A Novel Structured Kinetic Modeling Approach for the Analysis of Plasmid Instability in Recombinant Bacterial Cultures

William E. Bentley and Dhinakar S. Kompala
Department of Chemical Engineering, University of Colorado, Boulder, Colorado 80309-0424

Accepted for publication December 18, 1987

The instantaneous specific growth rate of a recombinant bacterial culture is directly calculated using a simple structured kinetic modeling approach. Foreign plasmid replication and foreign protein expression represent metabolic burdens to the host cell. The individual effects of these plasmid-mediated activities on the growth rate of plasmid-bearing cells are estimated separately. The dynamic and steady state simulations of the model equations show remarkable agreement with widely observed experimental trends in plasmid copy number and foreign protein content. The model provides an important tool for understanding and controlling plasmid instability in recombinant bacterial fermentations. The modeling framework employed here is suitable for studying the metabolism and growth of a variety of microbial cultures.

INTRODUCTION

Commercial exploitation of recombinant DNA techniques is dependent upon successful maintenance of recombinant strains in continuous and batch processes. Structural and segregational instabilities have been identified which can result in the dramatic loss of a plasmid-bearing, recombinant strain from continuous cultures.[1,2] A significant contribution to this loss is from a growth rate difference between the recombinant strain, with plasmid replication and foreign protein expression, and the plasmid-free host which has lost the ability to produce foreign proteins. Several experimental studies have established that as plasmid copy number and foreign protein overexpression increase, the growth rate of the recombinant host cell decreases.[3-7] It is generally recognized that the replication of high copy number plasmids and the overproduction of plasmid-encoded proteins represent an additional "metabolic burden" on the normal chromosome-directed metabolism of the bacterial cell. The metabolic intermediates, such as nucleotides and amino acids, which are normally used for replication of the chromosome and production of chromosome-encoded proteins, are being distributed in the recombinant cell between the chromosome- and plasmid-directed activities, thereby reducing the cell growth rate.

This reduction in growth rate, along with the generation of a small fraction of plasmid-free cells which occurs during cell-division of the recombinant strain, plays a key role in the loss of plasmid-containing cells and their replacement by plasmid-free cells in continuous fermentations. Imanaka and Aiba[8] examined these factors and showed that there could be a precipitous loss of the plasmid-carrying cells from a chemostat within ca. 30 generations. They assumed the following in their analysis: a smaller value for the specific growth rate of the plasmid-containing cells, μ^+, compared to the specific growth rate of the plasmid-free cells, μ^-, and a small constant, p, for the probability of generation of plasmid-free cells from plasmid-containing cells upon cell division. Their simulation results, using constant values for μ^+, μ^-, and p, are in good qualitative agreement with a large number of experimental observations on "plasmid instability" in continuous and extended batch cultures.

More recent experimental studies, however, indicate that the plasmid copy number and the extent of foreign protein expression are not constant for any given host–vector system. Instead, they depend strongly on the growth rate of the recombinant cell.[9-16] In almost all cases, there was a monotonic decrease in plasmid content (per genome equivalent) with an increase in specific growth or dilution rate. The only exception to this commonly observed trend is the work of Koizumi and co-workers[16] which indicates an increase in plasmid content with growth rate at certain temperatures and no change at others. This exception has caused some authors to suggest that there may be no unifying thread tying plasmid content to the growth rate or "state" of the cell. Some others have further indicated that it may not be accurate to compare the data of continuous culture studies to those using batch cultures with varying media. However, a good comparison exists between the chemostat results of Siegel and Ryu[10] (using *E. coli* K-12Δ H1 Δ*trp*:pPLc23 *trpAl*) and the batch results of Lin-Chao and Bremer[15] (using *E. coli* B/rA:pBR322) even though the experimental systems were different.

Biotechnology and Bioengineering, Vol. 33, Pp. 49–61 (1989)

CCC 0006-3592/89/010049-13$04.00

Thus, with the exception of results from Koizumi and co-workers, it would appear that there is a distinct trend in plasmid content and foreign protein expression with cell growth rate. As the balanced phase growth rate or continuous culture dilution rate is increased, the plasmid copy number and cloned gene product content decreases. In either case, the additional metabolic burdens of plasmid replication and foreign protein overexpression are both functions of growth rate and thus, so is the growth rate differential between the plasmid-carrying and plasmid-free cells.

In the present work, a simple structured kinetic model is proposed that is capable of predicting the growth rate of the bacterial cell mass as a dynamic function of the additional metabolic burdens. It is also shown to predict the widely observed trends in plasmid copy number and foreign protein expression as a function of growth rate. Furthermore, its value is demonstrated by predicting other experimentally observed phenomena in what is commonly referred to as "plasmid stability." The existing approaches for modeling the growth of recombinant bacteria are briefly discussed below so that the novel concepts employed in our approach become more clear.

Recently, Lee and Bailey[17] have suggested the following equation for expressing the growth rate of recombinant bacteria, μ_{rec}, as a function of plasmid content, foreign protein content, and the extracellular growth-limiting substrate concentration:

$$\mu_{rec} = \mu_{max}\left[1 - \frac{G}{G_{max}}\right]^m\left[1 - \frac{p}{p_{max}}\right]^n\left[\frac{S}{K_s + S}\right]$$

In this equation, G and G_{max} are the intracellular cloning vector concentration and the maximum attainable; p and p_{max} represent the existing product content and its maximum; K_s and μ_{max} are the typical Monod constants; and S represents the extracellular substrate concentration of the limiting nutrient. While this equation recognizes the independent and separable effects of plasmid replication and product expression on the growth rate of recombinant bacteria, it is clearly a phenomenological model, with a limited biochemical or mechanistic basis for the determination of G_{max}, p_{max}, m, and n.

A highly structured model for a single *E. coli* cell (Shuler and co-workers[18]) has been adapted to include equations for the mechanisms of plasmid replication and expression of foreign protein (ref. 19). This model is quite detailed biochemically in that it describes the anabolism and catabolism of 28 cellular components with 35 differential equations. The success of this model for determining the activity of an *Escherichia coli* cell has been well documented (Domach et al.[20]). More recently, Peretti and Bailey[21] have extended this model to encompass the kinetics and control of transcription and translation. In addition, they have adapted this model to simulate the effects of vector presence on cell metabolism.[22] The Cornell model and its extensions have been used to simulate single-cell bacterial growth mechanisms under various environmental conditions. Simulation results which yield favorable enhancements such as increased cloned gene product can be used to direct research on particularly promising areas, an example being transcription capacity.[22] The extensions of the Cornell model to predict population dynamics were, however, projected as too cumbersome for on-line control[23] and may prove to be too cumbersome for the design of optimal bioreactor strategies. The utility of such highly detailed models in studying the dynamic growth rate effects of plasmid-bearing versus plasmid-free bacteria in bioreactors may thus be limited by their size and complexity.

MODEL DEVELOPMENT

A novel approach is demonstrated here for directly calculating the instantaneous specific growth rate of a microorganism. A key feature of this modeling framework is the uncommon representation of the state of a microorganism in terms of the fractional mass levels of its intracellular constituents. The formalism of this method will be demonstrated after the rate expressions specific to recombinant bacteria are developed. The methodology for calculating the instantaneous specific growth rate, however, remains independent of the rate expressions used. Thus the methodology utilized here is equally applicable to a large variety of cultures, such as plasmid-free bacteria, yeasts, plant and animal cells, as well as transformed cell lines such as hybridomas.

The amount of detail included in our schematic representation of the metabolism of recombinant bacteria (see Fig. 1) strikes a middle ground between the simple phenomenological models and the more detailed, mechanistic models. In order to predict the reduction in growth rate

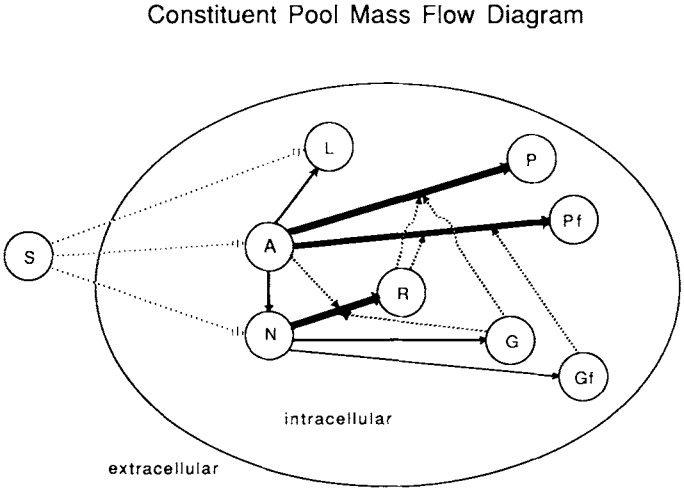

Figure 1. Constituent pool mass flow diagram. Cell mass is lumped into eight pools. Arrow width indicates the relative quantity of material flow between pools. Dotted lines indicate flow of information. Constituent pools include: (A) amino acids; (N) nucleotides; (P) protein; (Pf) foreign protein; (G) chromosomal DNA; (Gf) foreign DNA (plasmid); (L) lipid and membrane material; and (R) ribosomal RNA.

due to the additional metabolic burdens, and the functional variation of plasmid copy number and foreign protein activity with dilution rate, a mathematical model should of necessity include sufficient details of cellular metabolism. However, to keep such a model simple, the metabolic details are lumped into a minimal level of necessary variables and appropriate kinetic expressions.

Our lumped metabolic model of recombinant cells includes eight major intracellular constituent pools. These are: protein, P; foreign protein, P_f; chromosomal DNA, G; plasmid DNA, G_f; ribosomes, R; lipids, L; nucleotides, N; and amino acids, A. The width of the mass flow vectors in Figure 1 indicates the relative size of the mass flow. For example, the arrow projected from the amino acid pool towards the lipid pool is significantly thinner than the others indicating that only a minor fraction of the amino acids are converted to lipids. The flow of information is indicated by dotted lines. For example, a dotted line from G_f to P_f indicates that the synthesis of foreign protein, P_f, is related to the plasmid content, G_f.

Rate Expressions

The model consists of a differential equation for each pooled constituent (Tables I and II). Each differential equation consists of a rate of synthesis term and the appropriate terms for depletion or turnover. The synthesis rate expressions are comprised of saturation kinetic terms relating intrinsic cellular components as previously described by Shuler.[24] Parameters include a maximum rate and a saturation constant. In this way, the rate of formation of a particular constituent depends upon the precursor or reactant concentrations with the order varying between zero and one. Stoichiometric coefficients maintain the material balance between constituents.

External nutrients pass through the cell membrane and are consumed by the formation of lipids, amino acids, and nucleotides. For simplicity, transport mechanisms have been omitted and only the carbon substrate is assumed to be rate-limiting. The extracellular concentration of this substrate plays a major role in determining the synthesis rates of these three constituent pools.

The amino acid pool consists of all amino acids and their tricarboxylic acid (TCA) cycle precursors. Since TCA-cycle precursors were included in the same pool, synthesis of the amino acid pool is autocatalytic, in addition to having dependence on the extracellular substrate concentration. When the levels of TCA-cycle intermediates are low, the synthesis rate of the amino acids will be reactant limited, which is biochemically consistent. End-product inhibition of the biosynthetic pathways was included by a product inhibition term similar in form to that of Aiba and co-workers.[25] Thus, the overall synthesis rate expression is the product of an extracellular substrate dependent term and two amino acid dependent terms. The dependence on amino acids is similar in form to the microbial growth dependence on substrate in the substrate inhibition model proposed by Andrews.[26] Amino acids are depleted by the formation of lipids, proteins, and nucleotides. Depletion is accounted for by the appropriate stoichiometric constant relating the synthesis rate of the particular product to the amino acid consumption.

The nucleotide pool consists of both the deoxyribose and ribose moieties of the triphosphorylated and diphosphorylated molecules. Inorganic phosphate required for the interconversion of these molecules is also included. The synthesis rate expression includes a product inhibition term which accounts for the observed feedback control mechanism.[27] Nucleotide synthesis is also directed by the extracellular substrate concentration. Since the base portion of all nucleotides is primarily constructed from amino acids,[27] a saturation kinetic term with dependence on the amino acid level was included. Since the nucleotide pool consists of both deoxyribose and ribose moieties, pool depletion occurs by the production of rRNA and DNA, both chromosomal and plasmid. Again, stoichiometric coefficients maintain the material balance.

Proteins, both foreign and endogenous, are synthesized from amino acids. A saturation kinetic term was included to characterize the dependence of protein synthesis on amino acids. Since the rate of protein synthesis is directly proportional to the size of the "protein synthesizing system (PSS)",[28] and since we have represented the PSS principally as the rRNA pool, we have included its level directly in the synthesis rate expression. The dependence of the protein synthesis rate on mRNA content is included by assuming proportionality between mRNA and DNA over a wide range of growth rates and by multiplying the synthesis rate expression for proteins by the level of the appropriate DNA pool. Thus, the synthesis rate of foreign protein is proportional to the plasmid content rather than the mRNA transcribed from the plasmids. Similarly, the endogenous protein synthesis is proportional to the level of the gene pool, G. This is clearly a simplification of the known mechanistic details, resulting from our efforts to retain a minimal number of constituent pools. Turnover of the protein pools was included by first-order decay expressions. The equations describing both the foreign and endogenous protein pools are identical in form and differ only by parameter values. Secretion mechanisms were not included here but are being considered in a further work to model microorganisms which secrete products. Chromosomal and plasmid DNA synthesis depends upon the nucleotide pool level. Details are included below in the discussion of model dimensionality.

The RNA pool consists of that rRNA actively involved in protein synthesis. The synthesis of the rRNA pool depends on the amino acid level in addition to the nucleotide level. In this way, the role of amino acids as effector molecules through ppGpp[28] was included. Turnover of rRNA has been observed to vary with the growth rate of the cells. When cells were starved for substrate, the rate of synthesis was observed to be seventy percent higher than the accumulation. During other times, the rate of synthesis was observed to be only ten percent higher than the accumulation.[29] The difference is the rRNA turnover. The form

Table I. Synthesis rate expressions[a] for constituent pools.

$$\left[\frac{dA}{dt}\right]_s = k_1\left[\frac{K_A}{K_A + A}\right]\left[\frac{S}{K_{AS} + S}\right]\left[\frac{A}{K_{2A} + A}\right] \quad (1)$$

$$\left[\frac{dN}{dt}\right]_s = k_2\left[\frac{K_N}{K_N + N}\right]\left[\frac{A}{K_{NA} + A}\right]\left[\frac{S}{K_{NS} + S}\right] \quad (2)$$

$$\left[\frac{dP}{dt}\right]_s = \mu_1\left[\frac{A}{K_{PA} + A}\right]RG - K_{TP}P \quad (3)$$

$$\left[\frac{dP_f}{dt}\right]_s = \mu_4\left[\frac{A}{K_{P_fA} + A}\right]RG_f - K_{TP}P_f \quad (4)$$

$$\left[\frac{dG}{dt}\right]_s = \mu_2\left[\frac{N}{K_{GN} + N}\right] \quad (5)$$

$$\left[\frac{dG_f}{dt}\right]_s = \mu_5\left[\frac{N}{K_{G_fN} + N}\right] \quad (6)$$

$$\left[\frac{dL}{dt}\right]_s = \mu_3\left[\frac{S}{K_{LS} + S}\right]\left[\frac{A}{K_{LA} + A}\right] \quad (7)$$

$$\left[\frac{dR}{dt}\right]_s = \mu_6\left[\frac{N}{K_{RN} + N}\right]\left[\frac{A}{K_{RA} + A}\right]^G - K_{TR}R - K'_{TR}R\left[\frac{K_{TR_S}}{K_{TR_S} + S}\right] \quad (8)$$

[a] These expressions describe the synthesis of each constituent pool. Constituent synthesis is equal to the specific rate of synthesis, $\hat{r}_M$, where $M = A, N, P, P_f, G, G_f, L, R$, minus the rate of turnover. Turnover is included for both protein pools and the rRNA pool.

of our decay expressions was based on these observations.

The lipid pool consists of all remaining cellular membrane material and precursors. There are two major reactant groups contributing to lipid formation: first, fatty acids and glycerol which are represented by the external substrate pool; and second, serine, which is represented by the amino acid pool. Thus, the synthesis of lipids is influenced by both reactant groups. We have represented this by the multiplication of two saturation kinetic expressions. There are no significant depletions of this pool in our model.

Model Dimensionality

The level of each of the internal constituent pools is expressed in our model equations as mass fraction or gram constituent per gram dry cell mass. Since our equation dimensionality is not consistent with the majority of biochemical engineering literature, a detailed explanation of one differential equation follows. The synthesis rate expression for chromosomal DNA in a batch culture may be expressed as

$$\frac{d(GX)}{dt} = \mu_{G,\max}\left[\frac{N}{K_N + N}\right]X \quad (1)$$

where G is the mass fraction of DNA (g/g dry cell mass); $\mu_{G,\max}$ is the maximum replication rate (g/g dry cell mass, h); N is the nucleotide mass fraction, K_N is the saturation constant (g N/g dry cell mass); and X is the cell mass per unit volume, such that (GX) is the concentration (per culture volume) of chromosomal DNA. By invoking the chain rule on the left-hand side of eq. (1) and substituting in

$$\frac{dX}{dt} = \mu X \quad (2)$$

where μ is the instantaneous specific growth rate, the following result is obtained:

$$\frac{dG}{dt} = \mu_{G,\max}\left[\frac{N}{K_N + N}\right] - \mu G \quad (3)$$

The term, $-\mu G$, accounts for the expansion of the biomaterial and must be included for structured models employing *intrinsic* variables, as previously shown by Fredrickson.[30] Equation (3) has been represented by two entries, one each in Tables I and II. Table I lists the specific net synthesis rates for each constituent pool or $\mu_{G,\max}\,[N/K_N + N]$ in the above equation. Table II contains equations describing the overall dynamic response for each constituent pool. Table III is a list of all the numerical values for the stoichiometric coefficients, maximum synthesis rates, decay constants, and saturation constants. Appropriate references for the sources of these values are included. It is important to note that there are no adjustable parameters in this model. The parameters which govern the plasmid replication and foreign protein expression are

Table II. Dynamic equations for constituent pools.

$$\frac{dA}{dt} = \left[\frac{dA}{dt}\right]_s - \varepsilon_1\left[\frac{dN}{dt}\right]_s - \varepsilon_2\left[\frac{dL}{dt}\right]_s - \gamma_1\left[\frac{dP}{dt}\right]_s - \gamma_1\left[\frac{dP_f}{dt}\right]_s - \mu A \tag{1}$$

$$\frac{dN}{dt} = \left[\frac{dN}{dt}\right]_s - \gamma_2\left[\frac{dG}{dt}\right]_s - \gamma_2\left[\frac{dG_f}{dt}\right]_s - \gamma_2\left[\frac{dR}{dt}\right]_s - \mu N \tag{2}$$

$$\frac{dP}{dt} = \left[\frac{dP}{dt}\right]_s - \mu P \tag{3}$$

$$\frac{dP_f}{dt} = \left[\frac{dP_f}{dt}\right]_s - \mu P_f \tag{4}$$

$$\frac{dG}{dt} = \left[\frac{dG}{dt}\right]_s - \mu G \tag{5}$$

$$\frac{dG_f}{dt} = \left[\frac{dG_f}{dt}\right]_s - \mu G_f \tag{6}$$

$$\frac{dL}{dt} = \left[\frac{dL}{dt}\right]_s - \mu L \tag{7}$$

$$\frac{dR}{dt} = \left[\frac{dR}{dt}\right]_s - \mu R \tag{8}$$

Table III. Maximum rate, stoichiometric, and saturation constants.

Amino Acids					
k_1	1.69 hr^{-1}	Material Balance	K_A	0.125	6 x T.M.F.A., Ingraham,[28]
ε_1	$0.5485\frac{gA}{gN}$	Stryer,[27]	K_{AS}	0.01 g/l	Monod Constant
ε_2	$0.1418\frac{gA}{gL}$	Stryer,[27]	K_{2A}	0.001	"1/25 rule" [39]
γ_1	$1.167\frac{gA}{gP}$	Stryer,[27]			
Nucleotides					
k_2	1.19 hr^{-1}	Jensen,[40]	K_N	0.125	5 x T.M.F.N., Ingraham,[28]
γ_2	$1.056\frac{gN}{gG}$	Stryer,[27]	K_{NS}	0.01 g/l	Monod Constant
			K_{NA}	0.026	sensitive to A, T.M.F.N.,[28]
Protein					
μ_1	$154.4\frac{gm}{ggRgGhr}$	Ingraham *et al.*,[28]	K_{PA}	0.002	1/10 x T.M.F.A.
			K_{TP}	0.08 hr^{-1}	5% turnover,[28]
DNA					
μ_2	0.078 hr^{-1}	Ingraham *et al.*,[28]	K_{GN}	0.01	1/2 x T.M.F.N. (sens. to N)
Lipids					
μ_3	0.52 hr^{-1}	Shuler,[18]	K_{LS}	0.01 g/l	Monod Constant
			K_{LA}	0.026	(sens. to A) T.M.F.M.,[28]
rRNA					
μ_6	$19.64\frac{g}{ggGhr}$	Kjeldgaard *et al.*,[29] ,[28]	K_{RN}	0.026	T.M.F.N.
K_{TR}	0.147 hr^{-1}	10% turnover, high S,[28]	K_{TR_S}	0.01 g/l	Monod Constant
K'_{TR}	0.179 hr^{-1}	70% turnover, low S,[28]	K_{RA}	0.001	1/25 x T.M.F.N
Protein (For.)			*Plasmids*		
μ_4	1000	to be determined	μ_5	0.002	to be determined
K_{P_fA}	0.001	to be determined	K_{G_fN}	0.001	to be determined

to be determined for each replication locus and promotor/operator region of experimental interest.

Calculation of Growth Rate

Since the entire cell mass is divided into eight lumped constituent pools, the sum of all their fraction mass levels adds to unity at all times.

$$[A + N + P + P_f + G + G_f + L + R] = 1.0 \quad (4)$$

Furthermore, the time derivative of this sum is zero at all times.

$$\frac{d}{dt}[A + N + P + P_f + G + G_f + L + R] = 0 \quad (5)$$

These two results simplify growth rate calculations.

The summation of all eight equations in Table II yields:

$$\sum_{i=1}^{8} \frac{dC_i}{dt} = \sum_{i=1}^{8} \sum_{j=1}^{8} r_{ij} - \mu\left[\sum_{i=1}^{8} C_i\right] \quad (6)$$

where C_i denotes each constituent pool: A, N, P, P_f, G, G_f, L, and R, and r_{ij} corresponds to the synthesis and depletion terms shown in Table II. By substituting eqs. (4) and (5) into eq. (6), it can be readily shown that:

$$\mu = \sum_{i=1}^{8} \sum_{j=1}^{8} r_{ij} \quad (7)$$

This simple equation for calculating the instantaneous specific growth rate was first derived in general form by Fredrickson.[30] Thus, the instantaneous specific growth rate, which is calculated using the synthesis rate expressions (Table I) and stoichiometric coefficients (Table III), is given by:

$$\mu = \left[\frac{dA}{dt}\right]_s + (1 - \varepsilon_1)\left[\frac{dN}{dt}\right]_s + (1 - \varepsilon_2)\left[\frac{dL}{dt}\right]_s + (1 - \gamma_1)\left\{\left[\frac{dP}{dt}\right]_s + \left[\frac{dP_f}{dt}\right]_s\right\} + (1 - \gamma_2)\left\{\left[\frac{dG}{dt}\right]_s + \left[\frac{dG_f}{dt}\right]_s + \left[\frac{dR}{dt}\right]_s\right\} \quad (8)$$

The concept of fractional mass units was employed for a corrected form of Williams' two compartment model[31] by Bailey and Ollis.[32] To our knowledge, this was the only attempt at using this powerful approach for the structured modeling of biological systems. This article represents the first attempt at characterizing bacterial growth using a detailed metabolic and biochemically consistent structure within this general modeling framework. It is our aim to show that this simple technique can be put to great use in understanding the growth behavior of a variety of biological systems. While in this article, we restrict our attention to the growth metabolism and rate expressions for a recombinant bacterial culture, we have applied this methodology to modeling hybridoma growth dynamics, which will be published subsequently.[33]

Reactor Dynamics

The equations listed in Table II describe the kinetics of all intracellular constituents of the recombinant bacteria and consequently determine the instantaneous specific growth rate. This set of kinetic equations was incorporated into both batch and continuous flow stirred tank reactor (CFSTR) material balances. The following additional equations are required to describe batch reactor dynamics:

$$\frac{dS}{dt} = -\frac{1}{Y_s}\frac{dX}{dt} \quad (9)$$

$$\frac{dX}{dt} = \mu X$$

The cell mass, X, is in units of grams dry wt/L and the yield coefficient, Y_s, is in units of cell mass/g substrate. The calculated growth rate, μ, is of dimension h^{-1}. Chemostat dynamics are described by adding the following equations to those in Table II:

$$\frac{dS}{dt} = -\frac{1}{Y_s}\mu X + D(S_f - S) \quad (10)$$

$$\frac{dX}{dt} = \mu X - DX \quad (11)$$

In these equations, the dilution rate, D, and substrate feed concentration, S_f, are in units h^{-1} and g/L respectively.

Each of the equations in Table II remains unchanged for the chemostat dynamics as shown below. For chemostat operation with sterile feed, eq. (1) should be modified to:

$$\frac{d(GX)}{dt} = \mu_{G,\max}\left[\frac{N}{K_N + N}\right]X - D(GX) \quad (12)$$

By expanding the left-hand side of this equation and simplifying by the substitution of eq. (11) we have, as before:

$$\frac{dG}{dt} = \mu_{G,\max}\left[\frac{N}{K_N + N}\right] - \mu G$$

RESULTS AND DISCUSSION

Escherichia coli Cellular Composition

While the primary objective of this modeling effort is to quantify the individual effects of the additional metabolic burdens on the growth rate of the recombinant bacteria and their contributions to plasmid instability, the fundamental validity of this modeling approach is established first by comparing the simulations of the plasmid-free model with the available experimental data. The chemostat model was simulated with zero values for the parameters μ_4 and μ_5 to show how the intracellular composition of the plasmid-free *E. coli* varies as a function of growth rate. Model simulations of the major components rRNA, DNA, and protein are shown in Figure 2 along with the experimental data from Jacobsen.[28] Chemostat data and balanced growth batch data were superimposed and illustrate no disparities.

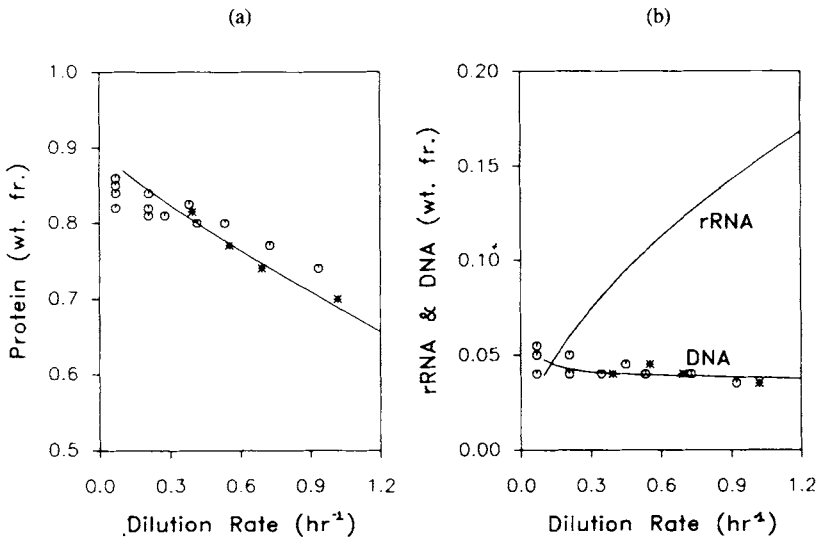

Figure 2. Cellular composition: (a) protein content as weight fraction of total dry cell mass in *Escherichia coli;* (b) weight fraction DNA and ribosomal RNA. Data from ref. 28; ovals and asterisks represent chemostat and balanced batch growth data, respectively; solid lines are model simulations.

The model simulations for protein and DNA contents agree very well with the experimental data. Ribosomal RNA data were not given by Jacobsen but could be estimated from the tRNA and total RNA data with an assumption about mRNA content. Data indicate an almost linearly increasing trend with growth rate, however the estimated values were higher than the model predictions by ca. 20% at low dilution rates (less than 0.2 h^{-1}). At higher dilution rates, model predictions agree to within 10%. Since ca. 15% rRNA is inactive and not involved in protein synthesis, these predictions are within 8% of the experimental estimates for active rRNA.

Growth Rate: Plasmid-Free Versus Recombinant Cultures

Using different values for the parameters μ_4 and μ_5, the chemostat model was simulated to determine the effect of plasmid copy number and foreign protein content on the growth rate. The calculated growth rate is shown as a function of the steady state substrate concentration in Figure 3. The uppermost line depicts the simulation results of a plasmid-free cell culture. Model predictions for a culture with an average of 50 plasmids (5.0 kb)/genome equivalent and no foreign protein expression ($\mu_4 = 0.0$) were not significantly different than the plasmid-free culture results. In fact, the results appear superimposed as one line, indicating that the bacterial growth rate is not significantly reduced by replication of ca. 50 plasmids/genome equivalent.

By varying the rate constant for foreign protein expression (or the promoter affinity for RNA-polymerase), μ_4, we can predict the metabolic stress placed on the cells from different levels of foreign protein expression. The constant μ_4 is promoter specific and needs to be determined for each experimental system. The bottom curves in Figure 3 are the model simulations with μ_4 equal to 1×10^3 and 2×10^3 h^{-1}. The corresponding foreign protein contents were 15 and 25% total protein content, respectively, at high substrate concentrations.

The same simulation results are plotted in double-reciprocal form in the inset figure which demonstrates that the growth rate differential between the plasmid-free host and the recombinant strain is not constant for all substrate concentrations. A comparison is also made to the Monod model. The deviation from the Monod model is similar to that for the more detailed Cornell model which was shown to agree with experimental results.[20]

These results indicate that increasing the foreign protein expression rate constant increases the foreign protein content and reduces the specific growth rate of the recombinant cell. Furthermore, overexpression of foreign protein reduces the bacterial growth rate much more significantly than the maintenance of a 50 copy/genome equivalent plasmid.

Dependence of rDNA Replication and Translation on Environmental Conditions

The model simulations of the variation in steady state plasmid copy number as a function of growth rate are shown in Figure 4. As mentioned previously, several experimental studies have indicated that there seems to be a distinct increasing trend in plasmid copy number (per genome equivalent) with decreasing growth rate. Our model contains two constants which dictate the plasmid synthesis response to environmental changes. The first, μ_5, is the maximum replication rate in units g plasmid/g cell mass h. The second, K_{G_fN}, is the saturation constant describing the dependence of plasmid replication on nucleotide content. Neither of these constants are known explic-

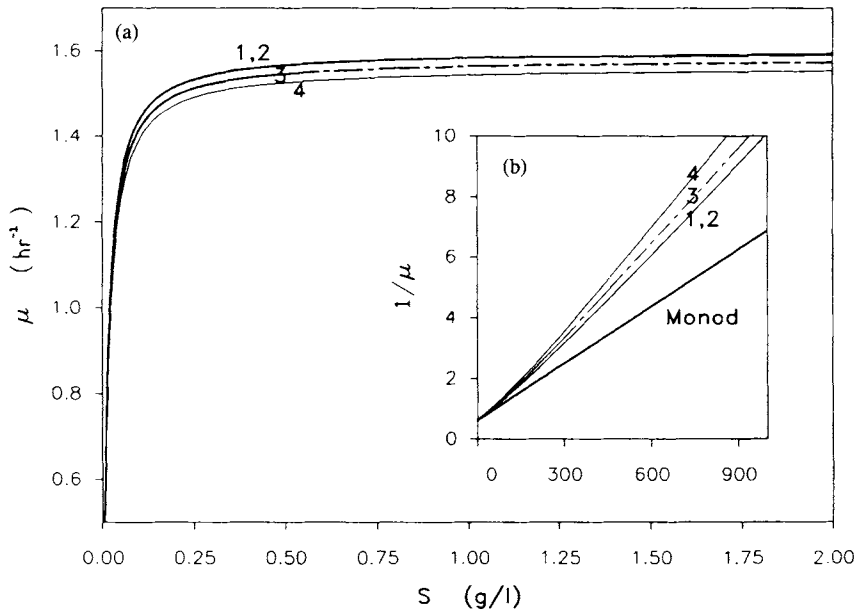

Figure 3. Growth rate dependence on substrate concentration. (a) Curve 1 represents the plasmid-free host: $\mu_4 = 0\ h^{-1}$, $\mu_5 = 0\ h^{-1}$. Curve 2 depicts simulation of plasmid-containing cell with no foreign protein expression: $\mu_4 = 0\ h^{-1}$, $\mu_5 = 0.0018\ h^{-1}$. Curves 1 and 2 appear superimposed. Curves 3 and 4 illustrate effect of increasing foreign protein expression with: $\mu_4 = 1000\ h^{-1}$, $\mu_5 = 0.0018\ h^{-1}$, and $\mu_4 = 2000\ h^{-1}$, $\mu_5 = 0.0018\ h^{-1}$, respectively. Plasmid copy number (per genome equivalent) is ca. 50 in curves 2, 3, and 4. Foreign protein levels are ca. 15 and 25% in curves 3 and 4, respectively. Inset graph, (b), illustrates deviation from Monod model with K_s = 0.01 g/L and $\mu_{max} = 1.596\ h^{-1}$.

itly since ours is their first application. Their values must be estimated from experimental data.

If those enzymes which replicate plasmids compete with equal vigor for nucleotides as those which replicate the host cell genome, then K_{G_fN} should be identical to K_{GN}, the analogous genome constant.

The dimension of the model's intrinsic variables also facilitates simple calculation of plasmid copy number, which was expressed per genome equivalent. There appears to be some variation among experimental reports regarding the units for plasmid content (per genome equivalent, per cell mass, per cell, per RNA, etc.). Since our objective is to characterize the effect of plasmid replication on the cell's normal metabolic activity, we need to express plasmid content in a manner which signifies the "burdensome quality" of the plasmids. Plasmid replication interferes with genome replication, thus by expressing plasmid content in plasmids per genome equivalent we examine the relative significance of plasmid activity to the cell.

Thus in Figure 4, we have taken the data from Seo and Bailey[9] and Siegel and Ryu[10] and expressed their results as number of plasmids per genome equivalent. Like Seo and Bailey, we assumed that chromosomal content varied with growth rate according to the Cooper–Helmstetter model.[34] The values of μ_5 and K_{G_fN} which best represent the data are 1.8×10^{-3} g plasmid/g cell mass h and 10^{-6} g G_f/g cell mass for Seo and Bailey's pDM247 (5.85 MD) and 1.07×10^{-3} g/g h and 7.5×10^{-4} g/g for Siegel and Ryu's pPLc23*trpAl* (6.5 kb). It is apparent that the model accurately describes the plasmid content in these systems. Also, there seems to be no inconsistency between chemostat and batch data, since the data of Seo and Bailey are from balanced phase batch cultures whereas the Siegel and Ryu data are derived from chemostat cultures and their trends agree qualitatively. In a further study using chemostat cultures, Seo and Bailey[12] confirmed at least qualitative agreement between batch and chemostat data. They have also suggested that a dilution rate exists where the plasmid content reaches a maximum. However, it is not clear whether this observed maximum is reproducible and part of a consistent trend. They have not provided an explanation for this occurrence. Our model simulations, based on the rate expressions shown here, exhibit no maximum in plasmid content.

The potential capabilities of the model, exemplified by successful predictions of plasmid content at various dilution rates, are further demonstrated by predictions of the corresponding product or foreign protein content. Experimental evidence confirms a trend similar to that of the plasmid copy number: increasing foreign protein content with decreasing growth or dilution rate.[5,9,10,35] Exceptions include the work of Koizumi and co-workers as previously discussed,[16] and Seo and Bailey,[12] who found a large qualitative discrepancy between batch and chemostat data. Only after a dramatic maximum was there a drop in foreign protein content with increasing dilution rate.

Model predictions depicted in Figure 5 illustrate that the highest foreign protein expression also occurs at low dilution rates. The model predictions of Peretti and Bailey[22] exhibit a similar qualitative trend. Simulations of the rate expressions utilized in this work can not explain the re-

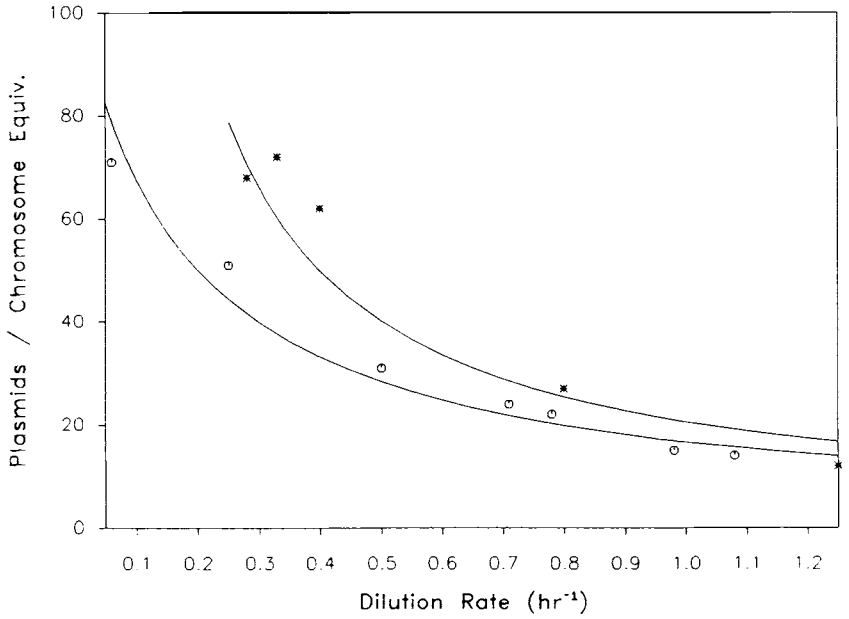

Figure 4. Plasmid dependence on growth rate. Experimental data included as ovals are from ref. 10 and asterisks from ref. 9. Solid lines are model simulations with $\mu_5 = 1.07 \times 10^{-3}$ h^{-1} and $K_{G_fN} = 7.5 \times 10^{-4}$ g/g cell mass, $\mu_5 = 1.8 \times 10^{-3}$ h^{-1} and $K_{G_fN} = 1 \times 10^{-6}$ g/g cell mass, respectively.

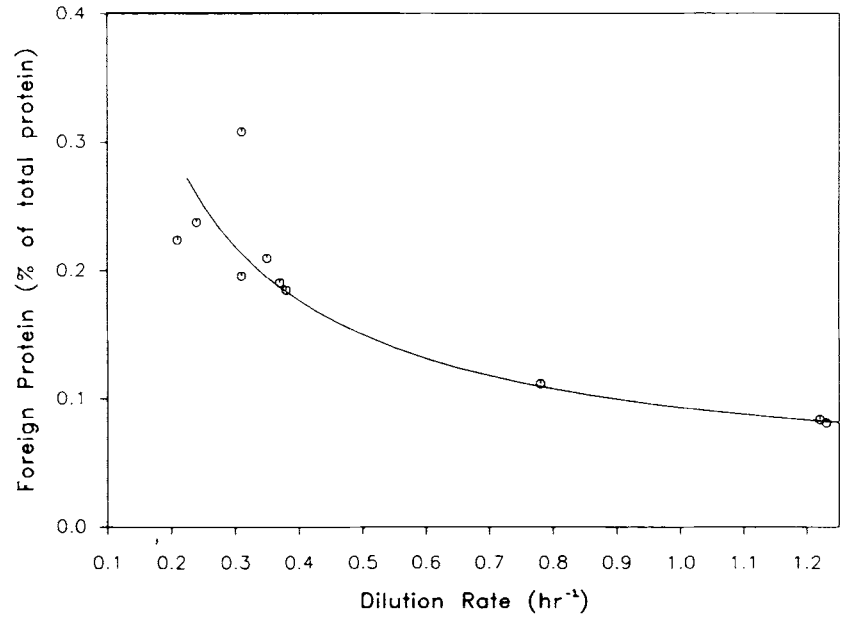

Figure 5. Foreign protein dependence on growth rate. Experimental data included as ovals are from ref. 9; the solid line is the model simulation with $\mu_4 = 4.5$ h^{-1} and $K_{P_fA} = 0.008$ g/g cell mass.

verse trend of Koizumi or the maximum in foreign protein content found by Seo and Bailey. Our foreign protein synthesis expression contains a maximum rate, μ_4, and a saturation constant, K_{P_fA}, which characterizes the translational rate dependence on the amino acid level. These constants must be determined for each foreign protein gene and promoter region. We found that with $\mu_4 = 4.5$ g and $K_{P_fA} = 0.008$ g/g our model predictions matched the batch experimental results of Seo and Bailey.[9] Their data were converted to weight percent with the assumption that 3500 U is equal to ca. 1.0 mg active β-lactamase. These two parameters may, in general, be functions of some other factors, such as temperature or inducer concentration, which can be manipulated in experiments to control the foreign protein expression.

Batch dynamic simulations of purely recombinant cultures illustrate that the plasmid copy number and foreign protein content in the cells increase during the stationary

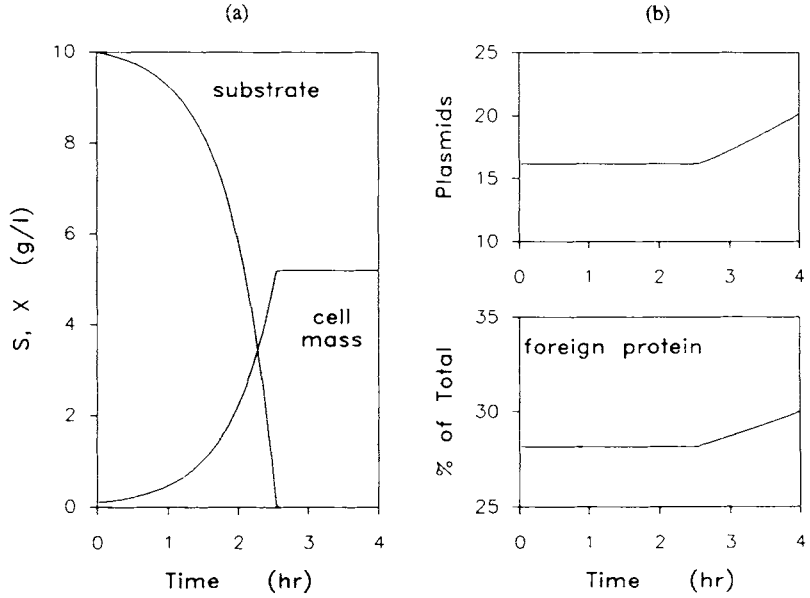

Figure 6. Batch culture profile: (a) substrate concentration (g/L) and cell mass (g/L) simulation versus time; (b) plasmid copy number (per genome equivalent) and foreign protein content (g foreign protein/g total protein) simulations increase during the stationary phase of a batch culture; $\mu_4 = 1500\ h^{-1}$, $\mu_5 = 0.002\ h^{-1}$, $K_{G_fN} = 1 \times 10^{-6}$ g/g cell mass, and $K_{P_fA} = 0.001$ g/g cell mass.

phase of the batch culture (Fig. 6). This was experimentally observed by Stueber and Bujard[14] with *E. coli* (C 600 r^- Δ lac M15:pBR322). An increase in the metabolic burden would result in a larger growth rate difference between the recombinant and plasmid-free host populations, which could lead to a higher "plasmid instability" in the post exponential phases of a batch culture. This phenomenon was experimentally observed by Pinches and co-workers[13] during the growth of *Bacillus subtilis*.

Mixed Populations: Stability of Chemostat Cultures

Among the most significant contributions of this work to the modeling of biological systems is its utility in predicting bioreactor dynamics and the ease with which these predictions can be made. We have predicted the culture stability of a chemostat containing two competing populations: the plasmid-free host and the plasmid-bearing recombinant strain (Fig. 7). In these simulations, we assumed that each population could be represented by an "average" cell. Simulations were performed by coupling two chemostat models, one simulating the plasmid-free host population ($\mu_4 = \mu_5 = 0$) and the other simulating the plasmid-bearing recombinant population. The initial population mixture was 99% recombinant cells and 1% plasmid-free cells. An assumption was made that no recombinant cells lost their ability to synthesize foreign protein (i.e., no structural or segregational instability). Hence, the population composition was strictly dependent upon the growth rate difference between plasmid-free and recombinant strains.

The fraction of recombinant cells remaining in the chemostat continually dropped due to the growth rate difference between cell types. Experimental studies describing chemostat instability found that it was greater at low dilution rates (Ataai and Shuler,[19] Godwin and Slater,[4] and Noack *et al.*[36]). In these studies, the inoculum was 100% recombinant strain, thus the initial appearance of plasmid-free cells was definitely due to improper segregation. However, the rapid takeover by the plasmid-free population is the result of two separable contributions: improper segregation and the dynamic growth rate differential between the competing populations. It is important to distinguish between these two effects in the analysis of plasmid instability in recombinant bacterial cultures. Our simulations, even with the assumption of no segregational instability, predict that a chemostat is more stable at higher dilution rates which is in good qualitative agreement with these experimental results. Future incorporation of plasmid segregation kinetics into the model framework presented here will provide a basis for a more quantitative description of recombinant culture instabilities.

Watson and co-workers[37] showed that the rate of fall in the proportion of plasmid-containing cells was independent of the growth rate since their data for different dilution rates collapsed to a single curve when plotted as a function of time rather than generation number. They concluded that the plasmid loss could not be ascribed to faulty partitioning during cell division, which would suggest that the observed instability was due to the growth rate differential between competing strains. Figure 8 shows our model simulation for the fall in the proportion of plasmid-containing

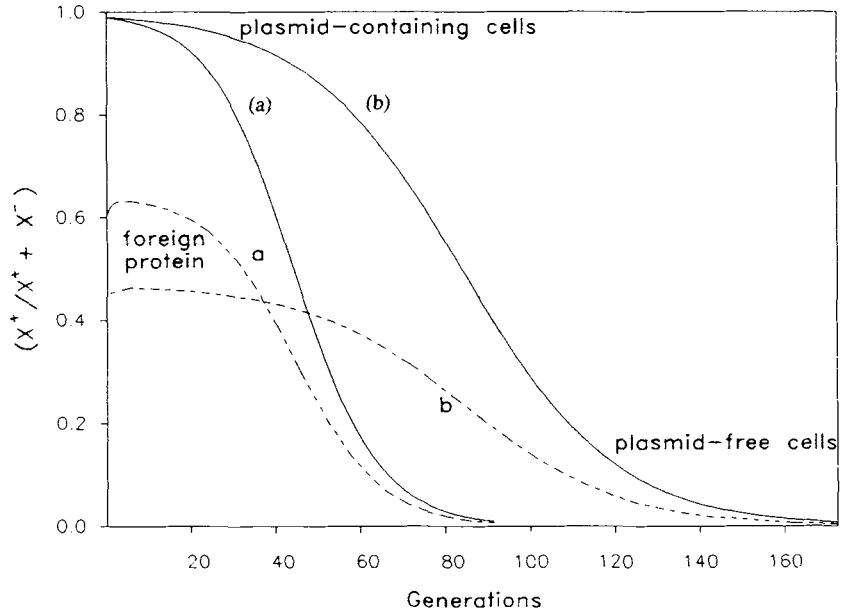

Figure 7. Chemostat stability profiles. Bioreactor stability (generations until washout of recombinant strain) is predicted greater at higher dilution rates (a = 0.5 h^{-1}, b = 1.1 h^{-1}). Product concentration is g foreign protein/g total protein in reactor. An assumption of no segregation was made, hence population dynamics depend solely on the varying growth rate differential between the plasmid-free and recombinant strains; μ_4 = 500 h^{-1} and μ_5 = 0.01 h^{-1}.

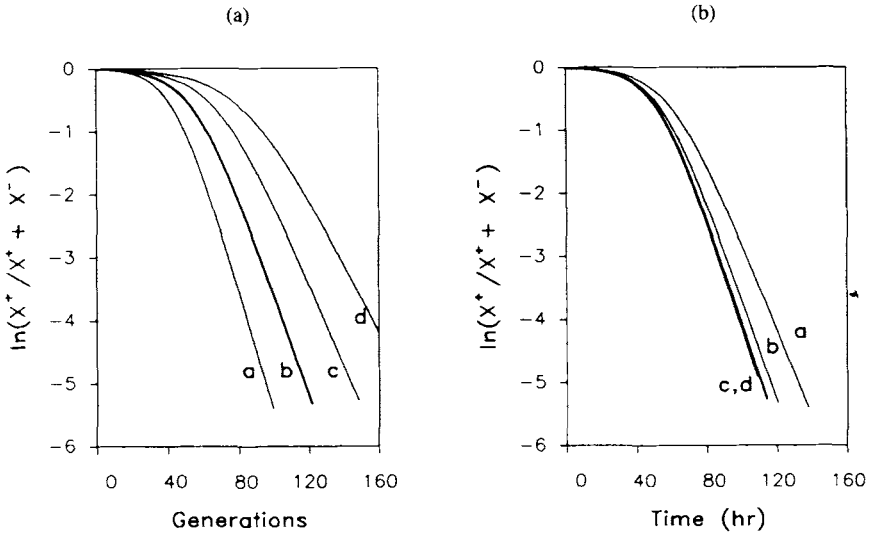

Figure 8. Chemostat population dynamics: (a) simulations similar to those in Figure 7 are depicted with dilution rates a, b, c, and d equal to 0.5, 0.7, 0.9, and 1.1 h^{-1}, respectively. (b) data collapse to one band when plotted on versus time scale, illustrating limited agreement with ref. 37. Segregation was assumed nonexistent.

cells as both a function of generation number and time for a range of dilution rates (0.5–1.1 h^{-1}). Implicit in this simulation is the assumption of a constant (zero) specific segregation rate. Thus, our results are comparable to those of Watson and co-workers. Note, however, in experimental studies concerning "plasmid" or "culture" instability the difference in growth rate between competing strains is not necessarily a constant over the course of a "steady-state" run at any fixed dilution rate.

This is made apparent by considering our simulation depicted in Figure 9. As the experiment progresses, the population mixture shifts toward the faster growing population. Thus the substrate concentration will shift gradually from an initial value of S^+ (corresponding to $\mu^+ = D$) to a

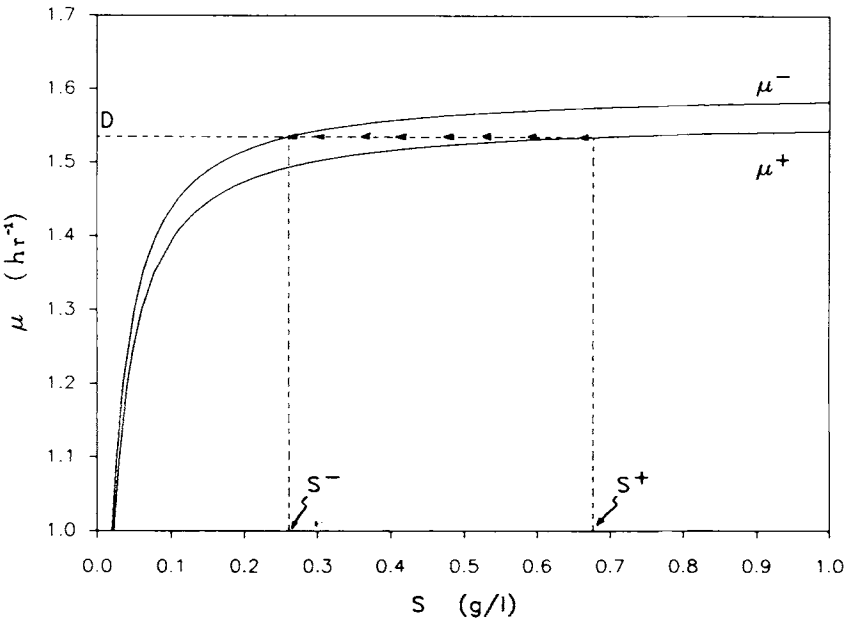

Figure 9. Substrate path in chemostat with mixed population. Substrate concentration, initially S^+, corresponding to a population of all recombinant cells, gradually shifts to S^-, the steady state substrate concentration for the plasmid-free cells, as the population mixture shifts to the plasmid-free (faster growing) cells. The replication and translation constants were the same as in Figure 3 (curves 1 and 4).

final steady-state value of S^- (corresponding to $\mu^- = D$). This results in a *dynamic* growth rate differential between cell types. Simulations presented here incorporate this dynamic behavior and illustrate the utility of this model in describing population dynamics.

NOMENCLATURE

A	amino acids
D	dilution rate
G, G_f	chromosomal and plasmid DNA, respectively
k_i	maximum synthesis rate (i = 1 for amino acids; i = 2 for nucleotides
$K_{i,j}$	saturation constant for dependence of constituent i on constituent j
K_s	Monod constant
K_{Ti}	degradation coefficient for constituent i
K'_{Ti}	substrate dependent degradation coefficient for rRNA
L	lipids and cell membrane precursors
N	nucleotides
P, P_f	endogenous and foreign protein, respectively
R	ribosomal RNA
r_{ij}	rate of synthesis of i or depletion of i by constituent j
S, S_f	substrate and feed substrate, respectively
$T.M.F.$	typical mass fraction
X	cell mass
$[dC_i/dt]_s$	rate of synthesis for constituent C_i
ε_i	stoichiometric coefficient for depletion of amino acids by nucleotides (i = 1) and lipids (i = 2)
γ_1	stoichiometric coefficient for depletion of amino acids by protein
γ_2	stoichiometric coefficient for depletion of nucleotides by RNA, DNA
μ, μ^-, μ^+	instantaneous specific growth rate; plasmid-free, and plasmid-bearing, respectively
$\mu_{i,\max}$	maximum rate of synthesis for constituent i

The authors wish to acknowledge valuable discussions with Professor Robert H. Davis and the support of this work by the Amoco Foundation Engineering Faculty Grant (to R. H. D.) and the National Science Foundation (grant No. ECE-8611305).

References

1. H. Tsunekawa, M. Tateishi, T. Imanaka, and S. Aiba, *J. Gen. Microbiol.*, **127,** 93 (1981).
2. T. Imanaka, H. Tsunekawa, and S. Aiba, *J. Gen. Microbiol.*, **118,** 253 (1980).
3. U. M. Nordstrom, B. Engberg, and K. Nordstrom, *Mol. Gen. Genet.*, **135,** 185 (1974).
4. D. Godwin and J. H. Slater, *J. Gen. Microbiol.*, **111,** 201 (1979).
5. S. H. Kim and D. D. Y. Ryu, *Biotechnol. Bioeng.*, **26,** 497 (1984).
6. T. Nakazawa, *J. Bacteriol.*, **133(2),** 527 (1978).
7. J. Pierce and S. Gutteridge, *Appl. Environ. Microbiol.*, **49(3),** 1094 (1985).
8. T. Imanaka and S. Aiba, *Ann. NY Acad. Sci.*, **369,** 1 (1981).
9. J. Seo and J. E. Bailey, *Biotechnol. Bioeng.*, **27,** 1668 (1985).
10. R. Siegel and D. D. Y. Ryu, *Biotechnol. Bioeng.*, **27,** 28 (1985).
11. B. Engberg and K. Nordstrom, *J. Bacteriol.* **123(1),** 179 (1975).
12. J. H. Seo and J. E. Bailey, *Biotechnol. Bioeng.*, **28,** 1590 (1986).
13. A. Pinches, M. E. Louw, and T. G. Watson, *Biotechnol. Lett.*, **7(9),** 621 (1985).
14. D. Stueber and H. Bujard, *EMBO J.*, **1(11),** 1399 (1982).
15. S. Lin-Chao and H. Bremer, *Mol. Gen. Genet.*, **203,** 143 (1986).
16. J. Koizumi, Y. Monden, and S. Aiba, *Biotechnol. Bioeng.*, **27,** 721 (1985).
17. S. B. Lee and J. E. Bailey, *Biotechnol. Bioeng.*, **27,** 1699 (1985).
18. M. L. Shuler, S. Leung, and C. C. Dick, *Ann. NY Acad. Sci.*, **326,** 35 (1979).
19. M. M. Ataai and M. L. Shuler, *Biotechnol. Bioeng.*, **30,** 389 (1987).
20. M. M. Domach, S. K. Leung, R. E. Cahn, G. G. Cocks, and M. L. Shuler, *Biotechnol. Bioeng.*, **27,** 203 (1984).

21. S. W. Peretti and J. E. Bailey, *Biotechnol. Bioeng.*, **28,** 1672 (1986).
22. S. W. Peretti and J. E. Bailey, *Biotechnol. Bioeng.*, **29,** 316 (1987).
23. M. M. Ataai and M. L. Shuler, *Biotechnol. Bioeng.*, **27,** 1051 (1985).
24. M. L. Shuler, *Chem. Eng. Commun.*, **36,** 161 (1985).
25. S. Aiba, M. Shoda, and M. Nagatani, *Biotechnol. Bioeng.*, **10,** 845 (1968).
26. J. F. Andrews, *Biotechnol. Bioeng.*, **10,** 707 (1968).
27. L. Stryer, *Biochemistry,* (Freeman, New York, 1981).
28. J. L. Ingraham, O. Maaloe, and F. C. Neidhardt, *Growth of the Bacterial Cell.*, (Sinauer Assoc. Sunderland, MA, 1983).
29. N. O. Kjeldgaard and K. Gausing, "Control of Ribosome Synthesis," in *Ribosomes,* M. Nomura, A. Tissieres, and P. Lengyel (Eds.) (Cold Spring Harbor Laboratory, NY, 1974).
30. A. G. Fredrickson, *Biotechnol. Bioeng.*, **18,** 1481 (1976).
31. F. M. Williams, *J. Theoret. Biol.*, **15,** 190 (1967).
32. J. E. Bailey and D. F. Ollis, in *Biochemical Engineering Fundamentals,* (McGraw-Hill, Inc., NY, 1986).
33. B. C. Batt and D. S. Kompala, Structured Kinetic Modeling of Hybridoma Growth and Monoclonal Antibody Production in Suspension Cultures (unpublished).
34. S. Cooper and C. E. Helmstetter, *J. Mol. Biol.*, **31,** 519 (1968).
35. R. H. Pritchard, M. G. Chandler, and J. Collins, *Mol. Gen. Genet.*, **38,** 143 (1975).
36. P. Noack, M. Roth, R. Geuther, G. Muller, K. Undisz, C. Hoffmeier, and S. Gaspar, *Mol. Gen. Genet.*, **184,** 121 (1981).
37. T. G. Watson, A. Pinches, and M. E. Louw, *Biotechnol. Lett.*, **8(10),** 687 (1986).
38. E. Remaut, P. Stanssens, and W. Fiers, *Gene,* **15,** 81 (1981).
39. M. M. Domach, "Refinement and use of a Structured Model of a Single Cell of *Escherichia coli* for the Description of Ammonia-Limited Growth and Asynchronous Population Dynamics," Ph.D. thesis, Cornell University, Ithaca, NY, 1983.
40. K. F. Jensen, "Metabolism of 5-phosphoribosyl 1-pyrophosphate (PRPP) in *Escherichia coli* and *Salmonella typhimurium,*" in *Metabolism of Nucleotides, Nucleosides, and Nucleobases in Microorganisms,* (A. Munch-Petersen, Academic, New York, 1983).
41. D. E. Pettijohn, C. R. Kossman, and O. G. Stonington, *J. Mol. Biol.*, **52,** 281 (1982).