cell is subjected to a shift up or a shift down, such as a nutritional change. The mathematical models demonstrate that, after a nutritional shift, the mean protein concentrations change most when their production is controlled by autoactivation and least when autorepression controls their production. In fact, the response from both constitutive and autorepressed genes is rather conservative and similar. However, a shift up or shift down can affect whether or not a particular autoactivated gene is expressed depending on binding affinity, gene position, and cell cycle time.

2. Concentration Changes for a Stable RNA

In this section we examine how the cell cycle of an exponentially growing cell affects the concentration of a stable RNA which is constitutively produced. This is the simplest mathematical model that can be developed to demonstrate how a gene doubling during the cell cycle affects the concentration of its resultant RNA. The theoretical model presented here examines a completely stable RNA that is not involved in the regulation of its own production, i.e. there is no feedback which affects the binding of the DNA-dependent RNA polymerase to the promoter region of its gene. Thus, the rate of production of this stable RNA depends on the number of active genes involved in its production times, the concentration of the RNA polymerase, or whatever the concentration of the rate-limiting component of transcription is.

2.1. THE MATHEMATICAL MODEL

The dynamics of a growing cell are very complex, with many interwoven control loops, so there are several modeling assumptions that are needed to simplify the mathematical model. The mathematical model assumes that the synthesis of RNA polymerase is regulated in order to maintain a constant concentration of free polymerase throughout the cell cycle. Several experiments support the concept that RNA polymerase is in excess (Matzura *et al.*, 1973; Gausing, 1980; Shepherd *et al.*, 1980). If RNA polymerase levels are subsaturating, then competition among the promoters for available RNA polymerase should be examined. However, subsaturation of RNA polymerase remains controversial (Nomura *et al.*, 1984; Neidhardt *et al.*, 1990: 439), and qualitatively, it should not affect our analysis of an individual gene.

The model analyzes an ideal cell which divides exactly in half at the end of its generation time, τ. It is assumed that the cell is growing exponentially (Cooper,

1988*a*, *b*, 1991) though other growth conditions could easily be handled. [In Appendix A, (A.2) shows how to enter different types of growth into the model, such as the linear cell growth suggested by Kubitschek & Pai, 1988.] With the assumption of exponential growth one obtains the following expression for the volume, $V(t)$, of the cell:

$$V(t) = V_0 2^{t/\tau} = V_0 e^{kt}, \qquad 0 \le t \le \tau$$

where V_0 is the initial size of the cell, τ is the generation time, and $k = \ln(2)/\tau$.

Let $c(t)$ denote the concentration of the stable RNA which is being produced, then a differential equation governing this concentration is given by:

$$\dot{c}(t) = \frac{rG_a(t)}{V(t)} - kc(t), \qquad 0 \le t \le \tau, \tag{1}$$

where r is a rate of production per gene site which depends on the concentration of the polymerase and is assumed to be constant for a given growth rate. The function G_a represents the number of active genes producing the RNA and is discussed in more detail below. Appendix A provides more information on the derivation of this equation and includes information on how to modify the equation for non-exponential cell growth. Equations for biochemical kinetics are often derived with concentrations as a result of standard rate laws (for example, see Banks, 1975); however, biologists often prefer mass units so this conversion appears in the Appendix. Given our assumption of an ideal cell dividing at the end of its doubling time, Appendix A shows that there is a unique attracting periodic solution to (1) satisfying the periodic boundary conditions:

$$c(0) = c(\tau). \tag{2}$$

These boundary conditions yield a solution where the concentration does not change at cell division.

The function G_a depends on gene location and cell cycle time and determines the copy number and time of replication for a specific gene during the cell cycle. Our discussion uses *E. coli* as the example because so much is known about its cell cycle (Helmstetter, 1987). Application of these ideas can be easily extended to other prokaryotic cells which have similar cell cycle characteristics such as constant C and D periods and precise timing of initiation. It is assumed that initiation of replication occurs at the origin of replication (*oriC*) and proceeds bidirectionally at a constant rate until the replication forks reach the terminus. This time is the C period. When initiation occurs and there are multiple *oriC*s present, then initiation events are assumed to occur synchronously at all origins in the cell. Experimental results of Skarstad *et al.* (1986, 1988) show that the timing of initiation in *E. coli* is synchronous. After termination of DNA replication there is a period of time, the D period,

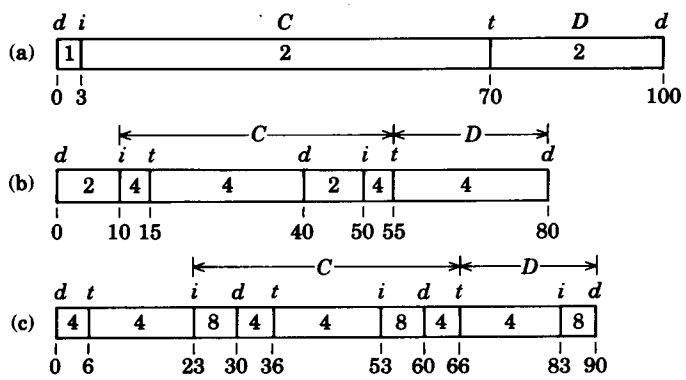

FIG. 1. Timing of initiation and termination for different cell cycle times, (a) 100 min; (b) 40 min; (c) 30 min. Below i is when initiation of DNA replication occurs, t is when the replication forks reach the terminus, and d is when cell division occurs. The C and D periods are labeled, and the middle numbers are how many *oriC*s are present. Data for the C and D periods can be found in Table 1.

before cell division occurs. Figure 1 shows the initiation and termination events for specific cell cycle times. The data used for this figure are taken from Bremer & Dennis (1987).

Gene location for *E. coli* is given in terms of minutes on the chromosome with *oriC* located at 84′ (Bachmann, 1990). Let x_g be the location of any specific gene in minutes around the chromosome from *oriC*. Because the *E. coli* chromosome consists of 100′, replication is bidirectional, and termination occurs approximately 50′ from *oriC*, it follows that $0 \le x_g \le 50$. Define $y = (C/50)x_g$, where C is the length of the C period, and define $t_g = C + D - y$ to be the time when the gene replicates relative to initiation of DNA replication, where D is the length of the D period. With this information

$$G_a(t) = \begin{cases} Q, & 0 \le t < t_0 \\ 2Q, & t_0 \le t \le \tau \end{cases}, \tag{3}$$

where Q and t_0 are defined below by the following:

$$\begin{aligned} Q=1 \quad & t_0 = y + \tau - (C+D), \quad \text{when } 0 < t_g \le \tau, \\ Q=2 \quad & t_0 = y + 2\tau - (C+D), \quad \text{when } \tau < t_g \le 2\tau, \\ Q=4 \quad & t_0 = y + 3\tau - (C+D), \quad \text{when } 2\tau < t_g \le 3\tau, \end{aligned} \tag{4}$$

Q is the number of genes in the newborn cell, and t_0 is the time in the cell cycle when the particular gene replicates. Thus, the rate of production in (1) is discontinuous which causes the concentration of RNA to fluctuate.

The solution to (1) with periodic boundary conditions (2) is found in Appendix A. From this analytical solution, one can readily find the time and value for the

minimum and maximum concentrations of RNA. These results are detailed in Appendix A. From this information we find that the ratio of the maximum to minimum concentration is given by

$$\frac{c_{\max}}{c_{\min}} = \frac{2}{e[\ln(2)]} \simeq 1{\cdot}061475, \tag{5}$$

which is constant independent of the generation time or gene position. Appendix A also gives an explicit formula for the average concentration of the constitutively produced stable RNA for an asynchronous culture of exponentially growing cells.

2.2. STUDY OF rRNA SYNTHESIS

In the previous section a model was developed for a constitutively produced stable RNA. The only stable RNA molecules in the cell are ribosomal RNA (rRNA) and transfer RNA (tRNA). Synthesis of rRNA under certain growth conditions can constitute over half of the RNA being transcribed which implies that the rRNA promoters are the strongest promoters in the cell. The metabolic cost of producing rRNA and the importance of rRNA in growing cells suggests that synthesis of rRNA is either the rate-limiting process in a growing cell or there is some control of the transcription rate.

Until recently, it was believed that the rRNA promoter was constitutive, but that its synthesis was controlled passively by competition for DNA-dependent RNA polymerase (Maaløe, 1979; Ingraham *et al.*, 1983). However, as noted before there is evidence that the RNA polymerase is present in the cell in excess. Another suggested control mechanism is a stringent response from the nucleotide ppGpp (Baracchini & Bremer, 1988; Sarubbi *et al.*, 1988). Jensen & Pedersen (1990) propose an alternative passive control where ppGpp inhibits the transcription of all genes and since stable RNA promoters are the strongest, they would be most affected by a decline in transcription rates. For this alternative passive control model we have the growth conditions establishing the level of ppGpp. The concentration of ppGpp would then determine the affinity of the promoters for the RNA polymerase, which would be reflected in the production constant r in the model given by (1). With r fixed based on growth conditions, the passively controlled rRNA synthesis satisfies our conditions for the constitutive model developed above.

We note that there are several other models for how rRNA synthesis is regulated. Gourse *et al.* (1985) propose a negative feedback mechanism responding to "free ribosomes" either directly or indirectly. Jensen & Pedersen (1990) claim that subsaturation or competition for RNA polymerase provides the regulatory control for rRNA synthesis. These models do not satisfy our hypotheses for constitutive production and would require changes in our mathematical model. However, we are assuming the passive control model to illustrate what information can be gleaned from the mathematical model using an actual biological example and do not claim that this is necessarily the best model based on current experimental findings.

In *E. coli*, there are seven operons which produce the stable rRNA molecules, and these are located at 87′ (*rrnA*), 90′ (*rrnB*), 85′ (*rrnC*), 72′ (*rrnD*), 90′ (*rrnE*), 57′

(*rrnG*), and 5′ (*rrnH*). As an example, consider the products of *rrnC* and *rrnG* located at 85′ and 57′ on the chromosome of *E. coli*. These operons produce stable rRNA molecules. For comparison, we also include a hypothetical gene located at 35′ producing a stable RNA product at the same rate located near the terminus (at 36′). These operons are located 1′, 27′, and 49′ from *oriC*, so (4) can be used to compute Q and t_0 given information on C, D, and τ. Figure 2 shows the variation in concentrations for these gene products for cell cycle times of 30 and 60 min, where data are given by Bremer & Dennis (1987) and Table 1. From the information on the ratio of c_{max}/c_{min} given in (5), each of the gene products individually vary about 6·15% between their minimum and maximum.

The graphs in Fig. 2 show that the gene products produced from genes close to *oriC* can contribute substantially more than the genes near the terminus. For each τ, a specific value of r is used so that the variations in concentration are a result of gene location only. From the values of the average concentrations of the gene located 1′ from *oriC* and the hypothetical gene located 1′ from the terminus we find that

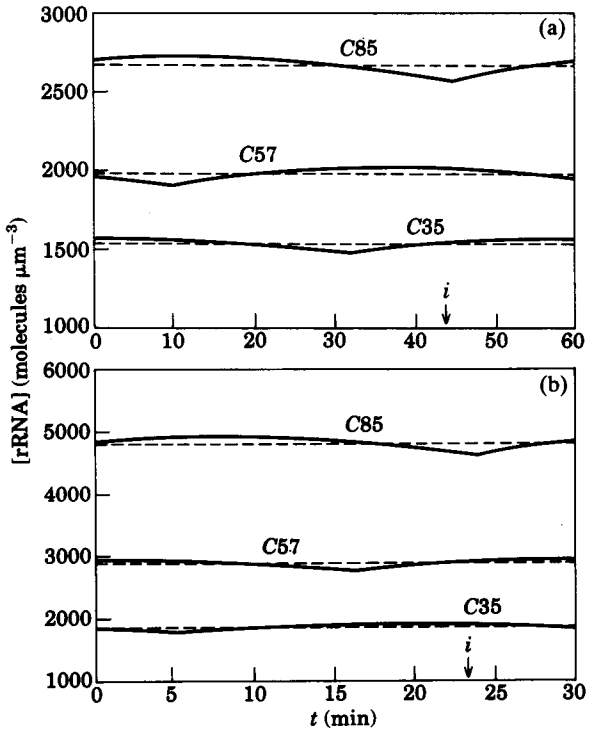

FIG. 2. These figures show the variation in concentration about the rRNA mean concentration (dashed line) for the *rrnC* (c_{85}) and *rrnG* (c_{57}) genes and a hypothetical gene (c_{35}) producing stable RNA at the same rate 1′ from the terminus; (a) is for a cell cycle time of 60 min, while (b) is for a cell cyle time of 30 min.

TABLE 1

Computations for products of the rrn *genes at different growth rates*

τ(min)	100	60	40	30	24
C period†	67	50	45	43	42
D period‡	30	27	25	24	23
V_0 §	0·437	0·577	0·815	1·15	1·63
$\bar{c}$ ‖	11·1	16·8	22·1	28·1	31·4
$\bar{c}_{85}$ ¶	1·73	2·68	3·65	4·81	5·57
$\bar{c}_{57}$ ††	1·36	1·98	2·43	2·87	2·96
$\bar{c}_{85}(\times 100)/\bar{c}$ ‡‡	15·6	15·9	16·5	17·1	17·7
$\bar{c}_{57}(\times 100)/\bar{c}$ §§	12·3	11·8	11·0	10·2	9·44
$\bar{c}_{85}/\bar{c}_{57}$	1·27	1·35	1·50	1·68	1·88
r ‖‖	3·75	10·27	22·56	38·42	56·98
c_{max} ¶¶	11·28	17·04	22·33	28·36	31·77
c_{max}/c_{min} †††	1·041	1·036	1·027	1·024	1·026

This table shows that as growth rate increases *rrn* genes located closer to *oriC* produce a higher percentage of the total rRNA. If promoter strength is approximately constant, then the *rrn* gene located at 57′ (farthest from *oriC*) actually produces less rRNA at higher growth rates. By having 7 *rrn* genes, the last line shows only a 2·4–4·1% variation in concentration of rRNA over the cell cycle.

† From Bremer & Dennis (1987).

‡ From Bremer & Dennis (1987).

§ Donachie & Robinson (1987) give the mean cell volume, V_μ, in μm^3, and from Appendix A, $V_0 = V_\mu/2\ln(2)$.

‖ × 1000 ribosomes μm^{-3} from Bremer & Dennis (1987).

¶ × 1000 ribosomes μm^{-3}, representing the average concentration of the product of the *rrn* gene located at 85′.

†† × 1000 ribosomes μm^{-3}, representing the average concentration of the product of the *rrn* gene located at 57′.

‡‡ Percentage of $\bar{c}$ produced by the *rrn* gene located at 85′.

§§ Percentage of $\bar{c}$ produced by the *rrn* gene located at 57′.

‖‖ Initiations min^{-1} gene^{-1} found from (6).

¶¶ × 1000 ribosomes μm^{-3}, representing the maximum concentration during the cell cycle.

††† Ratio of maximum to minimum concentration during the cell cycle.

when $\tau = 60$, the gene located 1′ from *oriC* produces about 74% more RNA than the gene located 1′ from the terminus. When $\tau = 30$, there is an increase of approximately 160% more RNA from the gene closer to *oriC*.

Table 1 shows a more detailed analysis of the *rrnC* and *rrnG* genes located at 85′ and 57′ (the *rrn* genes closest and farthest from *oriC*) for five different growth rates. Appendix A gives an explicit formula for the mean concentration of stable RNA, $\bar{c}$, in an asynchronous, exponentially growing culture of bacteria. We find the mean concentrations of $\bar{c}_{85}$ and $\bar{c}_{57}$ at each of the growth rates, assuming that the binding affinity of polymerase, and hence r, does not vary between positions. The following rows of the table use this information to show what percentage of the rRNA is produced at each of these gene locations for the different growth rates. The gene closest to *oriC* produces higher percentages of rRNA as the growth rate increases. From the table entry on the ratio of $\bar{c}_{85}/\bar{c}_{57}$, we see that the gene at 85′ produces 27% more rRNA than the gene at 57′ at the slowest growth rate, while it produces 88% more rRNA at the fastest growth rate. This shows that the location of a gene on the chromosome is important for changing growth rates.

Suppose the average rRNA concentration, $\bar{c}$, is known, and assume that all seven of the rRNA operons have the same initiation rate/gene, r, for a given growth rate. From Appendix A the individual gene products are shown to satisfy $\bar{c}_i = Q_i r\, e^{-kt_{0i}}/kV_0 \ln(2)$, $i = 1, \ldots, 7$, where Q_i and t_{0i} are found by (4). Note that $\bar{c} = \sum_{i=1}^{7} \bar{c}_i$. With this information, it follows that

$$r = \frac{\bar{c} k V_0 \ln(2)}{\sum_{i=1}^{7} Q_i e^{-kt_{0i}}}. \tag{6}$$

It is easy to combine the products of all seven of these genes and determine the total fluctuation in concentration as a function of cell doubling time.

A summary of the numerical results is presented in Table 1. A comparison of our computed r to the experimental values of Bremer & Dennis (1987) shows that most of the values are approximately 7% lower than the experimental values. This is expected as there is some turnover or degradation of the rRNA which is not considered in the model. With seven gene locations on the chromosome, Table 1 shows that the total variation of rRNA is 2·4–4·1% compared to the 6·15% variation of the individual genes given by (5). Thus, multiplicity of a specific gene positioned at several locations on the chromosome moderates the fluctuation in concentration over the cell cycle for the specific gene product. It is interesting to note that most of the rRNA genes, which are crucial for growth, are located closer to the origin of replication than the terminus.

3. Constitutive Protein with Decay

In this section, we examine a mathematical model for the production of a constitutive protein. The model in the previous section assumed that there was no degradation of the RNA; however, in the production of a protein, the mRNA usually has a half-life between 1 and 2 min. Once again, the mathematical model is an oversimplification of what is actually occurring in the cell, but this study does show the qualitative effects of gene doubling, molecular stability, gene position, and cell cycle time for protein synthesis with no regulatory control.

3.1. MATHEMATICAL MODEL FOR A CONSTITUTIVE PROTEIN

The mathematical model developed here is one of the simplest models that describes protein synthesis in a growing cell. This model considers degradation of the mRNA and protein, and it includes a time delay owing to transcription and translation. The model again assumes an exponentially growing cell with a volume, $V(t)$, as defined in the previous section. It is assumed that there are constant concentrations of DNA-dependent RNA polymerase and ribosomes throughout the cell cycle for a given growth rate. As noted in the Introduction, there is some controversy as to whether these are good modeling assumptions. However, these assumptions allow a detailed analysis of the mathematical model and, as a first approximation,

provide a valuable tool for assessing some qualitative features of constitutive protein production.

There are several kinetic parameters which must be defined to establish the differential equations for this model. G_a, the number of active gene sites, is defined as before. The parameter r_1 is the rate of initiation of successful transcripts from a particular gene site, and r_2 is the rate of initiation of successful translations. These rate constants are for initiations which proceed to completion. The parameters τ and k, which are the cell cycle time and dilution rate, respectively, are the same as the ones in the previous model. It is assumed that degradation is proportional to the concentration of the mRNA and protein with rate constants k_1 and k_2, respectively. The model also includes a time delay, T, for the time between the initiation of an mRNA and the completion of a protein (a composite of the transcription and translation times that is assumed to be constant). In bacteria, transcription and translation are coupled in that ribosomes begin translation almost immediately after the initiation of transcription, so the delay T is only slightly longer than the time required to complete translation of the mRNA. Thus, it is appropriate to define $c_1(t)$ to be the concentration of the mRNA along with all its initiated chains which will proceed to completion. The concentration of the completed protein is denoted by $c_2(t)$. The differential equations governing these quantities are given by the following system of equations:

$$\dot{c}_1(t) = \frac{G_a(t) r_1}{V(t)} - (k_1 + k) c_1(t), \qquad 0 \le t \le \tau \tag{7}$$

$$\dot{c}_2(t) = r_2 c_1(t - T) - (k_2 + k) c_2(t), \qquad 0 \le t - T \le \tau. \tag{8}$$

As shown in Appendix B, these differential equations are solved, and explicit formulae for the solutions are obtained. This Appendix also shows that the unique solution with periodic boundary conditions $c_1(0) = c_1(\tau)$ and $c_2(T) = c_2(\tau + T)$ is a globally attracting solution. The time shift in (8) is necessary to account for the delay in producing a completed protein. To have the solutions both correspond to one cell cycle, one can relate the solution of (8) on $[\tau, \tau + T]$ to the solution of (7) on $[0, T]$.

3.2. ESTIMATING RNA HALF-LIFE FROM THE MODEL

From the analytical solution to (7), the time and value for the minimum and maximum concentrations of mRNA, c_1, can be determined. Details of these calculations are found in Appendix B. [See eqns (B.4) and (B.5).] The ratio of the maximum concentration of mRNA, $c_1(t_{max})$, to the minimum concentration of mRNA, $c_1(t_{min})$, is given by

$$\frac{c_1(t_{max})}{c_1(t_{min})} = \frac{k_1(2 - e^{-k_1\tau})}{(k_1 + k)(1 - e^{-k_1\tau})} \left[\frac{k(2 - e^{-k_1\tau})}{k_1 + k} \right]^{k/k_1}. \tag{9}$$

If the ratio between the minimum and maximum concentrations of the constitutive mRNA in a cell growing with a certain cell cycle time is known, then, in principle, (9) provides a means to approximate the value of the decay rate, k_1. This could

be achieved experimentally with a well-synchronized culture of cells. However, the experimental measurements would only obtain a lower bound on k_1 (upper bound on the mRNA half-life), as variations in the cell cycle for a culture of cells would cause experimental measurements of the ratio in (9) to be underestimated. Using techniques from calculus, we find that (9) agrees with the ratio of the maximum to minimum concentration of stable RNA given by (5) as $k_1 \to 0$. The ratio in (9) tends to two as $k_1 \to \infty$, which is the limiting case of a very unstable RNA molecule and agrees with changes in concentration of the gene at $t = t_0$.

As an example of how in theory (9) could be used to determine a half-life, suppose that a synchronous culture of cells grows with a cell cycle time of 45 min. Assume that an experiment shows that the ratio between the maximum and minimum concentration of some mRNA over the cell cycle is $1{\cdot}7 \pm 0{\cdot}05$. Using a bisection method (see Dahlquist, & Björck, 1974: 220) on (9), we find that $k_1 = 0{\cdot}318\ \text{min}^{-1}$. This translates into a half-life of $t_{1/2} = 2{\cdot}18$ min. With the experimental error listed, the range for k_1 is $0{\cdot}250$–$0{\cdot}418\ \text{min}^{-1}$. This yields a range for the half-life of the mRNA between 1·66 and 2·77 min. This shows that a small error in the ratio between maximum and minimum concentrations can result in a large error in the half-life. As noted above, these results only provide an upper bound for the half-life of the mRNA.

The minimum concentration of the protein, c_2, occurs at $t_0 + T$. However, the equation for the concentration of the protein is too complicated for an explicit formula giving the maximum of c_2. Thus, numerical computations are used to find the ratio of the maximum to minimum concentration of the protein in our studies below.

For comparison to experimental data, the average concentrations of mRNA, $\bar{c}_1$, and protein, $\bar{c}_2$, are computed for an asynchronous culture of exponentially growing cells using the solutions to the differential equations (7) and (8). Information on how these formulae are derived can be found in Appendix B. As we saw in the previous section, information on the mean concentration can be used for computing quantities such as the initiation rates, r_1 and r_2.

In the previous section we noted that the initiation rates for rRNA of Bremer & Dennis (1987) were higher than the theoretical values presented in Table 1. If one assumes that the higher initiation rate compensates for degradation, then a comparison of initiation rate for a stable RNA [see (6)] to the initiation rate for an RNA including decay [see (B.6)] can be used to compute the half-life of the rRNA. As an example, Bremer & Dennis (1987) give r_1 to be 11 and 23 for generation times of 60 and 40 min, respectively. In the case where $\tau = 60$, r_1 is 7·1% higher than r, so $k + k_1 = 1{\cdot}071k$ or $k_1 = 0{\cdot}071k$. This gives a half-life for the rRNA, as $t_{1/2} = \tau/0{\cdot}071 \simeq 845$ min $\simeq 14{\cdot}1$ hr. When $\tau = 40$, a similar calculation yields $k_1 = 0{\cdot}02k$. Thus, $t_{1/2} = \tau/0{\cdot}02 \simeq 2000$ min $\simeq 33{\cdot}3$ hr, which corresponds well with turnover rates for stable proteins (Mandelstam, 1957; Goldberg & St John, 1976). Gausing (1977) showed that describing the stability of rRNA is a bit more complicated because its half-life matches that of mRNA with rapid degradation when it is not associated with ribosomal proteins, but appears to have the same stability as the ribosomal proteins, which are stable (Dennis, 1974), when combined to make a ribosome. He

conjectures that rRNA is produced in excess with high turnover, so that the cell can readily adapt to various growth conditions and that the ribosomal protein production is growth-rate regulated and determines the concentration of the mature ribosomes.

3.3. STUDY OF INDUCED *lac* OPERON

Most genes have some type of controls and, hence, fail to satisfy our modeling assumption of a "constitutive" gene. This makes it difficult to find representative experiments with which to compare our model. The experiments of Kennell & Riezman (1977) on the *lac* operon used sufficient quantities of isopropyl-β-D-thiogalactoside (IPTG) to fully induce the gene, and cAMP was added to minimize the effects of catabolite repression. This experiment makes the *lac* gene for β-galactosidase effectively constitutive in terms of our model.

Their experiment used cells which had a generation time of 48 min or $k = 0{\cdot}0144\ \text{min}^{-1}$. Linear interpolation from Table 1 yields $V_0 = 0{\cdot}72\ \mu\text{m}^3$, $C = 47$, and $D = 26$ min. As the β-galactosidase gene is located at 8′, or 24′ from *oriC*, $t_0 = 45{\cdot}6$ min. They found that the half-life of the mRNA transcript was 1·5 min (or $k_1 = 0{\cdot}462\ \text{min}^{-1}$) and that the average number of *lac* mRNA molecules cell^{-1} was 31 or $\bar{c}_1 = 30{\cdot}5$ molecules μm^{-3} as the average cell volume is $1{\cdot}015\ \mu\text{m}^3$. It follows that $r_1 = 7{\cdot}11$ initiations gene^{-1}min^{-1} [see (B.6)], assuming that this highly active constitutive promoter has not exhausted the polymerase available between the lowest and highest output of mRNA. They report 3·3 sec between initiations which, when combined with the 1·7 genes cell^{-1}, yields 11 initiations gene^{-1} min^{-1}. Thus, about 35% of the initiations must not proceed to completion. Numerical studies show that the ratio of the maximum to minimum concentration of mRNA is 1·78 with minimum and maximum concentrations of 22·2 and 39·3 molecules μm^{-3}, respectively.

The experiments of Kennell & Riezman (1977) on the *lac* operon found that 20 monomers of β-galactosidase sec^{-1}cell^{-1} were produced after full induction with IPTG. With the average of 31 mRNA molecules cell^{-1}, this implies that $r_2 = 38{\cdot}7$ initiations min^{-1}mRNA^{-1}. (This contrasts with their claim of 3·2 sec between initiations which would yield 19 initiations min^{-1}mRNA^{-1}.) β-Galactosidase is a stable protein and has a turnover rate of about 1–4% hr^{-1} (Mandelstam, 1957; Goldberg & St John, 1976). If a turnover rate of 2% hr^{-1} is used, then $k_1 = 3{\cdot}37 \times 10^{-4}\ \text{min}^{-1}$ and the half-life of β-galactosidase is 34·3 hr. It follows from Appendix B that $\bar{c}_2 = 84\,600$ molecules μm^{-3}. This value agrees reasonably well with the concentration given by Beckwith (1987) who reports 60 000 molecules of β-galactosidase per cell. Simulations of (8) show that the ratio of the maximum to minimum concentration is only 1·07 which indicates little change throughout the cell cycle of the concentration of β-galactosidase.

3.4. THEORETICAL STUDY FOR CONSTITUTIVE PROTEIN

A theoretical study is performed to determine the amount of variation in concentration of c_1 and c_2 when the cell cycle time, mRNA half-life, gene position, and protein stability were varied. When $\tau = 60$, the kinetic parameter, r_1, is taken to be

1 min^{-1}. This value corresponds to about 20% of the activity of the fully induced *lac* operon. The data of Bremer & Dennis (1987) give the mean number of RNA polymerase molecules $cell^{-1}$ to be 2800 and 8000 for generation times of 60 and 30 min, respectively. Assuming that r_1 is predominately determined by the concentration of RNA polymerase and using the volume ratios from Table 1 for generation times of 60 and 30 min, we compute the value of r_1 for a 30 min generation time to be 1·43 min^{-1}. The modeling assumption used here is that the concentration of free RNA polymerase is proportional to the total concentration of RNA polymerase. This assumption may affect some of the quantitative results between different generation times. However, it will not affect the qualitative results nor will it affect quantitative comparisons for a given generation time.

For $\tau = 60$, the kinetic parameter, r_2, is given by 30 min^{-1}. The r_2 value agrees reasonably well with the ribosome activity used above for the production of β-galactosidase. The data of Bremer & Dennis (1987) give the mean number of ribosomes $cell^{-1}$ as 13 400 and 45 000 for generation times of 60 and 30 min, respectively. If the production rate, r_2, for β-galactosidase depends primarily on the concentration of the ribosomes, then a similar calculation to the one for r_1 yields $r_2 =$ 50·4 min^{-1} for a generation time of 30 min. The model assumes that the free ribosome concentration is proportional to the total concentration of ribosomes. Again any weakness in this assumption should not affect the qualitative features of the model and should only create difficulties in comparing data from different generation times. The time delay, T, is taken to be 0·5 min in all cases.

The results of the simulations are presented in Tables 2 and 3. The third column in each table shows how the ratio of the maximum to minimum concentration of mRNA varies for two cell cycle times and two different half-lifes for the mRNA. As the stability of the mRNA increases, this ratio decreases. As the generation time decreases, so does this ratio. The gene position has no effect on the percent change

TABLE 2.

Variations in concentrations of mRNA, c_1 *and stable and unstable proteins,* c_2^s *and* c_2^u, *respectively, when the generation time,* $\tau = 60$

$\frac{\ln(2)}{k_1}$	Position	$\frac{c_1(t_{max})}{c_1(t_{min})}$	$\bar{c}_1$	$\frac{c_2^s(t_{max})}{c_2^s(t_{min})}$	$\bar{c}_2^s(\times 10^3)$†	$\frac{c_2^u(t_{max})}{c_2^u(t_{min})}$	$\bar{c}_2^u(\times 10^3)$‡
1	1′	1·86	4·27	1·059	10·8	1·18	3·69
	25′	1·86	3·24	1·059	8·16	1·18	2·80
	49′	1·86	2·45	1·059	6·19	1·18	2·12
2	1′	1·77	8·40	1·056	21·2	1·17	7·27
	25′	1·77	6·37	1·056	16·1	1·17	5·51
	49′	1·77	4·82	1·056	12·2	1·17	4·18

The ratio of maximum to minimum concentrations over the cell cycle and mean concentrations of mRNA and protein are presented for three different gene positions and two different half-lifes ($\ln(2)/k_1$) for the mRNA. The ratios show how molecular stability affects variation in concentration over the cell cycle with increased stability decreasing the ratio of variation. The mean concentrations show how gene position affects its expression with genes located closer to *oriC* having the highest expression.

† The stable protein is lost at a rate of 2% hr^{-1} (or a half-life of 34·3 hr).

‡ The unstable protein has a half-life of 30 min.

TABLE 3

Variations in concentrations of mRNA, c_1 and stable and unstable proteins, c_2^s and c_2^u, respectively, when the generation time, $\tau = 30$

$\frac{\ln(2)}{k_1}$	Position	$\frac{c_1(t_{max})}{c_1(t_{min})}$	$\bar{c}_1$	$\frac{c_2^s(t_{max})}{c_2^s(t_{min})}$	$\bar{c}_2^s(\times 10^3)$†	$\frac{c_2^u(t_{max})}{c_2^u(t_{min})}$	$\bar{c}_2^u(\times 10^3)$‡
1	1′	1·77	5·77	1·055	12·4	1·11	6·30
	25′	1·77	3·58	1·055	7·70	1·11	3·91
	49′	1·77	2·22	1·055	4·78	1·11	2·43
2	1′	1·63	11·2	1·047	24·0	1·095	12·2
	25′	1·63	6·94	1·047	14·9	1·095	7·57
	49′	1·63	4·31	1·047	9·27	1·095	4·70

This table compares to Table 2 to show how a decrease in generation time affects variation in concentration and gene expression. Note the amount of variation in concentration decreases; however, the effects of gene position are amplified with genes closer to *oriC* producing more than the ones further away, proportionately.

† The stable protein is lost at a rate of 2% hr^{-1} (or a half-life of 34·3 hr).

‡ The unstable protein has a half-life of 30 min.

between the minimum and maximum mRNA or protein concentration, though it does affect the average concentration with the average increasing for genes closer to *oriC*.

In both tables, the ratio of the maximum to minimum concentration of the stable protein is slightly less than 1·06 which is close to the value computed for a stable RNA. The value is less than the one computed for a stable RNA as the variation in the mRNA concentration is not as dramatic as the change in the gene concentration

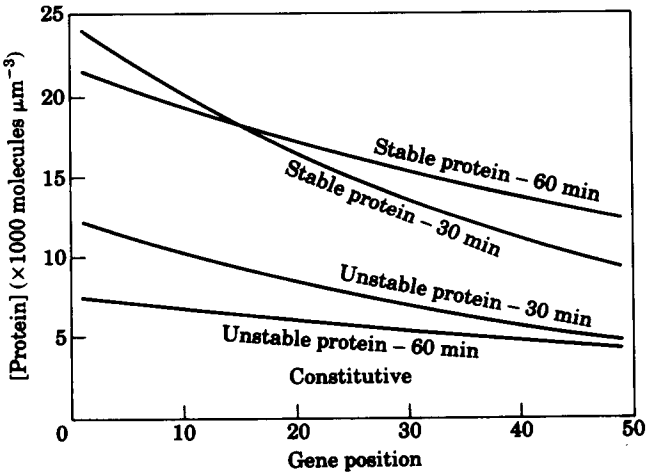

FIG. 3. This figure shows the mean concentration of stable or unstable protein as a function of gene position relative to *oriC* for a constitutive gene at generation times of 30 and 60 min. The kinetic parameters agree with the ones used to produce Tables 2 and 3 with the half-life of mRNA taken to be 2 min.

which is a step function (two-fold concentration change at t_0). The remaining ratios decrease as the growth rate increases. The computed ratios of the maximum to minimum concentrations of both mRNA and the proteins show that the less stable cases have greater variations in their concentrations.

Figure 3 shows how the mean concentration of stable and unstable protein varies with gene position for two different generation times. The graph illustrates that the protein concentration varies more with gene position at higher growth rates with a higher percentage of gene product from a gene closer to the origin of replication. At the slower growth rate all ratios of $\bar{c}_1$, $\bar{c}_2^s$, or $\bar{c}_2^u$ at the 1′ position to the 49′ position is the constant 1·74 or a 74% increase in production owing to the gene closer to the origin of replication. At the faster growth rate these ratios increase to almost 2·6 or a 160% increase in production owing to the gene closer to the origin of replication.

The simulations show that the mean concentration of stable protein, $\bar{c}_2$, increases about 14% for a gene located 1′ from *oriC* as the generation time is halved from 60 min to 30 min, while it decreases by about 23% for a gene located 49′ from *oriC*. However, this quantitative comparison between generation times is not likely to be robust as it depends heavily on our modeling assumptions for RNA polymerase and ribosomes, but the qualitative effect of the gene closer to *oriC* having a more significant role at higher growth rates is maintained.

4. Controlled Protein Production: Repression and Activation

Many authors have examined mathematical models of protein synthesis which are controlled by repression (for example, Goodwin, 1965; Othmer, 1976; Banks & Mahaffy, 1978). These models can exhibit oscillatory solutions about a steady-state solution. Biologists consider this negative feedback system important for maintaining a specific intracellular concentration of many proteins. However, because the gene doubles at some point in the cell cycle, there is not a steady-state solution. Rather, there is an intrinsic oscillation during the cell cycle owing to gene doubling. As seen in the earlier mathematical models which fail to include gene doubling, certain parameter values may produce rapid oscillations in the concentrations, distinct from the oscillation caused by the cell cycle.

At the other end of the spectrum is positive feedback or activation of a gene. Mathematical studies have focused on models of induction as the classical study for positive feedback. In these models it can be shown that there are no oscillatory solutions; however, multiple equilibria may exist (Othmer, 1976). In this section we examine mathematical models of protein synthesis controlled by the simplest negative and positive feedback systems, autorepression and autoactivation, and include the effects of gene doubling. Our studies have elucidated the response of negatively and positively controlled genes to variations in binding strengths and growth-related parameters. In *E. coli*, most genes use several types of control on production of their proteins; however, autorepression and autoactivation as presented here provide a basis for understanding the more complicated cellular controls.

4.1. MODELING AND SIMULATION OF AUTOREPRESSION

The mathematical model for autorepression with gene doubling is developed using basic biochemical kinetics and making similar simplifying assumptions as the constitutive model for the action of RNA polymerase and ribosomes. It is assumed that the protein endproduct which is produced from a particular gene can bind to the operator region of the gene and prevent transcription. Details of the biochemical kinetics used to formulate the production term for the mRNA of the protein can be found in Appendix C. The remaining production and degradation terms in the mathematical model are the same as the ones used in the model for a constitutive protein. The resulting mathematical model for an autorepressed protein is given by the following system of differential equations:

$$\begin{aligned} \dot{c}_1(t) &= \frac{r_1 K_r G_a(t)}{V(t)(K_r + c_2(t))} - (k_1 + k)c_1(t) \\ \dot{c}_2(t) &= r_2 c_1(t - T) - (k_2 + k)c_2(t), \end{aligned} \tag{10}$$

where $c_1(t)$ is the concentration of the mRNA and $c_2(t)$ is the concentration of the protein. The functions G_a and V are as before, as are the constants k, k_1, k_2, r_1, and r_2. K_r is the binding constant for the repressor to the operator region. This nonlinear system of differential equations cannot be solved explicitly as was done in the previous section, so is solved numerically. The results of the numerical simulations show that there is a unique globally attracting solution satisfying the periodic boundary conditions $c_1(0) = c_1(\tau)$ and $c_2(0) = c_2(\tau)$. Unlike the constitutive case, this mathematical result has not been proved.

Tables 4 and 5 summarize results from numerical simulations of the repression model for two generation times, 60 and 30 min. When $\tau = 60$, the kinetic parameters, r_1, r_2, and T, were taken to be 1 min^{-1}, 30 min^{-1} and 0·5 min, which are the same as the ones used to generate Table 2. When $\tau = 30$, r_1 and r_2 are changed to 1·43 min^{-1} and 50·4 min^{-1}, respectively, to reflect the increased concentration of the RNA polymerase and the ribosomes at the higher growth rate and agree with the values used to generate Table 3. As noted before, our modeling assumptions probably affect the quantitative results to some degree, but not the qualitative features. The stable protein is assumed to lose 2% hr^{-1} (or a half-life of 34·3 hr), while the unstable protein has a half-life of 30 min. Tables 4 and 5 show the effects of varying the strength of the binding of the repressor on a gene located 25′ from *oriC*. The strongest binding constant considered is $K_r = 10$, which corresponds well to the value given by Ptashne (1986) for the binding of the *cro* repressor. Since there is a negative feedback in the differential equations, the mRNA concentration is affected by the protein concentration; hence, the inclusion of columns for $\bar{c}_1^s$ and $\bar{c}_1^u$ which show the effects on the mean concentrations of the mRNA depending on whether the resulting protein is stable or unstable.

The result which most stands out from these tables is the constant ratios between the maximum and minimum concentrations of the mRNA or the protein at different repressor binding affinities. This shows that the strength of the repressor has little effect on how much the concentration varies through the cell cycle in contrast to the

TABLE 4

Effects of repression on the variations in concentrations of mRNA, $\bar{c}_1$, and stable and unstable proteins, c_2^s and c_2^u, respectively, when the generation time, $\tau = 60$

$\frac{\ln(2)}{k_1}$	K_r	$\frac{c_1(t_{max})}{c_1(t_{min})}$	$\bar{c}_1^s$ †	$\bar{c}_1^u$ ‡	$\frac{c_2^s(t_{max})}{c_2^s(t_{min})}$	$\bar{c}_2^s(\times 10^3)$	$\frac{c_2^u(t_{max})}{c_2^u(t_{min})}$	$\bar{c}_2^u(\times 10^3)$
1	10	1·86	0·081	0·187	1·060	0·277	1·17	0·162
	100	1·86	0·339	0·556	1·060	0·855	1·17	0·481
	1000	1·86	0·951	1·44	1·060	2·40	1·17	1·25
2	10	1·76	0·157	0·265	1·057	0·396	1·16	0·230
	100	1·76	0·483	0·802	1·057	1·22	1·16	0·694
	1000	1·76	1·40	2·19	1·057	3·54	1·16	1·90

The gene position is assumed to be 25′ from *oriC*, while other parameters were chosen to match those taken in Table 2. The ratio of maximum to minimum concentration over the cell cycle and mean concentrations of mRNA and protein are presented for three different binding strengths of the repressor, K_r, and for two different half-lifes for the mRNA. The ratios of variation decrease with increasing molecular stability, yet remain constant as K_r varies. The mean concentrations decrease with increasing repressor binding strength (decreasing K_r).

† This mean concentration of mRNA is associated with the stable (*s*) protein.

‡ This mean concentration of mRNA is associated with the unstable (*u*) protein.

TABLE 5

Effects of repression on the variations in concentrations of mRNA, c_1, and stable and unstable proteins, c_2^s and c_2^u, respectively, when the generation time, $\tau = 30$

$\frac{\ln(2)}{k_1}$	K_r	$\frac{c_1(t_{max})}{c_1(t_{min})}$	$\bar{c}_1^s$ †	$\bar{c}_1^u$ ‡	$\frac{c_2^s(t_{max})}{c_2^s(t_{min})}$	$\bar{c}_2^s(\times 10^3)$	$\frac{c_2^u(t_{max})}{c_2^u(t_{min})}$	$\bar{c}_2^u(\times 10^3)$
1	10	1·77	0·127	0·177	1·057	0·273	1·11	0·193
	100	1·77	0·386	0·530	1·057	0·830	1·11	0·578
	1000	1·77	1·08	1·38	1·057	2·32	1·11	1·60
2	10	1·65	0·178	0·248	1·050	0·382	1·10	0·281
	100	1·65	0·546	0·755	1·050	1·17	1·10	0·823
	1000	1·65	1·58	2·11	1·050	3·40	1·10	2·30

As in Table 4, the gene position is assumed to be 25′ from *oriC*. This table compares to Table 4 to show how a decrease in generation time affects the variation in concentration and mean concentration of mRNAs and proteins. The variation in concentration decreases as with the constitutive case. The change in mean concentration between growth rates varies less as the repressor binding strength increases.

† This mean concentration of mRNA is associated with the stable (*s*) protein.

‡ This mean concentration of mRNA is associated with the unstable (*u*) protein.

standing belief in biological literature that autorepression is an effective means of maintaining relatively constant concentrations of the protein. When these tables are compared against the results for the constitutive case in Tables 2 and 3, then it can be seen that there is a decrease in the ratios between the maximum and minimum concentrations for the proteins, but this effect is negligible. It appears that the principal reason why autorepression does not affect the variation in concentration over the cell cycle is because the amount of protein being replaced at any one time in the cell cycle is small compared to the total protein concentration. This is particularly true when the protein is stable. Thus, the variation in concentration reflects the gene

dosage change that is similar for both the constitutive model and the one which includes autorepression.

For an autorepressed gene, the mean concentrations of the resultant mRNAs and proteins decrease with increasing binding strengths of the repressor (decreasing K_r values). If a ratio of mean concentrations, c_2^s, is formed between two growth rates, $\tau = 30$ and 60, then a comparison of this ratio for the autorepressed model to the one for the constitutive model shows greater variation in the latter case. Specifically, when the half-life of the mRNA is 1 min, then there is a 6·0% change between growth rates for the constitutive gene; while the autorepressed gene with $K_r = 10$ demonstrates only a 1·5% change. Hence, autorepression moderates changes in mean concentration at different growth rates. This supports the Sompayrac & Maaløe (1973) model of autorepression for control of DNA replication, where they hypothesize that the cell maintains a relatively constant concentration of some important activator protein over a range of growth conditions. However, the variation in concentration observed in the mathematical models between different growth rates is close to the naturally occurring fluctuation between the maximum and minimum concentrations within a cell cycle, suggesting that the control by autorepression is not necessary for the Sompayrac & Maaløe model.

Figure 4 exhibits the variation in mean concentration of stable and unstable protein with respect to gene position for two different generation times. As in the case of the constitutive gene, the graph shows that the protein concentration varies more with gene position at higher growth rates. However, this graph demonstrates that the gene position does not affect protein production as dramatically for an autorepressed gene as it does for a constitutive one. The graphs show that for some gene positions the

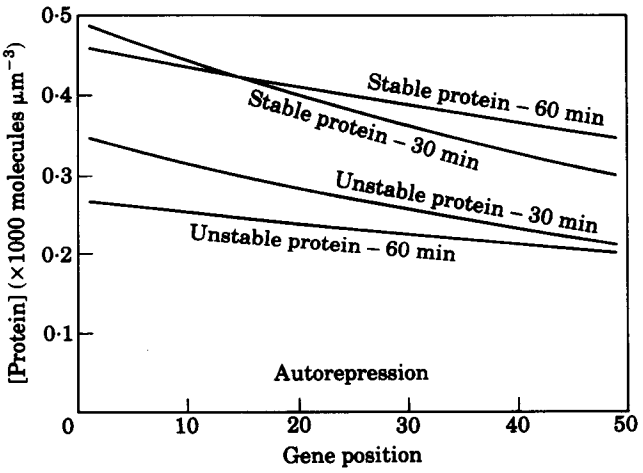

FIG. 4. This figure shows the mean concentration of stable or unstable protein as a function of gene position relative to *oriC* for an autorepressed gene at generation times of 30 and 60 min. The kinetic parameters agree with the ones used to produce Tables 4 and 5 with the half-life of mRNA taken to be 2 min and the repressor binding strength $K_r = 10$.

concentration of the stable protein is higher at the slower growth rate. This result is an artifact of how the rate constants r_1 and r_2 are chosen so the quantitative comparison between the generation times for a given gene position is not appropriate though the overall qualitative comparisons are valid.

4.2. MODELING AND SIMULATION OF AUTOACTIVATION

In Appendix C, the biochemical kinetics necessary for a mathematical model of autoactivation is developed. For this model it is assumed that the protein endproduct which is produced from a particular gene is needed to stimulate its own production. As noted in Appendix C, an approximation in the biochemical kinetics used to derive the production term results in a zero equilibrium where an actual cell would show low basal levels of the protein. The remaining production and degradation terms and the simplifying assumptions are the same as the ones used in the mathematical model for a constitutive protein. The resulting model for autoactivation is given by the following system of differential equations:

$$\dot{c}_1(t) = \frac{r_1 G_a(t) c_2(t)}{V(t)(K_a + c_2(t))} - (k_1 + k)c_1(t),$$
$$\dot{c}_2(t) = r_2 c_1(t - T) - (k_2 + k)c_2(t), \tag{11}$$

where $c_1(t)$ is the concentration of the mRNA and $c_2(t)$ is the concentration of the protein. The functions G_a and V are as before, as are the constants k, k_1, k_2, r_1, and r_2. K_a is the binding constant for the activator to the operator region. Like (10), this is a non-linear system of differential equations which must be solved numerically. Again the numerical studies show that there is a unique, globally attracting periodic

TABLE 6.

Effects of autoactivation on the variations in concentrations of mRNA, c_1, *and stable and unstable proteins,* c_2^s *and* c_2^u, *respectively, when the generation time,* $\tau = 60$

$\frac{\ln(2)}{k_1}$	K_a	$\frac{c_1(t_{max})}{c_1(t_{min})}$	$\bar{c}_1^s$ †	$\bar{c}_1''$ ‡	$\frac{c_2^s(t_{max})}{c_2^s(t_{min})}$	$\bar{c}_2^s(\times 10^3)$	$\frac{c_2^u(t_{max})}{c_2^u(t_{min})}$	$\bar{c}_2''(\times 10^3)$
1	100	1·86	3·19	3·12	1·059	8·06	1·17	2·70
	1000	1·86	2·19	2·07	1·059	7·16	1·17	1·80
	5000	1·86	0·964	0	1·059	3·12	—	0
2	100	1·77	6·33	6·25	1·055	16·0	1·17	5·41
2	1000	1·77	5·97	5·21	1·055	15·1	1·16	4·51
	5000	1·77	4·58	0·574	1·055	11·1	1·16	0·497

The gene position is assumed to be 25′ from *oriC*, while other parameters were chosen to match those taken in Table 2. The ratio of the maximum to minimum concentration over the cell cycle and mean concentrations of mRNA and protein are given for three different binding strengths of the activator, K_a, and two different half-lifes of the mRNA. The ratios of variation decrease with increasing molecular stability, yet remain constant as K_a varies. The mean concentrations increase with increasing activator binding strength (decreasing K_a). One entry shows that the gene may be turned off.

† This mean concentration of mRNA is associated with the stable (*s*) protein.

‡ This mean concentration of mRNA is associated with the unstable (*u*) protein.

TABLE 7

Effects of autoactivation on the variations in concentrations of mRNA, c_1, and stable and unstable proteins, c_2^s and c_2^u, respectively, when the generation time, $\tau = 30$

$\frac{\ln(2)}{k_1}$	K_a	$\frac{c_1(t_{max})}{c_1(t_{min})}$	$\bar{c}_1^s$ †	$\bar{c}_1^u$ ‡	$\frac{c_2^s(t_{max})}{c_2^s(t_{min})}$	$\bar{c}_2^s(\times 10^3)$	$\frac{c_2^u(t_{max})}{c_2^u(t_{min})}$	$\bar{c}_2^u(\times 10^3)$
1	100	1·77	3·54	3·50	1·055	7·61	1·11	3·81
	1000	1·76	3·08	2·67	1·055	6·71	1·11	2·91
	5000	1·76	1·26	0	1·053	2·70	—	0
2	100	1·63	6·90	6·86	1·047	14·8	1·095	7·48
	1000	1·63	6·48	6·03	1·047	13·9	1·094	6·57
	5000	1·63	4·62	2·46	1·047	9·93	1·089	2·65

As in Table 6, the gene position is assumed to be 25′ from *oriC*. This table compares to Table 6 to show how a decrease in generation time affects change in the model for autoactivation. The variation in concentration decreases as in the previous two examples. The change in mean concentration between growth rates shows a more pronounced effect than either the constitutive or repressed models.

†This mean concentration of mRNA is associated with the stable (s) protein.

‡This mean concentration of mRNA is associated with the unstable (u) protein.

solution satisfying the periodic boundary conditions $c_1(0) = c_1(\tau)$ and $c_2(0) = c_2(\tau)$.

Tables 6 and 7 show that once again the variation in concentration of the protein over the cell cycle is little affected by autoactivation when compared to the two previous models. As before, the mean concentration of the stable protein at $\tau =$ 60 min is higher than the concentration when $\tau = 30$ min, and this change is more pronounced than it was in the constitutive or autorepressed models. A striking difference from the previous models appears when the protein and the mRNA are

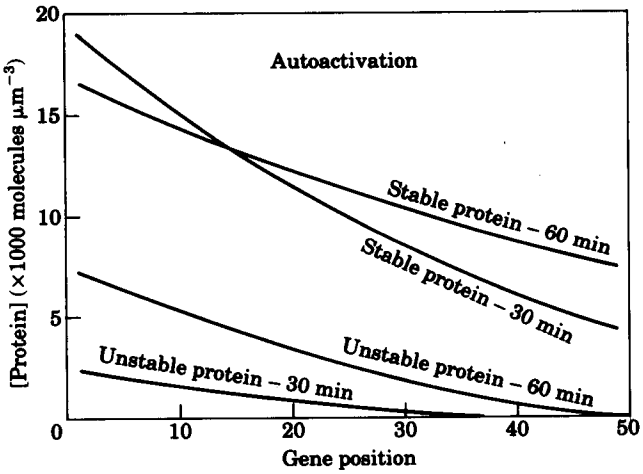

FIG. 5. This figure shows the mean concentration of stable or unstable protein as a function of gene position relative to *oriC* for an autoactivated gene at generation times of 30 and 60 min. The kinetic parameters agree with the ones used to produce Tables 6 and 7, with the half-life of mRNA taken to be 2 min and the activator binding strength $K_a = 5000$.

each unstable and $K_a = 5000$. In this case the zero equilibrium solution of the model becomes stable. This indicates that it is difficult for the gene to become activated (weak binding affinity), so the decay processes dominate driving the solution to zero. When $K_a = 100$, the affinity of the activator is quite strong and the resulting production closely matches the production rate in the constitutive case. As the binding affinity for the activator decreases, the corresponding production decreases as one would predict.

Figure 5 shows the variation in mean concentration of stable and unstable protein with respect to gene position corresponding to two different generation times. Once again, the graph demonstrates that the protein concentration varies more with gene position at higher growth rates, but with autoactivation this effect is much more significant. As seen in the case of an unstable protein with a generation time of 30 min, the gene can actually "shut off" for a gene position greater than 38′ from *oriC*.

4.3. SIMULATION OF CHANGING GROWTH CONDITIONS

A series of computer simulations are performed to determine the effects of a shift up or shift down by 10% in the rate constants for the cell. The main purpose of this study is to show how a growing cell might respond to a nutritional change depending on the type of control and position of the gene. Simulations include examples of neutral, negative, and positive control by considering a constitutive gene, an autorepressed gene, and an autoactivated gene, respectively. The simulations examine the mean concentrations of a stable and an unstable protein in three different gene positions for two different growth rates. The computer simulations changed only the rate constants r_1 and r_2 which govern the rate of mRNA and protein production, and would correspond biologically to a shift of 10% in the concentration or binding

TABLE 8

A shift up of 10% *is applied to the rate constants for producing mRNA and protein for each of the constitutive, autorepressed, and autoactivated genes previously studied*

Gene type	Position	$\bar{c}_1^s$	$\bar{c}_2^s(\times 10^3)$	$\bar{c}_1^u$	$\bar{c}_2^u(\times 10^3)$
Constitutive	1′	9·24	25·6	9·24	8·80
	25′	7·00	19·4	7·00	6·67
	49′	5·31	14·7	5·31	5·05
Repression with $K_r = 10$	1′	0·181	0·502	0·306	0·292
	25′	0·157	0·436	0·266	0·253
	49′	0·137	0·379	0·231	0·220
Activation with $K_a = 5000$	1′	7·43	20·6	3·97	3·78
	25′	5·19	14·4	1·73	1·65
	49′	3·51	9·73	0	0

The effects on the mean mRNA and protein concentrations are shown for various gene locations and protein stabilities. In this table the generation time is $\tau = 60$, and the mRNA half-life is 2 min. Production of mRNA and protein from a gene proximal to *oriC* is higher than a distal one with autorepression moderating the change in concentration and autoactivation showing a dramatic change in mean concentration. The constitutive gene is intermediate, as expected.

activity of RNA polymerase and ribosomes without a change in the rate of volume growth of the cell. RNA polymerase and ribosome concentration are seen to increase with increasing growth rates (Bremer & Dennis, 1987). Clearly, the actual biological response to a nutrient change consists of very complicated and interconnected controls, which would be very difficult to model. One explanation for changing the rate constants only in the model is that initially an enrichment of growth medium should cause the cell to increase production of crucial cellular components such as ribosomes

TABLE 9

A shift down of 10% *is applied to the rate constants for producing mRNA and protein for each of the constitutive, autorepressed, and autoactivated genes previously studied*

Gene type	Position	$\bar{c}_1^s$	$\bar{c}_2^s(\times 10^3)$	$\bar{c}_1^u$	$\bar{c}_2^u(\times 10^3)$
Constitutive	1′	7·56	17·2	7·56	5·89
	25′	5·73	13·0	5·73	4·46
	49′	4·34	9·86	4·34	3·38
Repression with $K_r = 10$	1′	0·180	0·409	0·305	0·238
	25′	0·157	0·356	0·265	0·206
	49′	0·136	0·309	0·230	0·179
Activation with $K_a = 5000$	1′	5·35	12·2	1·12	0·871
	25′	3·52	8·00	0	0
	49′	2·14	4·86	0	0

The effects on the mean mRNA and protein concentrations are shown for various gene locations and protein stabilities. In this table the generation time is $\tau = 60$, and the mRNA half-life is 2 min. This table is compared to Table 8. The changes between a shift up and shift down show minimal differences in the mean concentrations for an autorepressed gene, while there are major differences in the mean concentrations for the autoactivated gene including one case where the shift down causes the gene to shut off.

TABLE 10

A shift up of 10% *is applied to the rate constants for producing mRNA and protein for each of the constitutive, autorepressed, and autoactivated genes previously studied*

Gene type	Position	$\bar{c}_1^s$	$\bar{c}_2^s(\times 10^3)$	$\bar{c}_1^u$	$\bar{c}_2^u(\times 10^3)$
Constitutive	1′	12·3	29·1	12·3	14·8
	25′	7·64	18·1	7·64	9·16
	49′	4·74	11·2	4·74	5·69
Repression with $K_r = 10$	1′	0·226	0·535	0·162	0·383
	25′	0·178	0·421	0·127	0·301
	49′	0·139	0·330	0·100	0·236
Activation with $K_a = 5000$	1′	10·2	24·1	8·13	9·75
	25′	5·53	13·1	3·47	4·16
	49′	2·62	6·20	0·560	0·672

The effects on the mean mRNA and protein concentrations are shown for various gene locations and protein stabilities. In this table the generation time is $\tau = 30$, and the mRNA half-life is 2 min. A comparison of this table with Table 8 shows that the higher growth rate results in greater differences between the production of a gene proximal to *oriC* versus one distal to *oriC*. Again autorepression moderates this effect, while autoactivation accentuates it when compared to a constitutive gene.

TABLE 11

A shift down of 10% *is applied to the rate constants for producing mRNA and protein for each of the constitutive, autorepressed, and autoactivated genes previously studied*

Gene type	Position	$\bar{c}_1^s$	$\bar{c}_2^s(\times 10^3)$	$\bar{c}_1^u$	$\bar{c}_2^u(\times 10^3)$
Constitutive	1′	10·1	19·5	10·1	9·88
	25′	6·25	12·1	6·25	6·13
	49′	3·88	7·50	3·88	3·81
Repression with $K_r = 10$	1′	0·226	0·437	0·316	0·310
	25′	0·177	0·343	0·248	0·243
	49′	0·139	0·269	0·194	0·190
Activation with $K_a = 5000$	1′	7·49	14·5	4·96	4·87
	25′	3·67	7·10	1·15	1·13
	49′	1·29	2·50	0	0

The effects on the mean mRNA and protein concentrations are shown for various gene locations and protein stabilities. In this table the generation time is $\tau = 30$, and the mRNA half-life is 2 min. This table when compared to Table 10 shows a similar trend as between Tables 8 and 9. Notice that the higher growth rates in Tables 10 and 11 caused some autoactivated genes to become active when compared against entries in Tables 8 and 9.

before an increase in the rate of volume expansion would occur. The results are shown in Tables 8–11.

The effects of gene position are very dramatic among the different forms of control. When $\tau = 60$ min, a constitutive gene produces 74% more stable protein if the gene is located near *oriC* rather than the terminus. An autorepressed gene with $K_r = 10$ produces 32% more stable protein from the gene located proximal to *oriC* over the distal one, while the autoactivated gene with $K_a = 5000$ demonstrates a 112% increase because of gene position. When $\tau = 30$ min, the increase of the gene at 1′ from *oriC* over the one 1′ from the terminus for stable protein mean concentration is 160%, 62%, and 289% for the constitutive, autorepressed, and autoactivated genes, respectively. These results show that autorepression moderates the increase in the average concentration of a protein with respect to gene location, while gene position is very significant if the gene is autoactivated.

When the protein is unstable, the effects of gene position can be amplified. The constitutive or autorepressed gene producing an unstable protein shows results similar to the ones discussed above. However, the autoactivated gene with $K_r = 5000$ and $\tau = 60$ produces a mean concentration of 3780 molecules μm^{-3} at the gene near *oriC*, 1650 molecules μm^{-3} at a gene 25′ from *oriC*, and exhibits no production if it was located near the terminus when there is a shift up of 10%. With a shift down of 10% the autoactivated gene which is 25′ from *oriC* shuts down production as seen in Table 9. The datum associated with the autoactivated gene near the terminus for a 10% shift up when $\tau = 30$ min shows it to be expressed, which was not the case at $\tau =$ 60 min. However, this gene loses its ability to produce protein when there is a shift down of 10%, as seen in Table 11. Thus, an autoactivated gene with a relatively weak affinity for its activator can be strongly affected by both gene location and generation time.

An important question remaining in biology is how the cell manages to regulate its production of crucial proteins in response to a change in the growth media. A comparison of the entries in Table 8 to those in Table 9 shows that the mean concentrations of the mRNA, $\bar{c}_1^s$, have a difference of 22%, 1%, and 39–64% for the constitutive, autorepressed and autoactivated genes, respectively. A similar comparison of the mean concentrations of the stable protein, $\bar{c}_2^s$, shows differences of 49%, 23%, and 69–100% for the constitutive, autorepressed and autoactivated genes, respectively. Similar shifts in the mean concentrations are seen for $\tau = 30$ min when comparing entries in Tables 10 and 11. A comparison of the entries for the unstable proteins shows similar increases for the 10% shift up over the 10% shift down for the constitutive and autorepressed genes, but the increase is substantially higher for the autoactivated gene.

The analysis above demonstrates that an autorepressed gene responds very poorly to increases in the rate constants. Production of its protein is not very responsive to increases in concentrations of RNA polymerase or ribosomes, so this method of control acts as a buffer to changes in growth conditions and allows the cell to maintain relatively constant concentrations over a range of growth conditions. On the other hand, an autoactivated gene can have a very large response to small increases in the rate constants. This suggests that autoactivation could play an important role in a cell's response to changes in growth media. In particular, this would be the more appropriate choice of cellular control for a process that needs a rapid response to changing growth conditions such as timing of initiation of DNA replication.

5. Discussion

The mathematical models presented in this paper are simplifications of the actual biochemical processes controlling RNA and protein synthesis inside *E. coli*. However, as a first approximation, they provide useful information to compare the effects of type of control, gene position, molecular stability, and growth conditions for exponentially growing cultures of bacteria. Some of the results confirm previous intuition on this subject, but the mathematical models provide new quantitative information on the magnitude of these effects. On the other hand, several numerical simulations show qualitative behavior which has not been characterized or expected.

Surprisingly, our results demonstrate that the common belief that autorepression is the best way for a cell to maintain a relatively constant concentration is false. In fact, our calculations show that there is little difference in the percent variation of a constitutive gene when compared to one which is autorepressed. The computations performed in the previous sections show that stability of the RNA and protein molecules is the most significant factor in determining the amount of variation in concentration of a particular biochemical species through the cell cycle. Our results proved that the ratio between the maximum and the minimum concentrations for a constitutively produced stable RNA is about 6% independent of any factors such as growth rate or gene position. Similarly, the simulations for all cases (constitutive,

autorepressed, and autoactivated genes) show that a stable protein has approximately 5–6% variation in concentration over the cell cycle. As stability decreases for either the mRNA or the protein, the amount of variation between the minimum and maximum increases.

The mathematical models demonstrate that gene position is a very significant factor in determining what contribution a gene has to the overall protein content of the cell under a variety of growth conditions. The models show that at higher growth rates the genes proximal to the origin of replication contribute a higher percentage of their products to the cell than the ones more distal to the origin. For example, a constitutive gene located 1′ from the origin will produce 74% more stable RNA than a gene located 1′ from the terminus when $\tau = 60$ min. When τ is decreased to 30 min, the gene closer to the origin produces 160% more stable RNA than the one near the terminus. This holds the evolutionary implication that the genes required for growth of a cell should be located closer to the origin of replication. This certainly is the case for the *rrn* genes and may be a contributing factor to growth rate-regulated expression of these genes (see Cashel & Rudd, 1987 for a discussion of growth rate regulation).

The mathematical analyses of our model provide several methods for estimating parameters such as decay rates. A comparison of the initiation rates reported by Bremer & Dennis (1987) to ones calculated from the mean concentrations give estimates of the half-life for rRNA assuming rRNA is lost only by dilution and decay. If an experiment was designed to measure the amount of variation of the concentration for a particular mRNA through a cell cycle, then in theory our model could be used to find a bound on the half-life of the mRNA. However, the sensitivity analysis shows that any evaluation using this technique would be highly subject to error. Our analysis of the experiments on the *lac* operon by Kennell & Riezman (1977) allowed us to calculate the failure rate for production of mRNAs for β-galactosidase and estimate the average concentration of β-galactosidase in a fully induced cell. The computations from the mathematical models agree well with the results which were obtained experimentally.

The theoretical studies comparing constitutive, autorepressed, and autoactivated genes provide some insight into how these controls might best be used by a cell. The strength of autorepression had no effect on the percent of variation in concentration of the repressed protein. However, autorepression was shown to moderate the effects of a shift up or an increase in growth rate on the mean concentration of the resultant protein. This suggests that autorepression is most valuable for controlling proteins which need to be maintained in low, relatively constant concentrations under a variety of conditions. Proteins whose concentrations vary substantially over the cell cycle or need to respond rapidly to changing growth conditions should be unstable or require some alternative type of control.

An autoactivated gene has several interesting properties owing to the possible existence of both an "on" state and an "off" state. This type of control exhibited a much higher response to changes in the modeling parameters such as rate of production (shift up) or gene position. Our numerical studies showed that the autoactivated gene may be unexpressed at one growth rate but then become expressed at a higher

growth rate. In addition, when located at one position the gene may be unexpressed, implying that the zero solution to the model for autoactivation, (11), is stable, but when the same gene is moved closer to *oriC*, it may become active. In the latter case the solution to (11) is a stable limit cycle that oscillates about some positive mean concentration. This type of control allows the cell to respond rapidly to variations in growth conditions and could play an important role during a shift up or a shift down owing to nutritional changes.

Thus, crucial factors requiring a rapid response to increased growth conditions such as the key components for cell cycle control might utilize some type of autoactivation or some mathematically equivalent type of control. The *dnaA* gene has two promoters, *dnaA*1P and *dnaA*2P. The *dnaA*1P is repressed by DnaA protein, while the *dnaA*2P promoter appears to be activated by DnaA protein (Polaczek & Wright, 1990). Its not clear why the DnaA protein, which plays a vital role in determining when *E. coli* initiates replication, should adopt these two types of strategies for its expression. However, the mathematical analysis above indicates that autoactivation at one of the DnaA promoters might improve this gene's response to changing growth conditions. The response of the *dnaA* gene to varying growth rates in turn might affect changes in initiation rates for DNA synthesis and thus play a key role in cell cycle control.

The author wishes to thank Dr Judith Zyskind for her valuable input in the preparation of this manuscript. Her assistance has improved the biological relevance of this work dramatically.

REFERENCES

BACHMANN, B. J. (1990). Linkage map of *Escherichia coli* K-12, edition 8. *Microbiol. Rev.* **54,** 130–197.

BANKS, H. T. (1975). *Modeling and Control in the Biomedical Sciences, Vol. 6 of Notes in Biomathematics.* Berlin: Springer.

BANKS, H. T. & MAHAFFY, J. M. (1978). Global asymptotic stability of certain models for protein synthesis and repression. *Quart. appl. Math.* **36,** 209–221.

BARACCHINI, E. & BREMER, H. (1988). Stringent and growth control of rRNA synthesis in *Escherichia coli* are both mediated by ppGpp. *J. Biol. Chem.* **263,** 2579–2602.

BECKWITH, J. (1987). The lactose operon. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology,* Vol. 2 (Neidhart, F., ed.) pp. 1444–1452. Washington, DC: ASM.

BREMER, H. & DENNIS, P. P. (1987). Modulation of chemical composition and other parameters of the cell by growth rate. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology,* Vol. 2 (Neidhart, F., ed.) pp. 1527–1542. Washington, DC: ASM.

CASHEL, M. & RUDD, K. E. (1987). The stringent response. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology,* Vol. 2 (Neidhart, F., ed.) pp. 1410–1438. Washington, DC: ASM.

CHURCHWARD, G., BREMER, H. & YOUNG, R. (1982). Transcription in bacteria at different DNA concentrations. *J. Bacteriol.* **150,** 572–581.

COOPER, S. (1988*a*). Leucine uptake and protein synthesis are exponential during the division cycle of *Escherichia coli* B/r. *J. Bacteriol.* **170,** 436–438.

COOPER, S. (1988*b*). What is the bacterial growth law during the division cycle? *J. Bacteriol.* **170,** 5001–5005.

COOPER, S. (1991). *Bacterial Growth and Division: Biochemistry and Regulation of Prokaryotic and Eukaryotic Division Cycles.* San Diego, CA: Academic Press.

DAHLQUIST, G. & BJÖRCK, Å. (1974). *Numerical Methods,* translated by Anderson, N. Englewood Cliffs, NJ: Prentice-Hall.

DENNIS, P. P. (1974). *In vivo* stability, naturation, and relative differential synthesis rates of individual ribosomal proteins in *Escherichia coli* B/r. *J. molec. Biol.* **88,** 25–41.

DONACHIE, W. D. & ROBINSON, A. C. (1987). Cell division: parameter values and the process. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology*, Vol. 2 (Neidhart, F., ed.) pp. 1527–1542. Washington, DC: ASM.

GAUSING, K. (1977). Regulation of ribosome production in *Escherichia coli*: synthesis and stability of ribosomal RNA and of ribosomal messenger RNA at different growth rates. *J. molec. Biol.* **115,** 335–354.

GAUSING, K. (1980). Regulation of ribosome biosynthesis in *E. coli* In: *Ribosomes: Structure, Function and Genetics* (Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L. & Nomura, M., eds) pp. 693–718. Baltimore, MD: University Park.

GOLDBERG, A. L. & ST JOHN, A. C. (1976). Intracellular protein degradation in mammalian and bacterial cells: part 2. *Ann. Rev. Biochem.* **45,** 747–803.

GOODWIN, B. C. (1965). Oscillatory behavior of enzymatic control processes. *Adv. Enzyme Reg.* **3,** 425–439.

GOURSE, R. L., TAKEBE, Y., SHARROCK, R. A. & NOMURA, A. (1985). Feedback regulation of rRNA and tRNA synthesis and accumulation of free ribosomes after conditional expression of rRNA genes. *Proc. natn. Acad. Sci. U.S.A.* **82,** 1069–1073.

HELMSTETTER, C. E. (1987). Timing of synthetic activities in the cell cycle. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology*, Vol. 2 (Neidhart, F., ed.) pp. 1594–1605. Washington, DC: ASM.

INGRAHAM, J. L., MAALØE, O. & NEIDHARDT, F. C. (1983). *Growth of the Bacterial Cell.* Sunderland, MA: Sinauer Assocs.

JENSEN, K. F. & PEDERSEN, S. (1990). Metabolic growth rate control in *Escherichia coli* may be a consequence of subsaturation of the macromolecular biosynthetic apparatus with substrates and catalytic components. *Microbiol. Rev.* **54,** 89–100.

JINKS-ROBERTSON, S., GOURSE, R. L. & NOMURA, M. (1983). Expression of rRNA and tRNA genes in *Escherichia coli*: evidence for feedback regulation by products of rRNA operons. *Cell* **33,** 865–876.

JINKS-ROBERTSON, S. & NOMURA, M. (1987). Ribosomes and tRNA. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology*, Vol. 2 (Neidhardt, F., ed.) pp. 1358–1385. Washington, DC: ASM.

KENNELL, D. & RIEZMAN, H. (1977). Transcription and translation initiation frequencies of the *Escherichia coli lac* operon. *J. molec. Biol.* **114,** 1–21.

KUBITSCHEK, H. E. & PAI, S. R. (1988). Variation in precursor pool size during the division cycle of *Escherichia coli*: further evidence for linear cell growth. *J. Bacteriol.* **170,** 431–435.

MAALØE, O. (1979). Regulation of the protein-synthesizing machinery – ribosomes, tRNA, factors, and so on. In: *Biological Regulation and Development* (Goldberger, R. F., ed.) pp. 487–542. New York: Plenum.

MANDELSTAM, J. (1957). Turnover of protein in starved bacteria and its relationship to the induced synthesis of enzyme. *Nature, Lond.* **179,** 1179–1181.

MATZURA, H., HANSEN, B. S. & ZEUTHEN, J. (1973). Biosynthesis of the β and β' subunits of RNA polymerase in *Escherichia coli*. *J. molec. Biol.* **74,** 9–20.

NEIDHARDT, F. C., INGRAHAM, J. L. & SCHAECHTER, M. (1990). *Physiology of the Bacterial Cell.* Sunderland, MA: Sinauer Assoc.

NOMURA, M., GOURSE, R. L. & BAUGHMAN, G. (1984). Regulation of the synthesis of ribosomes and ribosomal components. *Ann. Rev. Biochem.* **53,** 75–117.

OTHMER, H. G. (1976). The qualitative dynamics of a class of biochemical control circuits. *J. math. Biol.* **3,** 53–78.

POLACZEK, P. & WRIGHT, A. (1990). Regulation of expression of the *dnaA* gene in *Escherichia coli*: role of the two promoters and the DnaA box. *New Biol.* **2,** 574–582.

PTASHNE, M. (1986). *A Genetic Switch: Gene Control and Phage λ.* Cambridge, MA: Cell Press and BSP.

SARUBBI, E. K., RUDD, E. & CASHEL, M. (1988). Basal ppGpp level adjustment shown by new *spoT* mutants affect steady state growth rates and *rrnA* ribosomal promoter regulation in *Escherichia coli*. *Mol. Gen. Genet.* **213,** 214–222.

SHEPARD, N. S., CHURCHWARD, G. & BREMER, H. (1980). Synthesis and activity of ribonucleic acid polymerase in *Escherichia coli*. *J. Bacteriol.* **141,** 1098–1108.

SKARSTAD, K., BOYE, E. & STEEN, H. B. (1986). Timing of initiation of chromosome replication in individual *Escherichia coli* cells. *EMBO J.* **5,** 1711–1717.

SKARSTAD, K., VON MEYENBURG, K., HANSEN, F. G. & BOYE, E. (1988). Coordination of intiation of chromosome replication in *E. coli*: effects of different *dnaA* alleles. *J. Bacteriol.* **129,** 1466–1475.

SOMPAYRAC, L. & MAALØE, O. (1973). Autorepressor model for control of DNA replication. *Nature New Biol.* **241,** 133–135.

APPENDIX A

Stable RNA Production

The differential equation for the concentration of a constitutively produced RNA is derived using a balance of mass law. When the RNA is stable, then the change in mass, $m(t)$, is simply the rate of RNA production per gene times the number of active genes. That is

$$\frac{\mathrm{d}m(t)}{\mathrm{d}(t)} = rG_a(t), \tag{A.1}$$

where r and G_a are defined in section 2. For $0 \le t \le \tau$, the concentration is given by $c(t) = m(t)/V(t)$ which by differentiation and (A.1) yields:

$$\begin{aligned} \dot{c}(t) &= \frac{\dot{m}(t)}{V(t)} - \frac{m(t)\dot{V}(t)}{V^2(t))} \\ &= \frac{rG_a(t)}{V(t)} - \frac{c(t)\dot{V}(t)}{V(t)}. \end{aligned} \tag{A.2}$$

For an exponentially growing cell, $\dot{V}(t) = kV(t)$, where $k = \ln(2)/\tau$. With this hypothesis and continuity of concentration at cell division, (A.2) becomes the boundary value problem,

$$\dot{c}(t) = \frac{rG_a(t)}{V(t)} - kc(t), \qquad c(0) = c(\tau). \tag{A.3}$$

In fact, one could use any periodic boundary conditions of the form $c(t) = c(t+\tau)$. This is a linear constant coefficient differential equation with an integrating factor of e^{kt}. Thus, the solution of (A.3) satisfies:

$$c(t) = \mathrm{e}^{-kt}\left[c(0) + \int_0^t \frac{rG_a(s)}{V(s)}\,\mathrm{e}^{ks}\,\mathrm{d}s\right] \quad \text{for } 0 < t < \tau,$$

with the periodic boundary conditions. Using $V(t) = V_0\,\mathrm{e}^{kt}$, $c(0) = c(\tau)$, and $\mathrm{e}^{-kr} = \frac{1}{2}$, we find

$$c(0) = \int_0^\tau \frac{rG_a(s)}{V_0}\,\mathrm{d}s.$$

Thus, the solution for (A.3) can be written as follows:

$$\begin{aligned} c(t) &= \mathrm{e}^{-kt}\left[\int_0^\tau \frac{rG_a(s)}{V_0}\,\mathrm{d}s + \int_0^t \frac{rG_a(s)}{V_0}\,\mathrm{d}s\right] \\ &= \frac{Qr\,\mathrm{e}^{-kt}}{V_0}\left[2\tau - t_0 + \begin{cases} t, & 0 \le t \le t_0 \\ 2t - t_0, & t_0 < t \le \tau \end{cases}\right], \end{aligned} \tag{A.4}$$

from the definition of G_a.

Consider a sequence of initial value problems (A.3) with c_j, $j = 1, \ldots,$ replacing c in the differential equation. For initial conditions take $c_1(0)$ to be arbitrary and

$c_j(0)=c_{j-1}(\tau)$, $j=2,\ldots$. Because of the leading exponential factor, one can readily show by Poincaré maps that this sequence of solutions rapidly converges to the periodic solution (A.4). Thus, (A.4) is a globally attracting solution to the differential eqn (A.3) with any initial conditions which makes it the appropriate definition of a steady-state solution.

The minimum and maximum values are found at the critical points of (A.4). It is not hard to see that the minimum occurs when $t=t_0$ with

$$c_{\min}=\frac{2\tau Qr\,e^{-kt_0}}{V_0}.$$

The maximum occurs at $t=t_m$ with $\dot{c}(t_m)=0$. This is readily found from (A.3) by letting $\dot{c}(t_m)=0$, inserting (A.4) with $t=t_m$, and solving for t_m. Thus,

$$0=\frac{rG_a(tm)}{V(t_m)}-kc(t_m)$$

or

$$k\left[2\tau-t_0+\begin{cases}t_m, & 0\le t_m\le t_0\\ 2t_m-t_0, & t_0\le t_m\le\tau\end{cases}\right]=\begin{cases}1, & 0\le t_m\le t_0\\ 2, & t_0<t_m\le\tau\end{cases}.$$

Solving for t_m, we find

$$t_m=t_0+\begin{cases}\frac{1}{k}-\tau, & t_m>t_0\\ \frac{1}{k}-2\tau, & t_m<t_0\end{cases}\simeq t_0+\begin{cases}0{\cdot}443\tau, & t_m>t_0\\ -0{\cdot}557\tau, & t_m<t_0\end{cases},$$

which always yields

$$c_{\max}=\frac{4Qr\,e^{-kt_0}}{e\,kV_0}.$$

The average concentration in an asynchronously, exponentially growing culture of cells is found by integrating the concentration, $c(t)$, over the cell cycle divided by the cell cycle time. The volume of the cell by assumption is given by $V(t)=V_0e^{kt}$, where $0\le t\le\tau$ is the age of the cell. The distribution in number of cells for a particular age $0\le t\le\tau$ is given by $P(t)=P_0e^{-kt}$ for some P_0. The average concentration, $\bar{c}$, is found by integrating $c(t)$ over the ages of the cells times the weighting factor of the volume times the population, which is then divided by the mean volume. The mean volume is given by the integral over the ages of the volume times the population. Thus,

$$\bar{c}=\frac{\int_0^\tau c(s)V(s)P(s)\,ds}{\int_0^\tau V(s)P(s)\,ds}=\frac{V_0P_0\int_0^\tau c(s)\,ds}{\tau V_0P_0}=\frac{1}{\tau}\int_0^\tau c(s)\,ds,$$

$$=\frac{Qr}{\tau V_0}\left[\int_0^{t_0}(2\tau-t_0+s)\,e^{-ks}\,ds+2\int_{t_0}^{\tau}(\tau-t_0+s)\,e^{-ks}\,ds\right],$$

$$=\frac{Qr\,e^{-kt_0}}{kV_0\ln(2)}.$$

APPENDIX B

A Constitutive Protein

The differential equations for the concentrations of the mRNA and protein in section 3 are derived in a manner similar to the one presented in Appendix A, where the principal difference is the inclusion of a linear decay term in each differential equation. We begin by solving (7) as follows:

$$c_1(t) = e^{-(k+k_1)t}\left[c_1(0) + \int_0^t \frac{r_1 G_a(s)}{V(s)} e^{-(k+k_1)s}\, ds\right] \quad \text{for } 0 \le t \le \tau,$$

$$= e^{-(k+k_1)t}\left[c_1(0) + \frac{Qr_1}{2V_0 k_1}\begin{cases} e^{k_1 t} - 1, & 0 \le t \le t_0 \\ 2e^{k_1 t} - e^{k_1 t_0} - 1, & t_0 < t \le \tau \end{cases}\right], \tag{B.1}$$

from the definition of V and G_a.

It is easy to show that for any initial condition $c_1(0)$ the solution (B.1) exponentially approaches the periodic solution satisfying the periodic boundary conditions, $c_1(0) = c_1(\tau)$. To see this let $c_1^*(t)$ be the periodic solution and $y_1 = c_1 - c_1^*$ be any perturbation of (7). It follows that y_1 satisfies the differential equation

$$\dot{y}_1 = \dot{c}_1 - \dot{c}_1^* = -(k_1 + k) y_1,$$

which exponentially approaches the zero solution. Thus, the periodic solution is a globally attracting solution. To satisfy the periodic boundary conditions

$$c_1(0) = \frac{Qr_1(2 - e^{-k_1(\tau - t_0)} - e^{-k_1 \tau})}{V_0 k_1 (2 - e^{-k_1 \tau})}.$$

Inserting the value of $c_1(0)$ into (B.1), we obtain

$$c_1(t) = \frac{Qr_1 e^{-kt}}{V_0 k_1 (2 - e^{-k_1 \tau})}\begin{cases} 2 - e^{-k_1 \tau}(1 + e^{-k_1(t_0 - t)}), & 0 \le t \le t_0, \\ 4 - 2(e^{-k_1 t} + e^{-k_1(t_0 - t)}), & t_0 \le t \le \tau, \end{cases} \tag{B.2}$$

where Q and t_0 are defined by (4).

Equation (7) is solved by the following:

$$c_2(t) = c_2(T)\, e^{-(k+k_2)(t-T)} + \int_T^t r_2 c_1(s - T)\, e^{(k+k_2)(s-t)}\, ds, \qquad T < t < T + \tau.$$

This equation is solved by expanding the integral using (B.2). An argument similar to the one given above shows that the periodic solution for (8) is globally attracting. The value of $c_2(T)$ for this attracting periodic solution is found by letting, $c_2(T) = c_2(T + \tau)$. This information is combined, and with the help of the symbolic

manipulator, Macsyma™, as a check on hand calculations, the results are given by

$$c_2(t)=\frac{Qr_1r_2\,e^{-(k+k_2)(t-T)}}{V_0k_1k_2(k_2-k_1)(2-e^{-k_1\tau})(2-e^{-k_2\tau})}\begin{cases}h_1(t), & 0\le t-T<t_0\\ h_2(t), & t_0\le t-T\le\tau\end{cases}, \tag{B.3}$$

where

$$h_1(t)=k_1(2-e^{-k_1\tau})\,[e^{k_2(t_0-\tau)}-(2-e^{-k_2\tau})\,e^{k_2(t-T)}]$$
$$+k_2(2-e^{-k_2\tau})[(2-e^{-k_1\tau})\,e^{k_2(t-T)}-e^{k_1(t_0-\tau)}\,e^{(k_2-k_1)(t-T)}]$$

and

$$h_2(t)=2k_1(2-e^{-k_1\tau})\,[e^{k_2t_0}-(2-e^{-k_2\tau})e^{k_2(t-T)}]$$
$$+2k_2(2-e^{-k_2\tau})\,[(2-e^{-k_1\tau})e^{k_2(t-T)}-e^{k_1t_0}e^{(k_2-k_1)(t-T)}].$$

As in Appendix A, the minimum and maximum values for c_1 are found at the critical points of (B.2). One critical value occurs at $t=t_0$ which is the minimum and is given by

$$c_1(t_{\min})=\frac{2Qr_1\,e^{-kt_0}(1-e^{-k_1\tau})}{V_0k_1(2-e^{-k_1\tau})}. \tag{B.4}$$

The other critical value, the maximum concentration, occurs when $\dot{c}_1(t_{\max})=0$. If $t_0<t_{\max}<\tau$, then from (B.2),

$$\frac{Qr_1}{V_0k_1(2-e^{-k_1\tau})}[-2ke^{-kt_{\max}}(2-e^{-k_1\tau})+2(k+k_1)\,e^{-(k+k_1)t_{\max}}\,e^{k_1t_0}]=0.$$

This equation is solved for $t_{\max}$, and the result is

$$t_{\max}=t_0+\frac{1}{k_1}\ln\left[\frac{(k_1+k)}{k(2-e^{-k_1\tau})}\right].$$

If $0<t_{\max}<t_0$, then a similar calculation shows that

$$t_{\max}=t_0-\tau+\frac{1}{k_1}\ln\left[\frac{(k_1+k)}{k(2-e^{-k_1\tau})}\right].$$

When $0<t_{max}<\tau$ is substituted into (B.2), the result is given by

$$c_1(t_{\max})=\frac{2Qr_1\,e^{-kt_0}}{V_0(k_1+k)}\left[\frac{k(2-e^{-k_1\tau})}{(k_1+k)}\right]^{k/k_1}. \tag{B.5}$$

Equation (B.3) is sufficiently complicated that finding minimum and maximum values is best done by numerically evaluating this equation over the interval $0<t<T$.

The average concentrations, $\bar{c}_1$ and $\bar{c}_2$, can be found in a manner similar to the one outlined in Appendix A though the calculations are slightly more complicated. These calculations were assisted by the use of the symbolic manipulator, Macsyma™,

and are given by the following expressions:

$$\bar{c}_1 = \frac{1}{\tau}\int_0^{\tau} c_1(s)\,ds = \frac{Qr_1\,e^{-kt_0}}{V_0(k+k_1)\ln 2},$$

$$\bar{c}_2 = \frac{1}{\tau}\int_0^{\tau} c_2(s)\,ds = \frac{Qr_1r_2\,e^{-kt_0}}{V_0(k+k_1)(k+k_2)\ln 2}.$$

These expressions can be rearranged to give information on the rates of initiation. For example,

$$r_1 = \frac{\bar{c}_1 V_0(k+k_1)\ln 2}{Q\,e^{-kt_0}}. \tag{B.6}$$

APPENDIX C

Production: Repression and Activation

In this Appendix we examine the biochemical kinetics needed to find the production terms used in (10) and (11). From before, we define G_a as the number of genes in the cell. Let $[O_T] = G_a/V$ be the concentration of the operator region for the gene that is being considered in all its possible states. For both models the operator region is in one of two states. Either it has a protein bound to it or it is unbound. In the case of repression, when the repressor protein is bound to the operator, then it is inactive. Otherwise, the operator is active. In the case of activation, if the activator is bound to the operator, then the operator is active. Otherwise, the operator is inactive. Let $[OX]$ be the concentration of the operator–protein complex and $[O]$ be the concentration of the unbound operator, then

$$[O_T] = [O] + [OX], \tag{C.1}$$

where X is either the repressor, R, or the activator, A.

Let $[X_T]$ be the total concentration of the repressor or activator protein. If the number of these molecules greatly exceeds the few bound to the operator region ($[X_T] \gg [OX]$), which is often the case, then we can write

$$[X_T] = [X] + [OX] \simeq [X], \tag{C.2}$$

where $[X] = [R]$ or $[X] = [A]$ depending on the case being considered.

Assuming that the formation of operator–repressor complex is not rate limiting, we can write the equilibrium expression

$$K_r = \frac{[R][O]}{[OR]}.$$

With (C.1) and approximation (C.2), it is easy to see that

$$[O_T]=[O](1+[R_T]/K_r),$$

or

$$[O]=\frac{K_r[O_T]}{K_r+[R_T]}. \tag{C.3}$$

This yields the free operator concentration which is ready for initiation of transcription. If we let $[R_T]=c_2$ and r_1 be the rate of initiation which assumes a constant polymerase concentration, then the rate of production of new mRNAs is the free operator concentration given by (C.3) times r_1 which yields the expression,

$$\frac{r_1K_rG_a(t)}{V(t)(K_r+c_2(t))}.$$

If the formation of operator-activator complex is not rate limiting, then we can write the equilibrium expression

$$K_a=\frac{[A][O]}{[OA]}.$$

With (C.1) and approximation (C.2), it follows that

$$[O_T]=[OA](1+K_a/[A_T]),$$

or

$$[OA]=\frac{[O_T][A]}{K_a+[A_T]}. \tag{C.4}$$

This gives the operator–activator concentration which is active and can begin transcription. If we let $[A_T]=c_2$ and r_1 be the rate of initiation which assumes a constant polymerase concentration, then using (C.4) we find that the rate of production of new mRNAs satisfies the expression,

$$\frac{r_1G_a(t)c_2(t)}{V(t)(K_a+c_2(t))}.$$

Note that the approximation (C.2) is not valid for very low levels of activator protein. By using this approximation, the model produces a zero equilibrium state rather than a low basal level of protein as would occur in a cell.