T H E   U N I V E R S I T Y   O F   M I C H I G A N
COLLEGE OF LITERATURE, SCIENCE, AND THE ARTS
Computer and Communication Sciences Department

Technical Report

ADAPTIVE BEHAVIOR OF SIMULATED BACTERIAL
CELLS SUBJECTED TO NUTRITIONAL SHIFTS

Erik David Goodman

with assistance from:

February 1972

PREFACE

Some introductory remarks seem to be desirable before presen-
tation of the model is begun. The author would like to describe
the reasons for his interest in the subject of this dissertation
and to thank those who have helped him during its preparation.

Work on the bacterial cell model described in this dissertation
was begun in 1968 by Dr. Roger Weinberg, as part of his post-
doctoral research in the Department of Computer and Communication
Sciences at The University of Michigan. Dr. Weinberg at that time
held a Ph.D. in microbial genetics from the University of Texas,
and was pursuing a second Ph.D. at The University of Michigan. In
talking with Dr. Weinberg at the Logic of Computers Group, where we
both had offices, I became interested in the model which he had
developed; in particular, the idea of building a model of a colony
of bacteria based on his model of a single cell appealed to me.
It seemed that a colony model in which the individual cells were
represented by a fairly extensive metabolic model would make
available for study a tremendous range of important questions
concerned with the interactions of cells with each other and with
their surrounding environment. My background in automata theory
suggested to me that a very powerful model of a colony could be
obtained by embedding Weinberg's model in a formal framework called
a cellular space or tessellation space. This sort of "checkerboard"
model could be realized as a simulation using a computer system avail-
able at the Logic of Computers Group. I developed a preliminary model
of this sort in the Fall of 1969. It soon became clear, however, that

ii

the performance of the colony model (called CELLSPACE) was being severely limited by the shortcomings of the single cell model (called ONECELL). Weinberg had developed the model to examine the necessity of various control mechanisms for obtaining certain behavior over short periods of time (ten seconds). Although he included some of the control mechanisms which are necessary to maintain realistic behavior over longer time periods, many were omitted, and still others were not "balanced" so as to achieve reasonable behavior over long periods of time. Two additional changes in approach were necessary: first, Weinberg's work did not include any mechanism for transport of materials into or out from the cell -- that is, the inputs to the model were particular *internal* concentrations of some biochemicals, and, second, the simulation operated with a time increment of 1 second or less, and this was much too short to allow extended runs of many hours of simulated time, particularly when many cells were concerned. In addition, advances in molecular biology made it possible to specify better models for some portions of the cell's metabolism. For these reasons, a new ONECELL model has been developed, over the period of the last two years. Work on the CELLSPACE model also progressed for a considerable portion of that time; however, the rate of changes to ONECELL has made it impractical to maintain compatibility between the two models.

As my understanding of the ONECELL model has deepened, my goals have changed. It has become clear that the ONECELL model as presently formulated represents an entirely different approach from the original one of Weinberg. During the past year, my effort has been concentrated almost entirely on the achievement of a good ONECELL model. Although

I still hold my original belief that much could be gained from a colony model composed of multiple ONECELL components, I now perceive that the goal of obtaining a reasonable ONECELL model is at least as worthwhile, and of course a prerequisite to the former goal.

helpful beyond the call of duty in preparing diagrams, formating, and typing this dissertation. Miss Monna Whipp has also assisted in the typing.

To the GO players and the musicians who have helped me remain reasonably sane, I am grateful.

My wife Denise has shouldered many burdens to allow me to do this work, and has buoyed my spirits when I have been low. I dedicate this dissertation to her; she has worked for it as much as I have.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

ADAPTIVE BEHAVIOR OF SIMULATED BACTERIAL
CELLS SUBJECTED TO NUTRITIONAL SHIFTS

by
Erik David Goodman

Chairman: Bernard P. Zeigler

The ONECELL simulation is designed to simulate the growth of

a single bacterial cell over many generations. It represents the cell

as a set of forty biochemical pools, interacting with each other and

with the extracellular environment to produce growth in a wide variety

of conditions. Environments modeled include glucose minimal, lactose

minimal, casamino acids, and broth media. Many parameters of the

simulation are set using data for growth in these media from the

microbiological literature; some result from attributing simple

forms to equations describing cellular processes. Extensive use is

made of experimental and theoretical work by Helmstetter, Cooper,

Pierucci, and Revelas (1968) and by Maaløe and Kjeldgaard (1966) on

the organisms *Escherichia coli* and *Salmonella typhimurium*, respectively.

The simulation is designed to permit representation of many cells

in different states, and is a successor to a model used in preliminary

simulations of the development of a bacterial colony (Goodman, Weinberg,

and Laing, 1970). Experiments with the simulation show that it is

capable of withstanding rapid changes of the simulated environment.

As a test of the flexibility of the simulation, instantaneous shifts

among various pairs of the media used for setting the parameters are

tried. The simulation performs well (that is, agrees well with

laboratory observations) in shifts from poor to rich media. In

simulated shifts from rich to poor media, the transient behavior is not as realistic as it is in the other shifts, but the control mechanisms are sufficient to bring the cell to the proper steady-state. A simulated shift to lactose medium demonstrates the capability of the simulation to handle the dynamics of enzyme induction, and to withstand the stresses of temporary deprivation of usable carbon-sources.

CHAPTER 1

Perhaps the first step in discussing a model and a simulation
should be to explain the reasons why the modeling is to be done.
Yates, Brennan, Urquhart, Li, and Halpern (1968) argue that models
serve several purposes. Central among them is the shaping of research
strategy. This is not to say that research should be directed only
toward confirming the predictions of models. In a very real sense,
the process of formulation of the model is the greatest benefit, if the
model is one which attempts to account for complex behavior in terms
of simpler, better-understood processes. The modeling process requires
explicit formulation and examination of ideas which the modeler may
never have solidified otherwise. Models allow examination of the
implications of sets of assumptions, often making evident inconsistencies
between assumptions which, taken individually, look quite reasonable.
The "bookkeeping" function of a model, bring together and codifying
a set of beliefs, is of considerable utility.

Rosen (1968, p. 26) points out that the models of the sort discussed
here are of a different sort than those traditionally employed by
biologists. While traditional biology may use physico-chemical models,
the relational biology (Rashevsky, 1960) which is used here abstracts
the chemical and physical laws from the actual material, and deals
only with them. This allows considerable freedom, but at the price
of difficulty of showing the correspondence of the abstract models
with physical reality.

If the need for a model is accepted, the next question concerns
the need for a simulation. A simulation can be called a realization

1

of the model, i.e., an analog of the modeled system. As such, it
should behave exactly as the modeled system does, to the extent that
the model captures the behavior of the modeled system. If a model
is of a sort that mathematical treatment allows formulation of the
behavior in closed forms, that is, the model's responses to inputs
can be solved for analytically, then a simulation is unnecessary.
However, in the much more common case (for complex models) where such
solution procedures are impossible or impractical, simulation offers
an opportunity to explore the implications of the model by examining
the behavior of the model under various input conditions.

A computer simulation can greatly facilitate the setting of para-
meters, given a model of a certain form. For example, it is possible
to use cumulative or long-term behavior of the real (modeled) system
to set model parameters for components which work on a much shorter
time-scale, since the simulation can be used to obtain the long-term
behavior of the model for comparison with the data on long-term behavior
of the real (modeled) system. Using this technique, parameters which
would be extremely difficult to measure experimentally can be assigned
values.

An early problem which confronts a would-be modeler is the question
what organism(s) should be modeled. At the level at which the ONECELL
model functions, there are, fortunately, many similarities among
organisms, making the possible sources of data for the model somewhat
less restrictive. The principle of the unity of biochemistry, which
asserts this similarity, has been acknowledged for many years to be
important. Of course, it is a general statement, and can be invoked
in specific instances only with considerable care, but as Luria (1960,

p. 4) points out, the principle is exemplified by many chemical results which are achieved by a unique pathway in many widely differing organisms. Specific areas where the units of biochemistry principle is not applicable are not hard to find (see, for example, Sanwal (1970, p. 24)). However, the general validity of the principle makes it possible to "fill in some holes" which might otherwise obstruct the modeling process.

The ONECELL model described in this dissertation relies on data from several similar bacteria. The most important source of information was research with *Escherichia coli*; also important were *Salmonella typhimurium* and *Aerobacter aerogenes*. *E. coli* and *S. typhimurium* are similar, not only in structure, but in rates of change of biochemical composition in rapidly changing environments, as noted for example, by Kjeldgaard (1967), p. 47. Many properties of interest are similar in these organisms, and with suitable scaling, some use has been made of data from all three in the ONECELL model.

The goal of modeling bacterial growth has been pursued by many biologists. There are, of course, a great variety of types and levels of models, each with its strengths and weaknesses. In order that it may be seen where the ONECELL lies in the range of these models, a few will be mentioned, with some of their characteristics. However, a few words about ONECELL are in order first.

ONECELL is a model of a single bacterial cell, perhaps closest in structure and function to *E. coli*. The model is more correctly called a model of an average cell of a particular age (since division) in a culture of cells engaged in exponential growth. If a culture could be completely synchronized, and held in synchrony, without

disturbing the "normal" metabolic processes, then ONECELL would represent the average cell in that synchronized culture. The model does not consider spatial distribution of components within a cell (in agreement, for example, with Fredrickson, Ramkrishna, and Tsuchiya (1967, p. 332)) but it does distinguish one cell from others in the culture which are in a different "metabolic state", having biochemical compositions different from the modeled cell. Of course, by using a number of "copies" of ONECELL, in different states, it is possible to simulate the behavior of an entire culture, and this is done in order to allow comparison of ONECELL with unsynchronized culture data. ONECELL does not keep track of "random" variations between cells in the culture; that is, there is no probability distribution or stochastic process used in the ONECELL model: it is deterministic. Marr, Painter, and Nilson (1969) provide a great deal of data concerning the variations between cells within a culture, but for the sake of simplicity, only culture-averages are used in ONECELL. The model is organized around a number (about 40) of pools of functionally-related biochemicals. Gross characteristics of culture growth, such as μ, the doubling rate, are not elements of the ONECELL model, but are observable only as statistics, and result from complex behavior of the biochemical apparatus. The model has a great variety of types of interactions between pools, rather than restricting all pathways, for example, to a single form of control or influence.

It might be argued at this point (but hopefully not later) that the model is unnecessarily complex. However the alternative--for example, prediction of μ from nutrient supply under a restricted set of circumstances--is of almost no utility except to allow interpolation

or extrapolation of values under the same circumstances. Henrici, as

quoted in Fredrickson, et al. (1967, p. 371) says, "Where these [i.e.,

mathematical analyses of the growth curves of bacteria] are not inter-

preted in terms of organisms, substrate, products of metabolism or

other definite factors, they do not seem to be very helpful to an

understanding of the phenomena... ."

There are great difficulties associated with studying a cell

at the level of 40 or more metabolic pools. The average number of pools

which *directly* influence the rate of change of any pool is five or six.

One of the chief mechanisms which makes such a model somewhat tract-

able is self-regulation which is exhibited by some pools. The mechanisms

of induction, repression and feedback (allosteric) modification

of enzymes, all of which are modeled in the ONECELL model, make it

possible to study certain properties of a subsystem in isolation from

other subsystems of the model. That is, the cell (and the model)

exhibit some stabilizing effects which reduce the sensitivity of

some pools to certain sorts of changes in the rest of the cell. Rosen

(1968, p. 61) makes this point very strongly, and Mesarovic (1968,

p. 73) discusses the importance of this sort of control in yielding

stability within the organism. Without these controls, the "network"

of interactions among pools would be much more difficult to model

(and so it should be, since one would be trying to develop the model

without considering some of the most important processes in the cell,

and this is somewhat akin to fistfighting with one hand tied behind

one's back). Models which have not included these controls have been

formulated (as discussed below), but they could not display much of

the power of the ONECELL model.

Despite the advantages derived from use of repression and feedback inhibition, formulation of a model which exhibits certain desired behavioral characteristics is still difficult. Speaking of his cell simulation, Heinmets (1966, pp. 14-15) says, "It requires a great deal of patience and experience in computer technology to organize a functional system of a model which consists of many entities (31 rate constants)." This statement has been borne out by the ONECELL model, with a much greater number of rate constants. He continues in a clear statement of the modeling problems shared by his model and many portions of ONECELL:

> "In order to obtain a functional cell growth system, the functional entities have to be in the proper quantitative relation to each other as a function of time. This presupposes that rate constants representing turnover and formation of various functional entities should also be in a proper relation to each other. Basically, we are here dealing with the flow system where continuous synthesis and decay of various elements takes place. If some element grows too fast and another too slow, then finally an imbalance will result and the system will be disorganized. As a matter of fact, in a system containing so many rate constants and variables, the most probable state is a disorganized state. Only by selecting specific rate constant values and initial conditions is it possible to obtain a functional system. Obviously there can be a certain amount of deviation in parameters, but those are limited. For example, it is possible to make some functional entity more unstable by letting it decay faster. This increased loss of entity can be compensated by faster synthesis. However, this procedure already introduces an imbalance in the system because increased synthesis of one entity will affect the other systems, since there is a competitive state between various synthetic processes. This means that if only one element is made more unstable, the whole system has to be altered in order to gain proper balance."

The ONECELL simulation is of the "recurrence" type, to use the terminology of Gordon (1969, p. 22). Values of variables at one time step are used in calculating values of variables at the next, and are called lagged variables. One simplifying restriction which reduces

the complexity of the model is that all bacteria modeled are growing

at 37° C. This allows the model to ignore the differing effects of

changing temperature on reactions, the rates of which might for some

rise linearly with T, some with T^2, etc. It is possible under this

restriction to represent complex reactions in fairly simple forms,

as will be discussed in Chapter II.

For setting parameters, the ONECELL model postulates that growth

(in volume) occurs exponentially. In particular, growth rate is assumed

to be primarily a function of size, not age (since division). This is

the experimental finding of Marr, Painter, and Nilson (1969, p. 245).

Ward and Glaser (1971) postulate that the cell volume actually rises

linearly until the cell reaches a certain age, at which the rate

of increase doubles; then volume rises again linearly. They state,

however, (p. 1064) that their data would not allow them to distinguish

linear from exponential growth. The model they support, that of

Donachie and Begg (1970), is one which postulates a certain site which,

when replicated, allows the growth rate to double. Many models have

used this type of discontinuous growth rate, for example, in expressing

the effect of the replication of a particular gene on the synthesis

rate of the mRNA for which it codes. While this is certainly an accurate

representation in most cases, it involves additional bookkeeping which

has not been incorporated into the ONECELL model, which instead assumes

that the rates of RNA formation respond to the total amount of DNA,

not to the number of instances of genes of a particular types. Thus

continuous curves for the rate of production of mRNA's result.

The ONECELL model and simulation can probably be understood more

completely if it is compared and contrasted with some other modeling

efforts.

Eakman, Fredrickson, and Tsuchiya (1966, p. 37) discuss a
classification of mathematical models for description of microbial
populations. They classify models as segregated or distributed,
according to whether or not the population is treated as protoplasmic
biomass distributed uniformly throughout the population's living space.
Further, segregated models are broken down into structured and
unstructured, according to whether or not differences between cells
are recognized. They propose a model which is segregated and structured,
as is the ONECELL model when viewed as modeling the growth of a culture.
That is, when ONECELL is used to generate behavior of cells of several
different metabolic states, and these behaviors are combined to yield
culture growth characteristics (as is done in this dissertation),
ONECELL is a structured (and segregated) model of a microbial population.
Eakman, et al. (p. 37) also classify models according to whether or
not they take explicit account of interactions between population
and environment. Their model does this also. However, there is really
very little further similarity between their model and ONECELL, as
they compress the entire metabolic state of the cell into a single
variable, cell mass. While this simplifies enormously the task of
obtaining data for use in the model, it obviously is not intended to
address itself to the broad range of situations which can be studied
with the ONECELL model. Their model interacts with its environment
only through a single variable, substrate concentration, allowing
none of the richness of behavior evidenced by ONECELL in growth in
different types of substrates.

Fredrickson, Ramkrishna, and Tsuchiya (1967) take a somewhat

broader view in another paper, and define a framework for describing
models using a vector of biochemicals to specify the physiological
state of a cell. They develop this framework as a set of probability
distributions, taking account of the variations between cells. They
do not develop a model, of a particular situation, *per se*, but derive
some results concerning the form of models of various sorts. They
make perhaps an overly strong statement about the difficulty of study-
ing nonrepetitive growth, saying that, "The fastest computers presently
available could possibly solve problems in a reasonable time (minutes
or hours at most) if the physiological state vector had no more than
three or four elements; this reduction in dimensionality must be made
if nonrepetitive growth is to be attacked mathematically." This
assertion is, no doubt, a product of their particular framework for
model formulation; it seems doubtful that mathematical attacks on
nonrepetitive growth need be made in that generality.

Ramkrishna, Fredrickson, and Tsuchiya (1967) in another paper,
do take a step in the direction of a biochemically structured model.
They consider two pools: nucleic acids and everything else. They
use this model to predict the occurrence of a lag phase in batch growth.
Their model is distributed rather than segregated. Along the continuum
of degree of structure, if their model is structured, ONECELL would
have to be called highly structured, and highly complex as a result.

A simple, nonprobabilistic model with two differential equations
is presented by Aiba, Nagai, Endo, and Nishizawa (1969). Their model
is distributed, and nonprobabilistic. They have used an analog computer
in some of their modeling work.

Powell (1967) proposes a model for growth rate μ in terms of

substrate concentration in *steady-state* growth, and compares his fit with others. Clearly this model is of a very different sort, not attempting to account for the reaction of μ to changing substrate concentration.

Simon (1970) has developed a model with many similarities to the ONECELL model. His model is of a single cell, and uses the replication model proposed by Helmstetter, Cooper, Pierucci, and Revelas (1968), as does the ONECELL model. His model considers three enzymes, and four additional biochemical pools. He does not model the feedback inhibition of enzymes. However, the enzymes are repressible in his model. He uses the Helmstetter-Cooper model, together with the ideas of substrate saturation and enzyme repression, to calculate a minimum cycle length (interdivision time) under conditions of substrate saturation. The model would not be suitable for studying the kinetics of a shift from one medium to another due to its lack of rapid-acting controls, but it is a step closer to the ONECELL approach than those models previously discussed. His equations are still of sufficient simplicity that he did not need to use a computer simulation to arrive at his minimum cycle length result.

Chance, Garfinkel, Higgins, and Hess (1960) have been engaged for some time in experimental study and modeling of the control mechanisms operating in ascites tumor cells. They have developed a model which describes the interrelation between glycolysis and respiration in these cells. Their model is a very detailed one, with components consisting sometimes of single enzymes; for this reason, it can model only a few pathways. This, of course, does not in the least diminish the worth of their model: its scope only needs to be broad

enough to include the problems of interest to them. They have used a digital computer to simulate the behavior of their multi-enzyme model, running over simulated time periods on the order of a minute. Their model is discussed in Rosen (1968), and Rosen states that, "it should not be surprising to see essentially the same basic model accounting for regulatory phenomena on quite distinct time scales." ONECELL, while not of a comparable level of detail of representation, is certainly an attempt to model the metabolism of a cell using the same fundamental approach as Garfinkel and his co-workers, but over a much longer time scale.

The model that probably bears the most resemblance to ONECELL (in terms of level of the entities which it chooses to group for representation as a single variable) is the "advanced cellular model-system" of Heinmets (1966).

This is a single-cell model which he develops from his earlier model, modifying it to include the control mechanisms of induction and repression. His earlier model identifies several metabolic pools, enzymes, genes, and messengers. It does not treat the process of cell division and genetic replication, however, so cannot be used to model extended growth periods; his advanced model does introduce mechanisms to allow cell division. In his earlier model, he is able to manipulate initial conditions and rate constants to determine the effect on the behavior of the model. He has realized the model on an analog computer, and has generated a large number of curves showing the relation of one pool to another. It should be pointed out that Heinmets does not "solve" his system so as to obtain initial conditions compatible with steady state growth: his curves typically show an

initial peak or dip, followed by a slow rise or fall. In his advanced model, Heinmets takes a somewhat different approach from that of ONECELL. In this model, Heinmets has many variables which have direct analogs in the ONECELL model, but he treats them somewhat differently. Each "reaction" in his model is between two reactants, and he treats each explicitly, with a rate constant associated with each. In contrast, ONECELL lumps a sequence of reactions together, and assigns a single rate curve to describe the rate of formation of end-product in terms of the concentrations of the multiple factors which appear in the reaction sequence. This considerably reduces the number of parameters which need to be specified in ONECELL, removing the need for knowledge of rates of formation of many intermediates and complexes, which appear explicitly in the models of Heinmets. It is probably this fact that makes it possible to implement ONECELL on a computer, while Heinmets says that his advanced model-system cannot, at the present time [1966], be analyzed on the computer. Of course, Heinmets was restricted by the fact that he used an analog computer; with a digital computer his model would not suffer from size limitations, but more likely from the shortage of information necessary to set the large number of parameters in his model. It should be pointed out, however, that Heinmets was not attempting to model the behavior of a particular cell, but to develop a model which exhibits cell-like behavior. ONECELL, while not at the other extreme, is directed much more toward a specific type of organism. Other areas in which the ONECELL model reflects a higher degree of correspondence with living cells are: the differentiation among nutrients and their transport mechanisms, the model for DNA replication and cell division, and factors such as the

surface-to-volume ratio, which does not appear in the Heinmets model.

Three radically different approaches from that of ONECELL will be mentioned, to indicate the breadth of attack which is possible in modeling a simple living system. Goodwin (1963) treats large portions of the metabolism as random variables, changing rate constants in the parts modeled explicitly. He looks at the occurrence of periodic phenomena as functions of the distributions of these random variable parameters. He uses statistical-mechanical notions to analyze the types of behavior his model could exhibit. Another formalism, that of automata theory, has been used to obtain two additional type of models. One, that of Stahl (1965), represents a cell as a Turing machine, writing metabolites on its tapes as symbols, forming macromolecular "words", and so on. He uses this model to describe the process of differentiation. The other automata-theoretic model is by Langer, Krohn, and Rhodes (1968), and represents metabolism in terms of the semigroup theory and machine theory in which Krohn and Rhodes have been heavily involved. They do not actually formulate a cell model, but present a framework for representing metabolism, and an example for a small portion of bacterial metabolism.

It should be fairly clear from the discussion above that the ONECELL model is not the first model which attempts to represent the growth of a cell, or even of a bacterial cell. It is unique, however, in the levels of organization and control it employs, the breadth of environmental conditions in which it is able to function, the many types of behavior it is able to exhibit, and the flexibility which is inherent in its organization. It is the author's belief that the particular form of "the model" at this point in time does not reflect

the best of all possible models; in fact, it does not utilize some large areas of current biological knowledge. However, the author feels that the *approach* exemplified by the ONECELL model, and the general level of metabolic representation used in the model, will prove to be increasingly important in the understanding of cellular growth and regulation, and will be used by biologists in further elucidation of the processes of life.

CHAPTER 2

Description of the Model and the Simulation

The description of the model and the simulation will be given

in several phases:

(1) Presentation of the principal pools used in the model

(2) Discussion of State Model Terminology

(3) Discussion of the state variables and the general forms

     of the state equations, showing what state variables enter

     into each state equation

(4) Discussion of approaches used in formulating the simulation

     and the model

(5) Presentation of the complete equations for an example pool,

     NUC.

Section (1) - Presentation of the principal pools used in the model

The principal pools used in the model are given in Goodman,

Weinberg, and Laing (1971). A graphical representation of the material

flows is shown in Figure 2.1.

This representation of the metabolism of the cell is based on

the grouping shown in Figure 2.2.

The abbreviations which will be used throughout this dissertation

are the variable names which appear in the ONECELL simulation. They

are listed here for convenient reference. Also listed are the enzymes

associated with various synthetic and transport systems. For each

enzyme EK(i) there also exists a mRNA, called RNK(i), which codes for
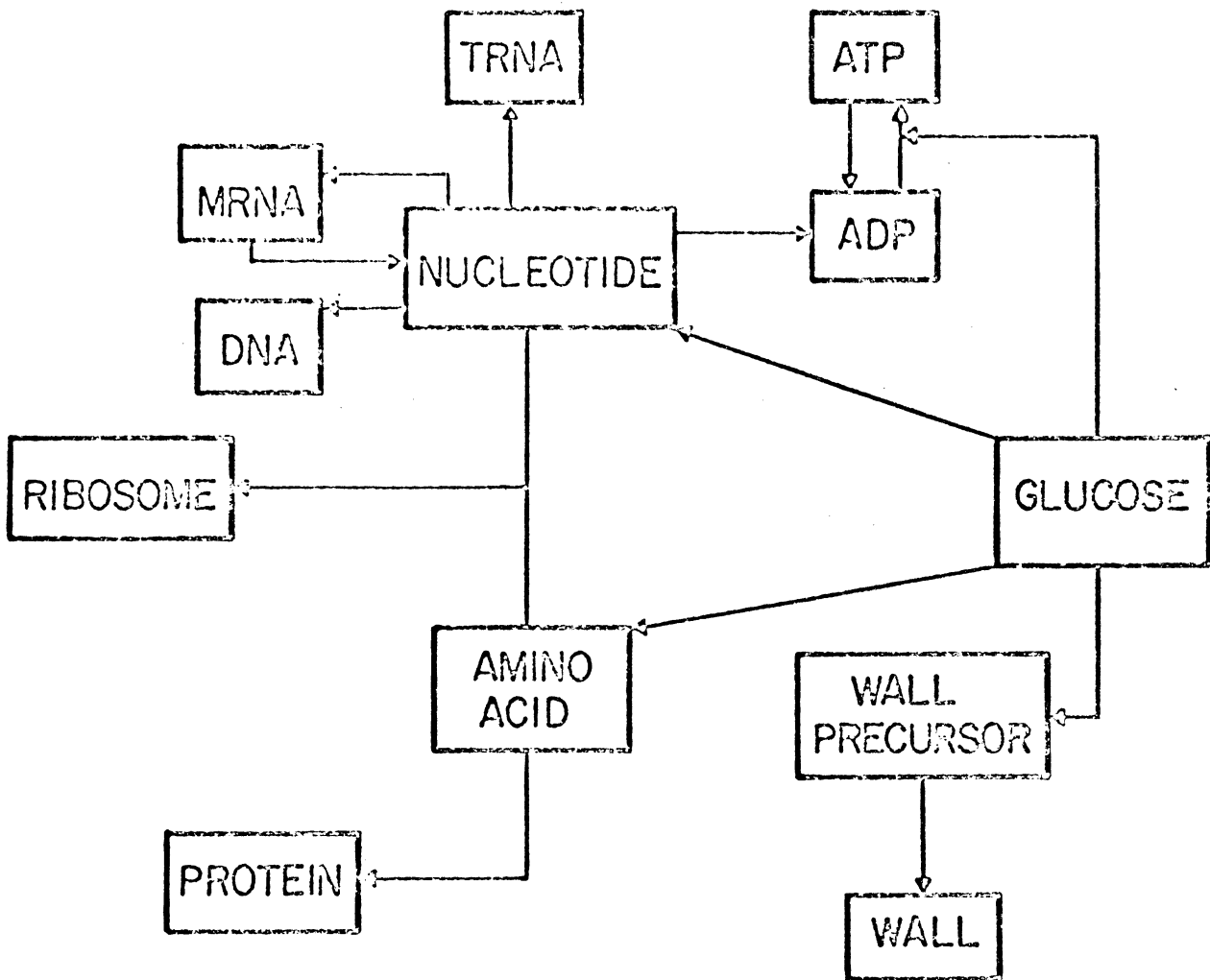
Figure 2.1: The Primary Chemical Pools used in the Model.
The arrows indicate the direction of the flow of
material between pools. The topology of this
flow preserves the metabolic topology of Figure 2.2,
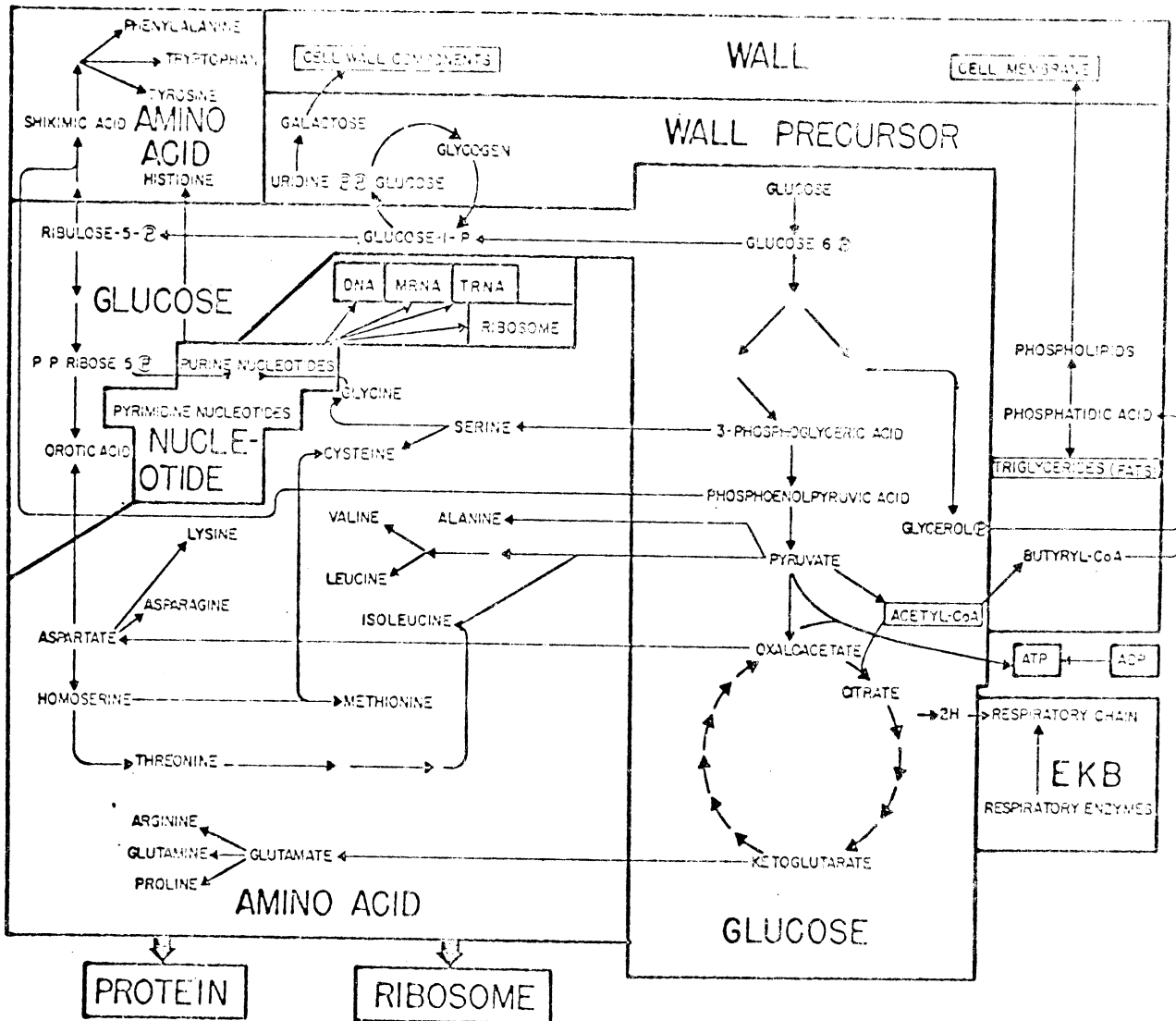in a manner explained in the text.

Figure 2.2:   Some Primary Constituents and Pathways in *E. Coli*.
These constituents are shown grouped together into
"pools", e.g., GLUCOSE, AMINO ACIDS, etc., which
serve as the entities for a model of a cell.

the synthesis of EK(i).

NUC   -   the concentration of (ribo- and deoxyribo-) nucleotides
          in the cell.

AA   -   the concentration of amino acids in the cell

ATP   -   the concentration of ATP in the cell (kept as a variable
          separate from NUC)

WALL -   the concentration of cell wall and membrane precursors in
          the cell

ADP   -   the concentration of ADP in the cell

DNA   -   the concentration of DNA in the cell (derived from DNA1,
          DNA2, DNA3 as discussed in Section (4))

PRTN -   the concentration of the sum of the enzymes EK(1) - EK(14)
          in the cell. Ribosomal protein is kept separately.

MRNA -   the concentration of the sum of mRNA's RNK(1) - RNK(14)
          in the cell

RIB   -   the concentration of ribosomes in the cell

TRNA -   the concentration of transfer RNA's in the cell

GLUC -   the concentration of glucose and its products (Figure 2.2)
          in the cell

LAC   -   the concentration of lactose in the cell

VOL   -   the cell volume

NO   -   the number of (identical) cells represented by this ONECELL.
          (NO is doubled at each simulated cell division.) NO is
          sometimes called COUNT in later sections.

MASS -   the weighted sum (in daltons) of the amounts of each pool
          in the cell. (An output.)

EK(1) -   the concentration of enzymes for NUC synthesis

EK(2) -   the concentration of enzymes for AA synthesis

EK(3) -   the concentration of enzymes for ATP synthesis

EK(4) -   the concentration of enzymes for WALL synthesis

EK(5) -   the concentration of enzymes for ADP synthesis

EK(6)   -   the concentration of enzymes for DNA synthesis

EK(7)   -   the concentration of enzymes for PRTN synthesis

EK(8)   -   the concentration of enzymes for MRNA synthesis

EK(9)   -   the concentration of enzymes for RIB synthesis

EK(10) -   the concentration of enzymes for TRNA synthesis

EK(11) -   the concentration of enzymes for GLUC transport

EK(12) -   the concentration of enzymes for LAC transport

EK(13) -   the concentration of enzymes for nucleoside transport

EK(14) -   the concentration of enzymes for amino acid transport

IN   -   the concentration of an initiator required for start of
            DNA replication

SVR   -   the surface-to-volume ratio

EXT(1) -   (EXT_{NUC}) external concentration of nucleosides (an input)

EXT(2) -   (EXT_{AA}) external concentration of amino acids (an input)

EXT(3) -   (EXT_{GLUC}) external concentration of glucose (an input)

EXT(4) -   (EXT_{LAC}) external concentration of lactose (an input)

With the exception of DNA, PRTN, MRNA, and SVR (which are derived

from finer-structured state variables), the EXT's (which are inputs),

and MASS, which is an output, the variables above are state variables,

as discussed in the next section.

For ease of comparison with published data, these output and

state variables are often transformed by converting concentrations

to amounts and taking logarithms. Crude representations of unsynchron-

ized cultures are done by making several runs and recording (and

changing environmental conditions) at different stages of the division

cycle. These procedures are described in Chapter 3 where they are

used in generating graphical output.

Section (2) - Discussion of state model terminology

It is possible to divide the numerical entities of the model into several classes: some are constant, some vary, some are integers, some are real numbers, some are solved for, some are estimated from experimental literature, and so on. One particularly important class of numerical entities is the set of *state variables*. Roughly speaking, these are the variables which may change during the course of time in the model, for which the model must retain a value from one instant to the next. That is, if time is treated as a set of discrete instants, a variable is *not* a state variable if it can be calculated at any instant of time using the inputs to the model at that instant of time, and its previous value is not needed to calculate the value of any variables. State variables, as the name implies, specify the *state* of the model at any time, and calculation of the values of any variables can be done using only the state variables and the inputs.

If a model can be expressed as a set of first order differential equations and algebraic equations, it is the variables whose derivatives appear in the differential equations that are the state variables. this specification of the state variables agrees with the description in the preceding paragraph, since in order to carry out a numerical solution to a differential equation, it is necessary to calculate the value of the state variable at a particular time using its value at a previous time (to which the change is added). Of course, the two

instants of time are brought arbitrarily close together in the limit, but there is no escaping the necessity for "remembering" a previous value to calculate a new one.

When one is considering a model with respect to preparing a simulation, the state variables are of particular importance. The values of the state variables are the only values which must be updated in the simulation. That is, all other entities are either constant or may be calculated at any time from only the constants, inputs, and state variables. The importance of this for computer simulations is very great. For example, if one wished to have the capability of running a simulation for some time and then jumping back to an intermediate state and trying the effect of a different sequence of inputs, he could do this easily by saving (on cards or some magnetic peripheral storage device, for example) the values of the state variables at the times to which he might wish to return. Other nonstate variables which may have been calculated need not be recorded, as they can be regenerated from the state variables (and inputs, if necessary).

For reasons such as the one above, it seems desirable to present a "skeleton" of the cell model, showing which variables are state variables, and showing, for each state variable, the set of state variables and inputs on which calculation of its value depends. The notation to be used is shown in the following example:

$$dx/dt = f(y,z)$$

This indicates that y and z are the state variables (or inputs) upon which the rate of change of x depends. One additional notational convenience will be used. Sometimes a state equation is easily

expressed, for example, in the form:

$$dx/dt = k \times x \times y \times dz/dt$$

In such a case, dz/dt is not a state variable, so dz/dt should not appear as a state variable on which x depends. In fact, dx/dt can be calculated from x,y, and the *set of state variables and inputs upon which z depends*. This should be clear since this set enables dz/dt to be calculated, then used to calculate dx/dt. The situation above will be expressed as $dx/dt = f(x,y,D(dz/dt))$, which should be interpreted as saying that the next value of x depends on the previous values of x,y, and of the state variables and inputs on which z depends. At some points, the notation $S \subset D(dz/dt)$ is used to indicate a *subset* of $D(dz/dt)$.

Section (3) - State equations

NUC

The pool called NUC represents free nucleotides present in the cell, except ATP and ADP, which are handled separately. The units used for this pools, and for most of the others, are molecules per unit volume, where the unit of volume is the volume of an *E. coli* cell in well aerated glucose minimal medium (called in this simulation environment 1).

The direct dependence of the rate of change of NUC on the state variables and inputs of the model is expressed as:

$$dNUC/dt = f_1(NUC,GLUC,STP,EK(1)) + f_2(D(dDNA/dt),D(dMRNA/dt),$$
$$D(dTRNA/dt),D(dRIB/dt),S \subset D(dADP/dt) \; S \subset D(dATP/dt))$$
$$+ f_3(EXT_{NUC},NUC,EK(13),D(SVR))$$

In the equation above, f_1 represents the synthesis of NUC, using

GLUC as substrate. The noncarbon substrates are assumed to be

nonlimiting. The ATP represents the energy required for synthesis

of NUC, and EK(1) represents the set of enzymes which catalyze synthesis

of various members of the NUC pool. The f_2 term expresses the usage

of NUC for synthesis of other pools. The f_3 term represents the

transport process, which enables the cell to use NUC molecules present

in the growth medium if the cell is growing in broth or another

nucleoside-rich medium. The EXT_{NUC} is an input to the model, namely

the concentration of NUC outside the cell. EK(13) is the set of

enzymes responsible for transport of NUC into the cell. The D(SVR)

term represents the set of state variables necessary to calculate SVR,

the surface-to-volume ratio, which is not chosen as a state variable.

The equation for dNUC/dt given above actually describes the change

of NUC per unit of volume at the beginning of the time step. The

new concentration of this pool is calculated also taking into account

the change in volume during the time step, so there is an additional

dependency of this and all the other pool equations on VOL and

D(dVOL/dt).

AA

The AA pool represents the amino acids present in the cell

(excluding those which have already been incorporated into structures

like protein or cell wall, etc.). The rate of change of the concen-

tration AA may be related to the other pools and inputs of the model

as:

$$dAA/dt = f_1(AA,GLUC,ATP,EK(2)) + f_2(D(dPRTN/dt),$$
$$D(dRIB/dt)) + f_3(EXT_{AA},AA,EK(14),D(SVR))$$

As was the case for NUC, the f_1 term represents synthesis of AA, the f_2 term handles usage of AA to form protein, and the f_3 term deals with transport of AA into the cell if it is present in the growth medium.

ATP and ADP

ATP represents the concentration of adenosine triphosphate molecules in the cell. It is used in the model as the principal supply of energy for the synthetic reactions of the cell. It is related closely to the ADP, or adenosine diphosphate, pool. ADP is one of the forms which results when ATP gives up part of its stored energy to drive a chemical reaction. ATP and ADP may both be viewed as high-energy derivatives of AMP, adenosine monophosphate, a mono-nucleotide which is included in the NUC pool. ADP is formed from NUC (at the expense of ATP). The energy for creating ATP from ADP comes from the GLUC pool, principally by way of the respiratory chain (aerobic growth is assumed in the model).

The dependence of the synthesis of ATP on the other state variables of the model is of the form:

$$d(ATP)/dt = f_1(ATP,ADP,GLUC,EK(3),EK(5))$$
$$+ f_2(S \subset D(dNUC/dt),S \subset D(dAA/dt),S \subset D(dWALL/dt),$$
$$D(dDNA/dt),D(dRNA/dt),D(dRIB/dt),D(dVOL/dt)),$$

where the f_1 term represents the synthesis of ATP from ADP, and $S \subset D(dNUC/dt)$ represents the state variables necessary for calculating the amount of NUC *synthesized* (not *used*), and similarly for $S \subset D(dAA/dt)$ and $S \subset D(dWALL/dt)$.

The f_2 term represents loss of ATP (becoming ADP). The f_2 term depends on the rates of synthesis of most of the pools within the cell, since ATP supplies energy to them.

The corresponding expression for change of ADP is:

$$d(ADP)/dt = f_1(ADP,ATP,NUC,EK(5)) + f_2(D(dATP/dt))$$

where the f_1 term represents synthesis of ADP from NUC, using enzyme EK(5), and the f_2 term represents net conversion of ADP to ATP (including both ATP formed from ADP and ATP becoming ADP).

WALL and VOL

Cell wall and cell membrane are lumped together in this model. The amount of cell wall and membrane is not kept explicitly as a variable of the model; rather, the cell volume is kept. The surface-to-volume ratio at any time can be calculated from a relatively simple geometric model of the cell, using the volume. Thus it is possible to calculate the amount of surface area (= volume × surface/volume ratio), which determines the amount of cell wall and membrane, assuming that the cell wall and membrane have a constant thickness and uniform composition. The rate of change of cell volume is calculated as a function of the concentration of precursors available for constructing wall and membrane. While the cell wall and membrane are composed of many types of substrates, the model will assume that the principal components are derived from lipids and carbohydrates. The model takes into account the need for the amino acids which are found in mucopeptides and lipoproteins, but they are not included as part of the precursor pool. The substrates for cell wall and membrane synthesis, consisting of lipids and carbohydrates not represented

elsewhere in the model (i.e., glucose and the products of the glycolytic pathway are not included in the cell wall precursor pool), will be grouped and called WALL. It should be remembered that WALL represents precursors, not the in-place cell wall and membrane. The model for cell volume utilizes the state variables shown below:

$$dVOL/dt = f(WALL, ATP, D(dPRTN/dt), VOL)$$

The dependence on VOL is autocatalytic; i.e., the rate of change of VOL goes up as VOL goes up.

The model for WALL, cell wall and membrane precursors, can be similarly represented as:

$$dWALL/dt = f(GLUC, ATP, EK(4), D(dVOL/dt), D(SVR), TERM, ITRM),$$

where EK(4) represents enzymes which enable synthesis of cell wall and membrane precursors, D(SVR) stands for the state variables which determine the surface-to-volume ratio, and TERM (time since termination of a DNA replication round), and ITRM, a variable set to 1 when cell division is imminent, are needed to determine what use of WALL, if any, is necessary for formation of additional wall and membrane near cell division time.

DNA and Initiator Protein IN

The model for DNA replication is based primarily on the model proposed by Helmstetter, Cooper, Pierucci, and Revelas (1968). First the functional dependencies of the variables involved in the DNA replication process will be presented here, mainly for reference purposes:

$$dDNA3/dt = f_3(NUC, ITRM, IN2)$$

$$dDNA2/dt = f_2(NUC, ITRM, ITRM2, IN1, IN2)$$

$$dDNA1/dt = f_1(NUC,ITRM,ITRM2,IN1)$$

$$dIN/dt = f_4(D(dPRTN/dt),IN,IN1,IN2,VOL)$$

$$dTERM/dt = f_5(ITRM,WATRT)$$

$$IN1 = f_6(IN,IN1,VOL)$$

$$IN2 = f_7(IN,IN1,IN2,VOL)$$

$$ITRM = f_8(DNA1,ITRM,GLUC)$$

$$ITRM2 = f_9(ITRM,DNA1)$$

$$DNART = f_{10}(NUC)$$

$$WATRT = f_{11}(GLUC)$$

IN1, IN2, ITRM, and ITRM2 are state variables, but are logical (not numerical) in form, and their behavior is not described by a differential equation (but is given later in this section). DNART and WATRT are not state variables, but instantaneous functions of NUC and GLUC, respectively.

The model of Helmstetter, et al. can be described by these assertions:

(1) Initiation of DNA replication occurs (at a fixed point on the genome) every τ minutes, where τ is the doubling time of the cells.

(2) DNA replication progesses at a constant rate after initiation, traversing the entire genome in 40 minutes, for cells with $\tau \le 60$ minutes.

(3) For cells with $\tau > 60$ minutes, replication of the DNA (after it is initiated) takes about $2/3$ τ.

(4) After a round of DNA replication is completed, cell division follows in 20 minutes (for $\tau \le 60$ minutes) or about $1/3$ τ (for $\tau > 60$ minutes).

(5) DNA replication is initiated after a fixed amount of initiator per chromosomal origin has accumulated, regardless of the growth conditions.

(6) This initiator is accumulated as if it were the result of protein synthesis; when protein synthesis is blocked, accumulation of initiator stops. Thus, it seems that a total of τ minutes of unrestricted protein synthesis are necessary for accumulation of the required amount of initiator.

An example of the behavior implied by these assumptions may help to clarify their meaning. For a cell which is in exponential growth with a doubling time of 30 minutes, the state of the DNA replication apparatus just after cell division is as shown in Figure 2.3.
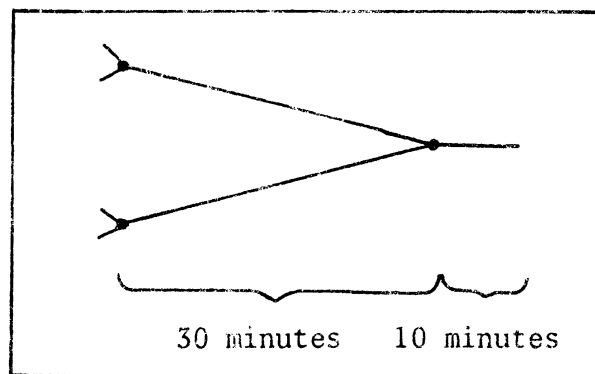

Figure 2.3: DNA Configuration after Cell Division, τ = 30. The circular chromosome is represented as opened flat (at the site where replication is initiated). The dots on the left end represent points where DNA replication has just been initiated; the third dot shows a replication point that has progressed thirty minutes down the chromosome.

If division is at t = 0, then DNA replication must have been initiated at t = -(20+40) = -60 minutes. Then, since τ = 30, replication must have been initiated again at t = -30 minutes. At t = -30, then there were 3 points moving along the chromosome. At t = -20, the first round of DNA replication must have been completed, leaving the second replication point 10 minutes (1/4) along the chromosome. Division must then occur 20 minutes later, that is, at t = 0. But another round is also initiated at t = 0 (30 minutes after the last round started). The round initiated at t = -30 has progressed 30 minutes down the chromosome (i.e., 3/4 of its length). Thus we find the situation must be as shown in Figure 2.3, where the dots represent the moving sites of DNA replication. The same logic can be used to convince oneself that the situation at t = +30 will be the same as that at t = 0. Since at a growth rate τ = 30 minutes there are always two sites where DNA replication must be initiated, the amount of initiator to be accumulated during each 30 minute period is 2 units (where 1 unit per origin is required for initiation). It is assumed that initiation "uses up" or breaks down the (presumably protein) initiator, so that the entire 2 units must be synthesized during the division cycle.

Like most models, this model of DNA replication and its regulation is only applicable under certain conditions. Some experimental evidence, for growth under the conditions within the scope of the rest of the cell model presented here, lends considerable support to the replication model of Helmstetter, et al. This work is cited and discussed in their paper. On the other hand, the work of Ward and Glaser (1970),

indicates that control of the process of initiation is not (in their strain, at least) independent of the synthesis of DNA, as predicted by the model used here.

The ONECELL model for DNA replication and cell division includes the assumptions of the model of Helmstetter, et al., and requires some additional ones. In particular, the time needed for DNA replication and the time between completion of DNA replication and the occurrence of cell division in slow growth ($\tau > 60$) must be functions of the metabolism of the cell. Also, the control over the rate of initiator synthesis must be made explicit.

The model of Helmstetter, et al. has been restructured into a form which is easily implemented in a computer simulation. The DNA of a cell is divided into three categories:

(DNA1): DNA between the terminal end of a chromosome and the replication point closest to reaching it. In the case of no replication, DNA1 includes the entire chromosome.

(DNA2): DNA between the furthest-advanced replication point and the next replication point (on each strand leading from the furthest replication point). If there are no additional replication points, DNA2 extends to the origins.

(DNA3): DNA between the DNA2 region and the origins of the chromosome.

When a round of replication is completed, DNA1 is, of course, 0. DNA which was up to that time designated as DNA2 now becomes DNA1, and that DNA3 becomes DNA2. DNA3 is now 0, assuming that

no generation times τ < 20 are considered. Figure 2.4 illustrates
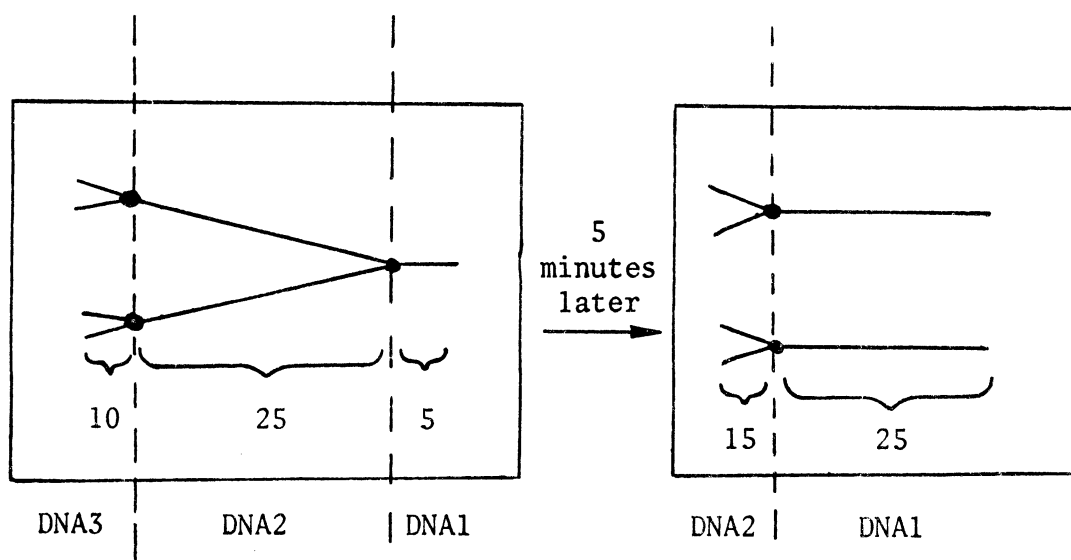this partitioning of DNA for two configurations of DNA.

Figure 2.4: Representation of DNA Configuration. Amounts of DNA1,
DNA2, DNA3 are kept as fractions of a single
nonreplicating chromosome (so the value of DNA1 in the
left hand cell is 5/40 = .125, and the value of DNA2
is 2 × 25/40 = 1.25).

The initiator substance is assumed to be a protein, and its

concentration is denoted IN. In order to obtain an *amount per*

chromosomal origin, IN must be multiplied by VOL and divided by the

number of origins. Several additional variables are used in the model.

One is ITRM, which is set to 1 when a round of replication is finished,

and to 0 when cell division occurs. Others are IN1 and IN2. IN1 is

1 when any replication point is present on the chromosome, and 0 other-

wise, IN2 is 1 whenever there are multiple forks on a chromosome

(i.e., DNA3 is being formed). ITRM2 is set to 1 when a round is

completed but the (normally) 20-minute post-termination sequence from

the previous round has not been completed (which can occur only under

slow-growth conditions).

The slow-growth behavior is controlled by DNART and WATRT, which

are functions of NUC and GLUC respectively. DNART determines the rate

at which a DNA replication point moves along the chromosome. DNART

is 1 when NUC is above a threshold value (which it is whenever $\tau > 60$),

and a replication round takes 40 minutes. When NUC is lower, DNART

falls off sharply and a round takes longer. (The equation for DNART

is DNART $= \min(NUC/(.9 \times NUC0),1.0)^{2.4}$, where NUC0 is the initial

concentration of NUC in environment one. Thus low NUC causes DNART

to be low, and the replication points move more slowly, so the rate

at which DNA1 yields DNA2 is decreased, as is the rate at which DNA2

yields DNA3. The numbers .9 and 2.4 used in the equation were chosen

to give a DNART that falls off fairly sharply for values of NUC below

those in environment one.) WATRT has a similar effect on the rate at

which TERM, the time between completion of a round of replication and

the cell division which follows it, is decremented. WATRT is a function

of GLUC, and low GLUC causes WATRT to be low (WATRT is linear with

GLUC below the threshold), so the time before TERM goes to 0

(signaling cell division) is longer. In a sense, both DNART and

WATRT can be viewed as measures of the rate of flow of "physiological

time", not unlike the concept of "degree-days" which has been employed

in studies of growth (Caswell, Koenig, Resh, and Ross, 1971).

PRTN

The protein pool is actually a sum of 14 separately-maintained

pools, the pools, EK(1) through EK(14). Each of these pools represents

the concentration of a group of enzymes which catalyze a particular

set of reactions associated with one of the pools of the model. The

protein present in cell membrane and in ribosomes is not included in

the PRTN pools; the model deals with these proteins elsewhere. A

large part of the total biochemical apparatus of the cell is involved

in the production of substrates for or regulation of protein synthesis,

so it should not be surprising that the protein synthesis function in

the cell model contains a large number of terms. Abstractly, the

equation for synthesis of each type of enzyme pool is:

$$dEK(i)/dt = f(RNK(i),EK(7),ATP,TRNA,RIB,EK(i),PRTN),$$

where RNK(i) represents the mRNA's that code for enzymes in EK(i),

and EK(7) represents the enzymes that catalyze translation of mRNA

into protein.

MRNA

The pool MRNA, like PRTN, is a sum of smaller pools, called

RNK(1) through RNK(14). Each RNK(i) stands for the concentration of

messenger RNA's that code for enzymes in EK(i).

The equations for synthesis of the RNK's are of the general form

$$dRNK(i)/dt = f(Y(i), NUC, DNA, EK(8)ATP, AA, RNK(I), MRNA),$$

where Y(i) stands for the pool for which EK(i) catalyze the synthesis or transport. The dependence on NUC is for precursors. The role of DNA is as a template for forming mRNA. The Y(i) concentrations, for some i's, enter into the mRNA synthesis by repressing or inducing (or derepressing) synthesis of the corresponding mRNA's. The total MRNA is used in calculating an overall rate-constant for synthesis of mRNA.

RIB

The pool RIB represents the concentration of ribosomes in the cell. The units are number of 100 S (sedimentation constant of 100) particle-equivalents per unit volume, where each 100 S particle is viewed as the equivalent of 2 70 S ribosomes. The reason for using these units is a holdover from the past, but there is no adverse effect, since the half-size number is used throughout all places where RIB appears.

Ribosomes are composed of both protein and RNA (called rRNA for ribosomal RNA). There is experimental evidence (see, for example, Kelley and Schaechter (1968)) that control of synthesis of the protein fraction can be exacted separately from control of synthesis of the rRNA fraction. This model, however, will treat RIB as if it were a single entity with a specified composition, and will not allow for synthesis of an excess of ribosomal RNA over the supply of ribosomal protein, for example.

The model for ribosome synthesis is of the form:

$$dRIB/dt = f(RIB,NUC,AA,EK(9),ATP,D(dPRTN/dt)),$$

where $D(dPRTN/dt)$ denotes all of the state variables of which $dPRTN/dt$ is a function.

TRNA

The transfer RNA pool, TRNA, is composed of the low-molecular-weight RNA molecules which play a role in interfacing amino acids with mRNA templates in the protein synthesis process. There is a different tRNA specific for each distinct type of amino acid, but in this model, they are all lumped as TRNA. The synthesis of TRNA is assumed to be of the form:

$$dTRNA/dt = f(TRNA,NUC,AA,DNA)$$

GLUC

The GLUC pool consists of the intracellular glucose and intermediates in the breakdown of glucose along several pathways. The model for the concentration GLUC is of the form:

$$dGLUC/dt = f(EXT_{GLUC},GLUC,LAC,EK(11),EK(12),D(dNUC/dt),$$

$$D(dADP/dt),D(dAA/dt),D(dATP/dt),D(dWALL/dt),D(SVR)),$$

where EXT_{GLUC} represents the concentration of glucose in the medium in which the cell is growing, LAC represents the concentration of lactose in the cell, EK(11) is the concentration of enzymes responsible for transport of glucose into the cell, and EK(12) is the concentration of enzymes for conversion of LAC to GLUC. $D(SVR)$ denotes the state variables on which the surface-to-volume ratio depends.

A more detailed description of the model for GLUC is

$$dGLUC/dt = f_1(GLUC, EXT_{GLUC}, EK(11), D(SVR)) + f_2(LAC, EK(12))$$

$$+ f_3(D(dNUC/dt), D(dAA/dt), D(dATP/dt), D(dWALL/dt),$$

$$D(dADP/dt))$$

of GLUC in synthesis of NUC, AA, ADP and WALL, and of ATP from ADP.

It is assumed that these uses of GLUC do not overlap; for example,

energy available as ATP when GLUC is used in AA formation is ignored.

LAC

The LAC pool represents the concentration of lactose molecules

in the cell. Lactose is brought into the cell, and its only method

for entering into the general cellular metabolism is assumed to be

by conversion to glucose and galactose, which in turn enters the GLUC

pool. The form of the LAC model is thus:

$$dLAC/dt = f_1(EXT_{LAC}, LAC, D(SVR), EK(12)) + f_2(LAC, EK(12)),$$

where EXT_{LAC} represents the concentration of lactose outside the cell,

and EK(12), represents both the enzymes for transport of LAC into

the cell (β-galactoside permease) and for conversion of LAC to GLUC

(β-galactosidase and enzymes associated with galactose conversion

to components of the GLUC pool).

The function f_1 represents the process of transport of LAC into

the cell (and the exit process as well), and f_2 represents conversion

of LAC molecules to the GLUC pool.

Section (4) - Approaches in Formulating the Model and Simulation

The ONECELL model has at its core a set of differential equations. These equations express the rate of change of many of the cell's vital constituents, as the cell grows and changes. Most of the variables which appear in these equations represent concentrations of groups of functionally-related biochemicals. The extent to which the model captures essential features of the life of the cell depends upon the adequacy of the set of variables chosen in representing the activities of the cell and the degree of validity of the relations assumed to hold among these variables.

Aggregation

First it should be noted that there is inaccuracy inherent in the lumping of related types of molecules to consider them as a single pool. An obvious difficulty is the necessity to reconcile the differences in weight among various molecules. For example, the AA pool (amino acids) is assigned a certain concentration (number of molecules/unit volume), but removal of a particular amino acid, say arginine, will result in a change in the average weight, as well as the number, of molecules remaining. This is a problem when numbers of molecules used by a particular synthetic pathway are estimated using the average molecular weight of the product divided by the average molecular weight of the substrate pool. If a large inaccuracy results, it can be decreased by considering the concentration of the pool to be the number of average-weight-equivalents of each type of molecule. for example, if a pool of average M.W. = 200 contained 25 molecules

of M.W. = 400, these 25 molecules could be considered, where necessary, to be 50 average-weight-equivalents.

The inaccuracy attributable to differences in weights is small by comparison to the other difficulties which may arise out of lumping to form pools of functionally-related molecules. The larger problem is that the different chemical species lumped may *not* have similar functions with respect to some aspects of the cell's metabolism. The pools chosen represent an attempt to limit this divergence of behavior within a pool, but it should be clear that this can only be accomplished to a limited degree.

The rule for choosing chemicals to be grouped as a pool is that which mathematicians might call a homomorphism rule. Lightly stated, it says that one should not put together what the cell can subsequently pull apart. More accurately, it requires that the grouping of two chemicals A and B should be done only if the behavior of the cell is influenced only by the *sum* of A and B, and not by the proportion of A+B that is A, or how much of the total is B. This is obviously a very stringent requirement, and whatever aggregation it is possible to do with strict adherence to this criterion reflects unnecessary complexity in the real cell. However, if one considers the behavior of the cell under a restricted set of conditions, and if one is willing to apply the rule loosely, as a guideline, it is possible to perform a great deal of lumping. The formalism for expressing the notion of a homomorphism and its application to this model are discussed in detail in Zeigler and Weinberg (1970).

Equation Forms

The differential equations used in the model are developed in part from examining the chemical kinetics of the reactions being modeled. For example, the chemical reaction $A+B \xrightarrow{k_1} C$ would be represented as $d[C]/dt = k_1*[A]*[B]$.

Similarly, a reaction in which three molecules must meet simultaneously could be represented by $d[D]/dt = k*[A]*[B]*[C]$. This equation is *not* the one which would ordinarily be used to represent the formation of an AB complex which subsequently reacts with C. However, by ignoring the temperature dependence of the reaction (the ONECELL model always models growth at $37°$ C) and by *not* requiring that k be the product of the k's of the individual steps, it is possible to consider the equation

$d[D]/dt = k*[A]*[B]*[C]$ to model a sequence

$A+B \rightarrow AB$

$AB+C \rightarrow D.$
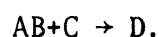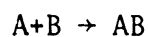
This is the approach used throughout the formulation of the model, with the rate constant k replaced by a function K of the concentration D.

In many instances, it is desirable to include substrates, energy-yielding molecules, enzymes, and other chemical species necessary for a synthetic pathway all in the same equation. This is done by multiplying in factors for the effects of each of these on the reaction rate. The simplest factor used is the concentration itself; it is often assumed that a reaction is linearly proportional to each of the chemical species it requires. Of course, this is not true in general, but it would be extremely difficult to characterize the effect of each species on the reaction rate. The factors used, then, repre-

sent an attempt to introduce some effect, which hopefully is a better

representation of the reaction than would be had were the factor omitted

entirely. In addition, if the concentrations appearing in the equation

are following the pattern they exhibit in the range of environments

used to fit the overall reaction rate curve (discussed next), the

nonlinear variation of the reaction rate with the concentrations

is included in the rate curve calculated.

Rate Curves

One of the principal control mechanisms of a cell is the feedback

control provided by the mechanism of allosteric modification of enzyme

activity.

This control mechanism was incorporated in Weinberg's original

model (Weinberg and Berkus, (1968)). The simulation of the effect of

end-product feedback is done using a variable K for each of the major

synthetic pathways simulated. The K's are subscripted, with $K(I)$

being the effective activity of the enzyme $EK(I)$ in catalyzing formation

of product $Y(I)$. Actually each K is a lumping of many factors, and

is used as the rate "constant" for synthesis of a pool. Of course,

K is *not* constant; it varies with the concentration of the end-products

of the reaction sequence it is associated with, and is recalculated

for each time step. The shape of each K curve, that is the value of

$K(I)$ for each concentration of product $Y(I)$, is determined by solving

a set of simultaneous equations. Values of K are estimated in each

of three environments, using the following procedure:

The amount of product (I) to be produced in one second, per unit

volume, in order to keep $Y(I)$ constant in this environment is estimated.

This amount is called $DY(I)$. The concentrations of all the pools

entering into the synthesis of Y(I) are known. DY can be divided by each of these concentrations, leaving only K. (That is, if DY(I) = K(I)*A*B*C, then if A,B, and C and DY(I) are known for a particular set of conditions, K(I) can be solved for in that set of conditions.)

After the values of K(I) are known for all 3 environments being used to determine K(I), the values are used in a set of three equations in seven unknowns, which arise from the model chosen for allosteric modification. Specification of four additional values allows solution for the K(I) curve, as a function of Y(I). The four values which are to be specified correspond to the rate constants of an enzyme-product complex with no, one, two, and three molecules of product, respectively. Since these values are generally unknown, particularly when they represent rates for sets of enzymes, appropriate values are determined in this model by trying possible values until a K(I) with certain properties is obtained. The K(I) curve is assumed to be monotonic nonincreasing with Y(I) and to flatten out for Y(I) above and below its extreme values in the three environments.

One of the major difficulties in obtaining stable behavior in all three environments is the problem of fitting the K(I) curves. Monotonicity, while necessary, is not sufficient to yield overall stability; that is, the behavior of one pool is affected by the behavior of others, and K curves must be adjusted as a group, not just individually (as discussed in Chapter I).

A detailed description of the equation used for calculating the K(I) in each environment is given in Zeigler and Weinberg (1970). One fact not mentioned there is that the K curve reflects not only

allosteric inhibition, but also some of the nonlinearities of the reaction rate with respect to the substrate and other reactants entering into the synthesis. For example, saturation of the reaction with substrate in one of the environments would be reflected in the K curve.

Repression

Repression (or induction) of enzyme synthesis is also included in the ONECELL model. It is assumed that the enzymes for NUC,AA,ATP, WALL,ADP, and LAC may be repressed. The technique used to implement repression is to sense the level of the controlling pool, say NUC, for example. If it is above a certain level, the rate constant for synthesis of the mRNA that codes for NUC-synthesizing enzymes is cut to a small fraction of its fully-induced value. This results in a rapid reduction of the corresponding mRNA, because of the high lability of mRNA. The level of the associated enzyme pool will, of course, fall more slowly.

In the case of LAC, the "normal" state in the model is the repressed one, and derepression occurs when lactose is present in the growth medium and glucose is not (the "glucose effect" described in Beckwith and Zipser (1970) and elsewhere). Derepression results in a rapid increase in the rate of formation of the mRNA which is transcribed from the LAC operon.

Solution of Difference Equations

The model is largely composed of differential equations. In order to realize the model as a simulation, it is necessary to solve the

differential equations. There are a variety of techniques for numer-
ically integrating a differential equation. The simplest, and that
used in this simulation, involves assuming the derivative is constant
throughout a time interval DT, and using its value at the beginning
of the time interval times the length of time DT to get the change
in Y, called DY. That is, given an equation

$$dy(t)/dt = f(y(t),x(t),z(t),...),$$

the approximate numerical solution at time $t+\Delta t$ can be calculated
by converting it to a discrete-time form:

$$\Delta y(t)/\Delta t = f(y(t),z(t),x(t),...),$$

so

$$\Delta y(t) = f(y(t),z(t),x(t),...) \times \Delta t$$

and

$$y(t+\Delta t) \approx y(t)+\Delta y(t).$$

Clearly, if Δt is chosen to be small enough, any desired accuracy
can be obtained. In the simulation, Δt is called DT, and Δy is called
DY. In effect, the differential equations are converted to difference
equations.

In order that the computer time required to run many hours of
simulated time is not excessive, DT is set at 60 seconds. This results
in poor accuracy in situations where the derivative is changing rapidly
relative to DT, but for most processes, this inaccuracy is not large
relative to the other errors of approximation used in the model.
For four pools, GLUC,LAC,ADP and ATP, this step size is too large,
and they are handled differently (as explained later in this section).

A better estimate of the solution of the equation

$$dy(t)/dt = f(y(t),z(t),x(t),...)$$

can be obtained by using:

$$y_1(t+\Delta t) = f(y(t),z(t),x(t),...) \times \Delta t + y(t)$$

$$y_2(t+\Delta t) = f(y_1(t+dt),z(t),x(t),...) \times \Delta t + y(t)$$

$$y(t+\Delta t) = (y_1(t+\Delta t) + y_2(t+\Delta t))/2$$

This solution is more accurate particularly when $y(t)$ tends to

be oscillatory, that is, where the simpler formula $y(t+\Delta t) =$

$f(y(t),x(t),z(t),...) \times \Delta t$ tends to over-correct and make y alternately

large and small. This is often the case when the step size Δt is

chosen large, and the better approximation is used in the simulation

of GLUC transport where the oscillation problem is severe.

One additional technique is used to reduce the oscillations

introduced by a large DT. For some pools, the DY is "lagged", that

is, saved from one time step to the next. At each time step, a new

DY is calculated and saved, using $\Delta y = f(y(t),x(t),z(t),...) \times \Delta t$.

The old and new DY's are averaged and used as the actual DY to be added

to Y for this step. The result is that each amount DY calculated

is added to the pool Y, but over two time steps, half at a time. Thus,

conservation of mass is maintained, in that an amount of Y used in

some reaction, which appears as a negative contribution to DY, is

finally subtracted in full from Y, although it requires two time steps

to do it.

Sampled Simulation

In a sense, there are two "sides" to the concentration variables

in ONECELL; that is, two ways in which a concentration functions in

the model. The first way is as what a systems theorist might call

an intensive variable. That is, a concentration of one pool acts as a stimulating factor (or retardant) or the synthesis of some other pool. This role may be performed as a substrate, enzyme, template, or in other ways. The importance of the concentration viewed in this way is independent of the actual volume (that is, of the absolute number of molecules), but depends only on the number *per unit volume*. The type of variation that is generally important in this context is the proportion (or per cent) by which the concentration changes, regardless of the absolute numbers involved.

On the other hand, when a pool A is being used in a reaction, but does not limit its rate (so long as A is nonzero), the important aspect of pool A is not the *proportion* by which its concentration changes, but the absolute number of molecules per unit volume which are removed. An example may clarify the distinction being made: suppose pool A consists of 10 moles of A in 1 liter of water. Then removal of 2 moles results in a change of 20% in the concentration: the change is 2 moles/liter. However, if pool A is at 5 moles per liter originally, and if 2 moles are removed, the *proportion* by which the concentration changes is 40%, although still 2 moles/liter change.

The reason for making this distinction is to lay the groundwork for discussion of the time-sampled simulation of some of the pools in ONECELL.

The decision to simulate growth using a large time-step (one minute) between successive modifications of the state of the cell, entails many difficulties. The principal one is that many of the

pools in the cell exhibit turnover rates far in excess of one per

time step (some turn over in less than a second). The difficulty with

this is that internal to the real cell there are control mechanisms

which operate very rapidly, whereas the most rapid operation in the

simulated cell can only be once per minute. This results in the follow-

ing problem: for pools with high turnover rates, a small difference

between rate of formation and rate of use (or rate of entry and rate

of use) in the real cell is compensated for quickly, and the concen-

tration does not change a great deal. However, such a difference,

when projected over a whole minute in the simulation, results in a

difference which is very large relative to the size of the pool, even

though the proportional difference between the synthesis rate and use

rate may be small.

The procedure which is used in the situations with the behavior

described above is to "sample" their equations for only a portion of

the time step. That is, rather than calculating the amounts of pool

A which would be synthesized and used during one minute at particular

rates, the rates are used to determine how much of pool A will be syn-

thesized or used in the first fraction of one minute (or sometimes,

fraction of a second). This fraction of the full time step is called

the sampling time. By means of a shorter sampling time, the difference

between the amount of a pool used and the amount synthesized is kept

to a reasonable fraction of total amount in the pool.

A critical step in this procedure is the choice of the sampling

time. The two factors relevant to the choice of the sampling time

are sensitivity and stability. Obviously, if the sampling time is

made small enough, the concentration of a pool can be made arbitrarily

stable. However, the pool is then insensitive to changes in the sub-
strates, enzymes, etc. which should affect its rate of formation,
and insensitive to the rates of reactions which consume the pool.
Increasing the sampling time somewhat will result in a pool which re-
sponds, but too slowly. Of course, response times of less than one
minute are automatically excluded from the simulation, but responses
should not be "damped" any more than necessary. On the other hand,
sampling times which are too long do not alleviate the original insta-
bility problem. For example, a long sampling time could "over-exhaust"
a pool - that is, use more of the pool than was there to begin with,
causing a "negative concentration" to result. This is obviously a
situation to be avoided. Short of this, a too-long sampling time
can cause a pool to undergo very high amplitude oscillations, with
the model's control mechanisms only able to over-react and shuttle the
concentration back and forth about its correct value. This wild
oscillation has detrimental effects on other pools of the model, tending
to make them unstable also.

There is another problem associated with using a sampled simulation
of a particular equation. For example, let pool A be sampled for only
one second each time step (i.e., one-sixtieth of the full rate) The
concentration of pool A may be maintained quite well by this procedure.
However, consider the use of another pool, say ATP, necessary for the
formation of pool A. The amount of ATP used during the one-second
sampling interval may be calculated, but that is not the amount by which
ATP should be decremented. The amount of ATP used can be calculated
from the synthesis rate *over the whole time step* of one minute. The
problem of a difference (between rate of synthesis and rate of use)

being exaggerated does not occur at this stage, since no difference must be calculated. Of course, the ATP pool may (in fact, it does) require sampled simulation.

In summary, then, the use of "sampled" simulation of some pools does not affect the rate at which those pools consume resources (that is, other pools). The use of resources is calculated over the entire time step, rather than over the shorter sampling time.

Transport

The model used for describing the transport of nutrients into cells is based on the work of Kepes and Cohen (1962). The model accounts for both uptake and turnover; that is, an exit mechanism is included as well as an entrance mechanism. Entrance is regarded as an energy-requiring enzyme-catalyzed process with Michaelis kinetics; that is, it exhibits a saturation curve characterized by a maximum rate of uptake V_{max} and a Michaelis-Menten constant K_m. The exit mechanism is assumed to be kinetically passive; that is, it shows no tendency toward saturation. The overall model, in the notation of Kepes and Cohen, is given by:

$$\frac{dG_{in}}{dt} = V_{in}^{max} \times \frac{G_{ex}}{K_m + G_{ex}} - k_{ex}G_{in},$$

where G_{ex} represents the concentration of G, the substance being transported, *outside* the cell, G_{in} is the concentration *inside* the cell, V_{in}^{max} is the maximum rate of entry of G, K_m is the Michaelis-Menten constant (the concentration of G_{ex} at which entry is at half its maximal rate), and K_{ex} is the exit rate constant.

In ONECELL, the Kepes-Cohen model is augmented by inclusion of the surface-to-volume ratio as a multiplicative factor in the active transport term. This is consistent with the notion that transport is a function of the membrane surface area, while the internal concentration depends on the volume; thus, as the surface-to-volume ratio increases, the amount of material transported in *per unit volume* should increase.

The energy-supply ATP for the transport is assumed not to be a rate-limiting factor, so ATP does not appear.

There are four separate transport mechanisms in the model: NUC, AA,GLUC, and LAC. Each of these has its own set of transport constants. In addition, the NUC transport model has a modified exit mechanism, which is described in the NUC discussion in Section 5. Kepes and Cohen provide most of the information which has been used in setting the constants for the LAC, GLUC, and AA transport models. Their paper also gives an indication of the degree of similarities and diversities among many different transport systems.

The three environments used in setting rate constants, etc., are glucose minimal medium, casamino acid medium, and broth medium, called environments one, two, and three, respectively. Media of these types are defined in Schaechter, Maaløe, and Kjeldgaard (1958, p. 225), (their media numbers 14,7, and 6, respectively.) An additional medium, lactose minimal medium, is used in simulated shifts. The concentrations assumed in these media are EXT_{GLUC} = 0.4%(W/V) in environments one, two, and three, and 0 in lactose medium; EXT_{AA} = 1.0 (arbitrary units) in environment two, 1.1 in environment three, 0 elsewhere; EXT_{NUC} = 1.0 (arbitrary units) in environment three, 0 elsewhere; EXT_{LAC} = 0.2%(W/V) in lactose medium, 0 elsewhere.

Initial Concentrations

Setting the parameters of the simulation requires knowledge of
the initial state of the pool concentrations in each of three environ-
ments. The particular initial states used are those which occur in a
cell just after a cell division; however, in determining the initial
concentrations of the pools (with some exceptions), it is assumed that
the concentrations remain constant throughout the division cycle during
steady-state exponential growth in a given environment. Since many
of the values which are reported are average *amounts* of a pool *per
cell* in an unsynchronized culture, it is necessary to apply a trans-
formation to obtain the values for the pools as concentrations. The
culture-average figure must be divided by the average volume of the
cells in the culture, assuming that the concentration of the pool is
constant over the cell cycle.

The average volume of a cell in exponential growth, which has
doubled when t reaches 1, and has initial volume V_0, can be calculated
as:

$$\text{avg. vol.} = \frac{1}{1} \int_0^1 V(t)dt = V_0 \int_0^1 e^{t \log_e 2} dt = V_0 \left. \frac{e^{t \log_e 2}}{\log_e 2} \right]_0^1$$

$$\frac{V_0}{\log_e 2} \quad (2\text{-}1) = \frac{V_0}{\log_e 2}$$

If V_0 and the average amount of pool per cell are known, the
concentration can now be calculated as: concentration = avg. amount
per cell/average volume per cell = avg. amount per cell/$(V_0/\log_e 2)$ =
avg. amount per cell $\times \log_e 2/V_0$ = concentration of the pool.

This is the initial value of the pool which is actually used in
determining rate constants and initializing a simulation run. (Of

course, different average amounts of pool and V_0's are used for different
nutritional environments.) During a simulation run, the concentration
of each pool varies as is expected, but tabulation of the average
value and calculation of the average amount per cell (that is, gener-
ating data comparable to the original amount per cell data) indicates
that very little error has been introduced by assuming constant concen-
trations when deriving the transformation procedure.

Sources of Data

Data used in setting parameters for the ONECELL were taken from
a variety of sources. Data on cellular size and composition were
taken largely from Watson (1971), Maaløe and Kjeldgaard (1966), and
Mandelstamm and McQuillen (1968). Data concerning DNA replication
were taken from Helmstetter, Cooper, Pierucci, and Revelas (1968).
Transport data is mainly estimated from Kepes and Cohen (1962) and from
Dean and Hinshelwood (1966). Weinberg used Mandelstamm and McQuillen
(1968) as a reference in assigning values for ATP requirements, but
better estimates are available in Gunsalus and Shuster (1961).

Marr, Painter, and Nilson (1969) provide information on the increase
of cellular volume. Data on the lactose pool and the associated enzymes
is found in Knorre (1969), Kjeldgaard (1967), and in the volume by
Beckwith and Zipser (1970). Ward and Glaser (1970) discuss the
Helmstetter-Cooper DNA-replication model, and in another paper (1971),
discuss the rate of change of volume.

The role of tRNA in RNA synthesis is discussed in Kjeldgaard
(1967) and Williams and Neidhart (1971). Kelley and Schaechter (1968)
provide information on ribosomes. Tempest and Ellwood (1969) deal

with the cell wall pool, and discuss the surface-to-volume ratio.

Section (5) - Equations for NUC

To enable the reader to get an idea of the form of the equations used in simulating the individual pools, the equations for the NUC pool will be given here. It may be viewed as a typical pool - some pools have more complicated equations; some, less.

For the reader who is not familiar with FORTRAN, a few remarks are provided here to enable him to read the description of the NUC equations.

Variables are written in capital letters. They may be subscripted, by enclosing the subscript in parentheses following the name of the variable. For example A(I) means a_i in more conventional mathematical notation, and B(3) means b_3. An equation written in FORTRAN has a special meaning. A = B+C means add the values of variables B and C together, and make that sum the value of A.

Functions are expressed similarly to subscripted variables. In mathematical notation, we denote a function f with argument x as f(x); in FORTRAN, we express it as F(X). Functions may be defined in FORTRAN; for example, f(x) = 3x+2 would be defined as F(X) = 3*X+2, where the * denotes multiplication. After this function F is defined, it can be used as follows (for example):

W = F(5)

Q = F(2.2)

These statements cause W to acquire the value 17(=3×5+2), and Q to become 8.6 (=3×2.2+2).

For further information on FORTRAN, the reader may consult the
IBM 1800 FORTRAN Reference Manual, or any other FORTRAN manual.

The NUC (nucleotide) pool is modeled using a set of equations
which account for NUC synthesized, incorporated into nucleic acids,
recovered from breakdown of mRNA, and transported into the cell (or
out of the cell). Diagrammatically, the NUC pool can be seen in relation
to a subset of the model as shown in Figure 2.5.
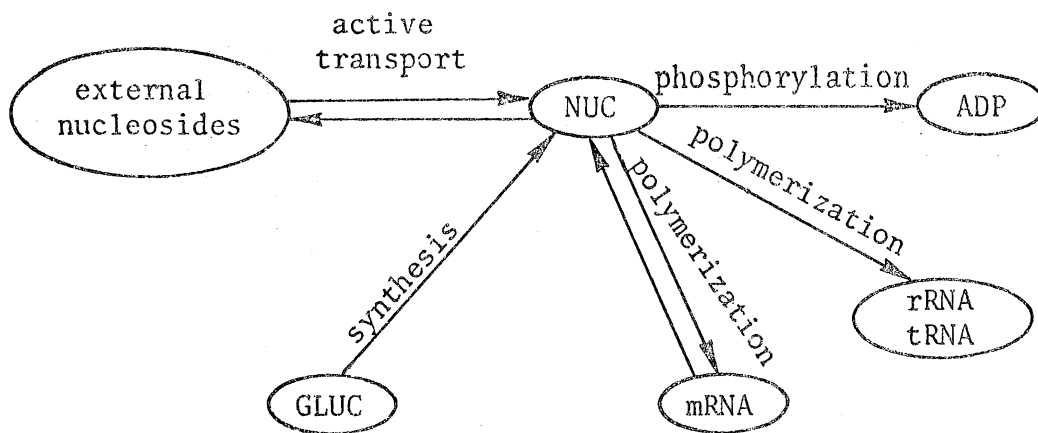
Figure 2.5: Mass-flows involving the NUC pool.

Figure 2.5 shows only the mass-flows involving the NUC pool;
there are other pools which influence the rates at which these flows
occur. The equations which govern the NUC pool can be written in
the form:

(1) $\text{DNUC} = \text{K}(1)*\text{GLUC}*\text{EK}(1)*\text{ATP}*\text{DT}$

$-(2.5\text{E}9/660.)*\text{DDNA} - (1.\text{E}6/660.)*\text{DMRNA}$

$-(2.5\text{E}4/660.)*\text{DTRNA} - (2.\text{E}6/660.)*\text{DRIB}$

$-\text{DPOOL}$

(2) XNUC = (EK(13)*K2(1)*EXT(1)*SUVLR(VOL)/(KM(1)+EXT(1)))

        -KEX(1)*EXNUC(NUC)*NUC)*DT

(3) NUC = VOL/(VOL+DVOL)*(NUC+DNUC+XMAX(XNUC,0.0))

The first term in equation (1) represents synthesis of nucleotides,

considering members of the GLUC pool to be the substrates. K(1) is

the rate curve, expressing feedback inhibition by NUC of its own

synthesis. As was discussed in Section (4) of this chapter, K(1)

is hand-fitted to obtain a monotonically nonincreasing rate curve,

when graphing rate (y-axis) versus NUC (x-axis). EK(1) represents the

concentration of enzymes which are used in NUC synthesis. This con-

centration is affected by repression, as well as by the general level

of protein synthesis in the cell. The ATP term is included to reflect

the need for high-energy bonds in the NUC synthesis reactions. DT

is the time step (60 seconds).

The remaining terms in equation (1) represent loss of NUC due

to its incorporation into macromolecules. DDNA is the amount of DNA

synthesized this time step (per unit volume). The multiplicative

factor (2.5E9/660.) has the following meaning: 2.5E9 (to read 2.5×10^9)

is the estimated weight of one complete DNA chromosome in daltons

(Watson, 1970, p. 99). The 660 divisor was used (by Weinberg, originally)

as the average weight of a NUC molecule. In fact, this is the average

weight of a nucleotide pair, as given by Watson (p.99), and the figure

for NUC has been corrected to 330 in a later version of the simulation.

The DMRNA term represents the change in MRNA per unit volume, and may

be either positive or negative; that is, a small decline in the mRNA

synthesis rate will allow the breakdown of mRNA (which has a half-life

on the order of a minute) to numerically exceed the synthesis, resulting in a net flow from mRNA to NUC (Watson, 1970, p. 452). The multiplier (1.E6/660.) is intended to convert the change in mRNA to the corresponding change in NUC. The terms for tRNA and rRNA, are handled similarly, with molecular weight estimates taken from Watson (1971, p. 85). The DPOOL term represents AMP (adenosine monophosphate) which is phosphorylated to become ADP, some fraction of which becomes ATP within the same time-step. Since one NUC is consumed in producing one ADP, no multiplicative factor is needed.

Equation (2) represents transport of nucleosides from the external medium into the cell. This process cannot bring in nucleosides, of course, unless they are present in the medium, as in nutrient broth. The form of the transport model is basically that described by Kepes and Cohen (1962, p. 187), as discussed in Section (4).

The specific values assigned to the NUC transport model parameters were arrived at in a number of ways. Since NUC is transported into the cell in only one of the environments used to set parameters (environment 3, broth), it was *not* possible to solve a set of simultaneous equations to obtain values for several parameters after fixing only a few and knowing the rate at which transport must proceed in each environment. The only information known is the net rate at which NUC must enter the cell in environment 3 calculated from the rate of use of NUC and assuming a steady-state concentration of NUC internal to the cell) and the external concentration of nucleosides EXT(1) (in arbitrary units, therefore chosen to be 1). Of the remaining parameters - KM,K2, and KEX, two could be set and the third solved for. However, preliminary attempts at this technique showed that

adequate control of NUC could not be maintained, even in a constant environment, by this model with a linear exit term (i.e., a constant KEX times NUC). It was decided that a nonlinear function, called EXNUC, would be introduced, and would be of the form

$$EXNUC(NUC) = 1. + 1.*(NUC-NUC03)/NUC03,$$

where NUC03 is the initial internal concentration of NUC in steady-state growth in environment 3. Thus, the entire exit term, KEX*EXNUC(NUC)*NUC, allows exit to proceed at a rate above or below KEX*NUC, depending on whether NUC is high or low.

The values for KEX and KM are set at .005 (that is, .5% of the NUC pool per second) and .1, respectively. The KM of .1 was chosen to indicate that the external concentration of nucleosides (EXT(1), = 1.) is saturating in environment 3. The KEX is chosen arbitrarily, small enough not to expend a large amount of metabolic energy in "wasted" transport, but large enough to provide adequate control of the internal NUC level. K2 can now be solved for in terms of the other quantities.

Initial values for NUC, called NUC0, must be set in each of the three environments used for calculating the rate curves K(i). Watson, (p. 85), gives a value for the average amount of NUC in a cell growing in environment 1; this is 1.2E7 molecules. Converting this figure to an initial concentration in a newly-divided cell, according to the method described in Section (4) yields 1.2E7*LN2/VOL01, which will be used as the initial value of NUC in environment 1.

The value in environments 2 and 3 are estimated using the observation that RNA formation appears to be nearly saturated with NUC in all three environments, but it is not expected that the NUC pool

is a great deal above the saturation level, since for transport of NUC into the cell to maintain a higher-than-useful level of NUC would require unnecessary use of energy. The implication of these assumptions is that NUC does not change by a large factor within the range of the three environments. However, NUC must rise significantly from environments 2 to 3 in order that repression of NUC synthesis may operate. Therefore, it is assumed that NUC0 in environment 2 is 1.4E7*LN2/NUC03), and in environment 3, NUC0 is set at 2.1E7/*LN2/VOL03)

Equation (3) handles the updating of the NUC concentration after the DNUC and XNUC terms have been computed. It uses VOL, the volume of the cell *last* time step, and DVOL, the change in volume, to arrive at a new volume, and uses NUC (the old value) and DNUC and XNUC to arrive at a new amount of nucleotides. The result is the new concentration, since the equation essentially reduces to:

$$\text{new concentration} = \frac{\text{old volume} \times (\text{old concentration+change in old concentration})}{\text{new volume}}$$

The models of the other pools are of similar forms, and will not be described in detail. The interested reader can consult the listing of the program for details of the other pools.

CHAPTER 3

In this chapter the results of the simulation will be presented, and comparisons will be made with data on real cell growth. The performance of the simulation will be given first in the simulated environments which were used in setting the parameters; second, in shifts between these environments; and third, in an entirely new environment. Most of the output of the simulation will be presented in graphical form, for ease of comprehension. Parameters are set in the ONECELL model using three sets of values. Each of these sets is an estimate of the state vector of the cell growing in a particular environment. For example, one environment represents glucose minimal medium (sometimes called glucose mineral medium or glucose salt medium), composed of glucose and other constituents as given in Schaechter, Maaløe, and Kjeldgaard (1958, p. 225). An estimate is made of the concentration of each of the pools in a newly divided cell which is in exponential growth, in a well-aerated culture. There is no single source of all the data required, so estimates are made as described in Chapter 2, using information from various sources.

The two other steady-state (repetitive) growth environments for which parameter estimates are made are casamino acid medium and broth, also described in Schaechter, et al. (the broth medium is their medium number 6 (p. 225)). These media are discussed in Chapter 2.

Once the parameters have been set, it is possible to run the simulation to determine whether the values of the parameters, together with the form of the model, do indeed produce balanced growth. That is, growth can be simulated over long time periods to see if it exhibits

the characteristics it is supposed to: namely, balanced growth with all pools, doubling in the appropriate time period, and a close agreement of the average pool values with those originally input. The clearest way to present this behavior of the simulation is to follow the concentration of each pool. If the concentration of each pool remains constant throughout the course of the simulation, and if the volume doubles each T, where T is the average interdivision time of real cells in the medium being simulation, then it is clear that the simulation is producing "steady-state" growth. However, these conditions are far too restrictive. A cell in steady-state growth does not maintain all of its concentrations constant throughout its division cycle. In fact, such factors as the surface-to-volume ratio and a nonexponential DNA-replication apparatus make it imperative that the cell *not* maintain constant concentrations. What *is* required of steady-state growth is that the cell volume (that is, the sum of the volumes of the individual cells produced) double every T minutes, and that the *average* concentrations of each pool remain the same from one cycle to the next (or, a somewhat stronger condition, that the concentration of each pool be the same at time t+T as it was at time t.

For glucose minimal medium (called environment 1 in the simulation), the behavior of the simulation is shown in Figures 3.1 through 3.4. Each of these figures follows several concentrations through two complete division cycles. As can be seen from Figure 3.1, there is a wide range of types of behavior--some pools remain relatively constant, while others show large (but regular) changes within the cell division. The units for all the curves are arbitrary, chosen so the curves do not overlap. The top curve, for the NUC pool, shows
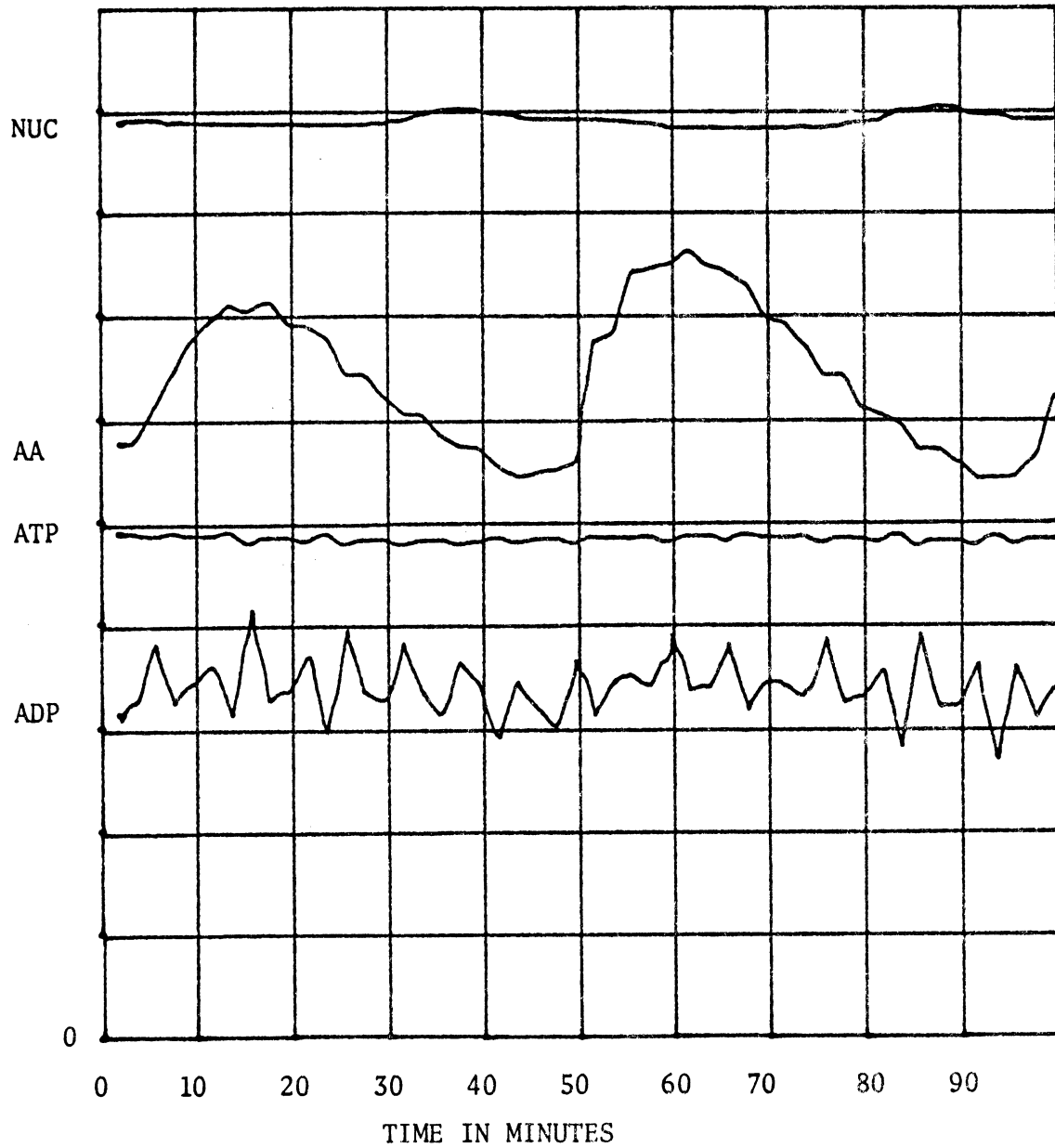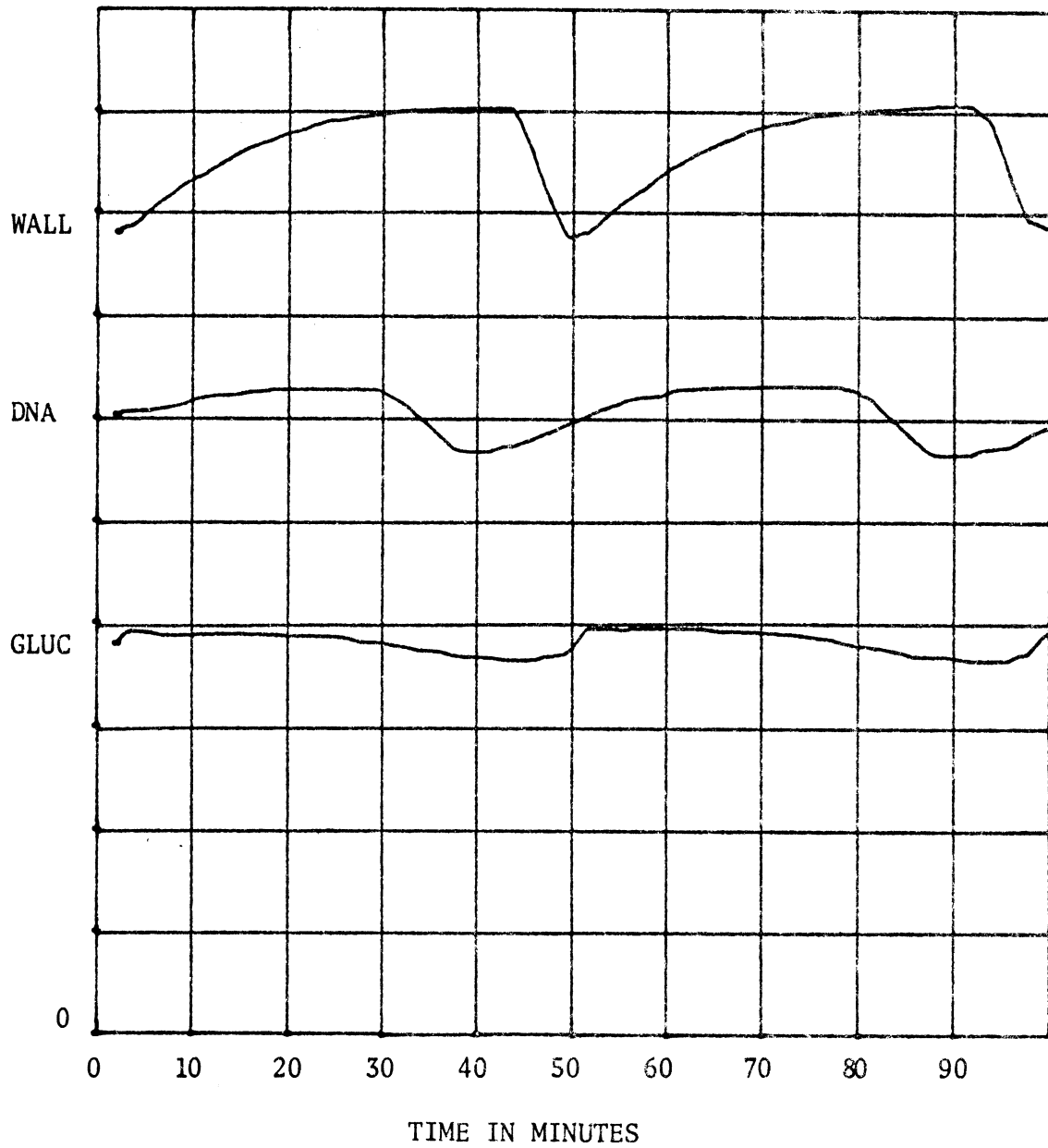
Figure 3.1: NUC, AA, ATP, ADP in Simulated Growth in Environment One. Scale is linear, units for each pool are arbitrary. Sample points are two minutes apart. Recording was begun after one simulated cell division.

that NUC remains fairly steady during much of the cell cycle, rising only for about ten minutes near the end of the 50-minute cycle. These 10 minutes are the period during which no DNA replication is occurring.

The AA (amino acid) curve shows much more change during the cycle. Its fluctuation is mostly attributable to the availability of glucose-pool precursors, which varies with the surface-to-volume ratio, which is a function of cell size (which, of course, depends on the amount of time since the last division, among other things). Although the AA pool attains a higher value during the second cycle, longer runs have shown that the second peak on Figure 3.1 is approximately the maximum attained. The reason the first peak is lower is that the graph begins just after the completion of the first simulated division, and some pools have not quite reached their steady-state growth values. This could be eliminated with a better estimation of the cell's initial state, but it is more practical to begin with a rougher estimate and allow a short amount of running to adjust the values.

The ATP pool shows essentially constant concentration throughout the cycle. The fluctuations are due to the coarseness of the simulations, not to any attempt to model cyclic fluctuations associated with ATP. The ADP pool exhibits similar behavior, but the amplitude of the fluctuations is higher, because the size of the pool is smaller.

Figure 3.2 shows the WALL pool, which represents precursors used in cell membrane and wall formation. The large dip near the end of each cycle results from a high rate of use of WALL to form the septum which will divide the daughter cells. In the ONECELL model in environment 1, it is assumed (in setting parameters) that the septum formation goes on during the last 5 minutes prior to cell division.
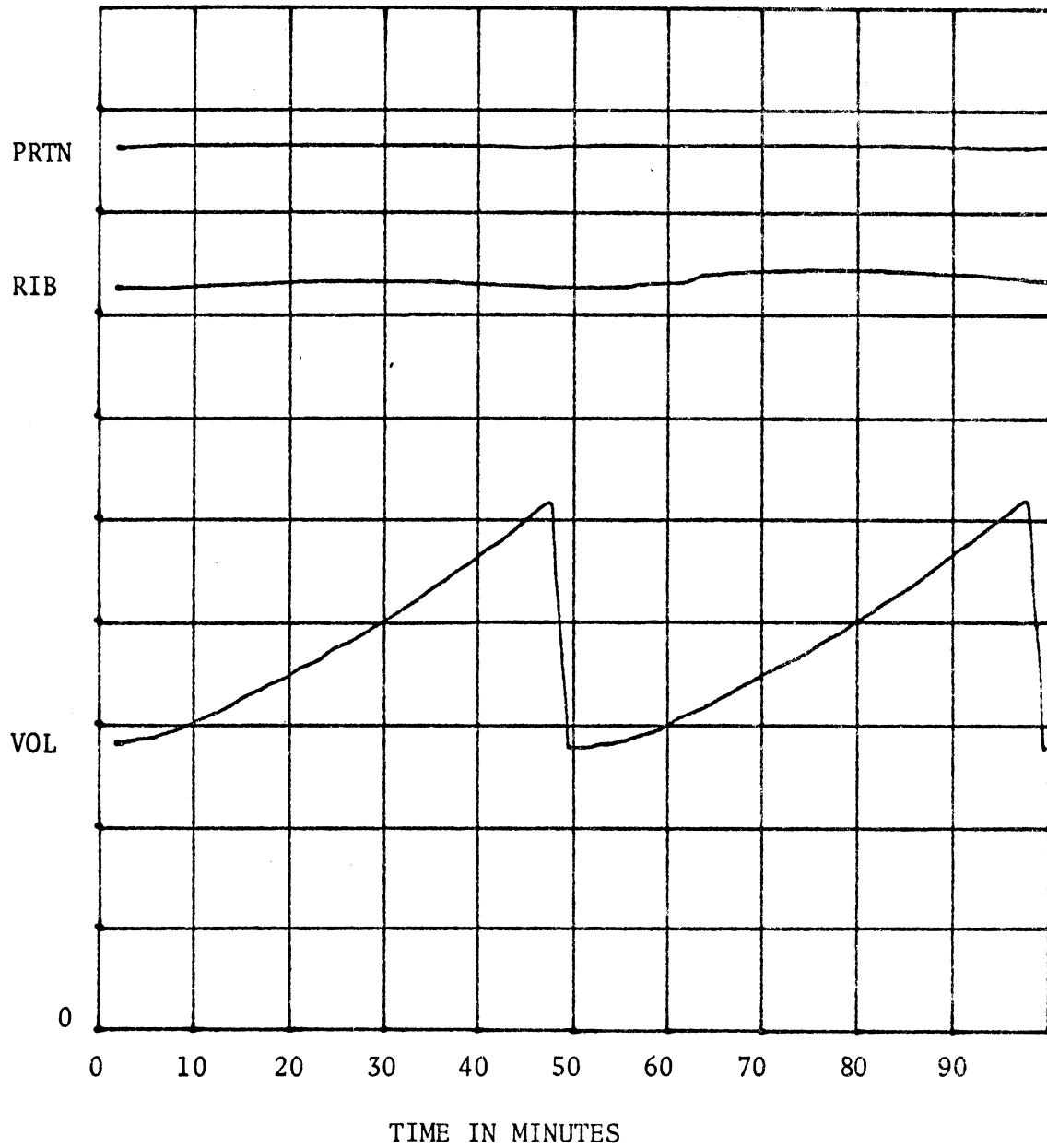
Figure 3.2: WALL, DNA, GLUC in Simulated Growth in Environment One.
Scale is linear, units for each pool are arbitrary.
Sample points are two minutes apart. Recording was
begun after one simulated cell division.

It is also assumed that elongation does not cease during this period, so a higher demand for WALL occurs during the five-minute period.

The curve for DNA represents the concentration of DNA, not the amount per cell. This figure is derived by dividing the amount of DNA (calculated in accordance with the Helmstetter-Cooper model (Helmstetter, Cooper, Pierucci, and Revelas, 1969) by the cell volume. The concentration can be seen to fall near the end of a cycle, at which time no DNA replication is going on, but volume is increasing. When replication is initiated, the concentration increases toward its former value.

The GLUC curve shows the concentration of glucose and the products of the pathways which metabolize glucose (up to the points where they are considered to be NUC, AA, etc.). The variation in the concentration is due to the changing surface-to-volume ratio, which affects the rate at which glucose is transported into the cell.

The protein pool, PRTN is shown in Figure 3.3. The large size of the pool relative to its synthesis rate makes the concentration relatively insensitive to small changes in the synthesis rate, so a very smooth line results. Furthermore, since the protein synthesis rate-determining factor in most growth situations is not concentration of precursors, it displays very little variation with the AA level, unless AA drops far below its value in environment 1.

The RIB curve, representing the number of ribosomes, shows somewhat greater sensitivity to its substrates than does PRTN. In part this is due to the fact that the ribosome synthesis equation does not synthesize a protein component, an RNA component, and then "join" them. Instead, it uses a simpler model in which the precursors for both

Figure 3.3: PRTN, RIB, VOL in Simulated Growth in Environment One.
Scale is linear, units for each pool are arbitrary.
Sample points are two minutes apart. Recording was
begun after one simulated cell division.

protein and RNA synthesis enter. Undoubtedly a more sophisticated model would give a more realistic picture of the mechanism of ribosome synthesis. However, for the present purposes, the simple ribosome equation appears to function adequately.

The curve for VOL represents the volume of the cell. This curve increases in a roughly exponential fashion until cell division, at which point it is halved. It has been argued (see Ward and Glaser (1971)) that the increase in volume within the cycle is linear, not exponential, and that the slope doubles at a particular point (associated loosely with the start of a DNA replication round). However, they indicate that their data could not distinguish stepped exponential increase from the stepped linear increase. In any case, the VOL curve appears to be a reasonable description of the cell volume in environment 1.

Figure 3.4 shows the MRNA pool (among others). It shows strong fluctuations because it is a small pool, and undergoes rapid turnover. These factors mean that when its synthesis rate falls, with the breakdown of MRNA going on at a high rate (half-life on the order of a minute (Kjeldgaard, 1967, p. 452)), the pool shows a sizeable change in concentration. There are several factors which can influence the MRNA synthesis rate. The principal ones influencing this curve are DNA and AA. The AA, while not a substrate for MRNA synthesis, is observed to affect its rate, as discussed in Maaløe and Kjeldgaard (1966, pp. 125-163). The DNA concentration enters the MRNA synthesis equation so that when multiple copies of a gene are present, its rate of transcription can increase. Of course, the ONECELL model does not consider the replication of individual genes, but using the DNA
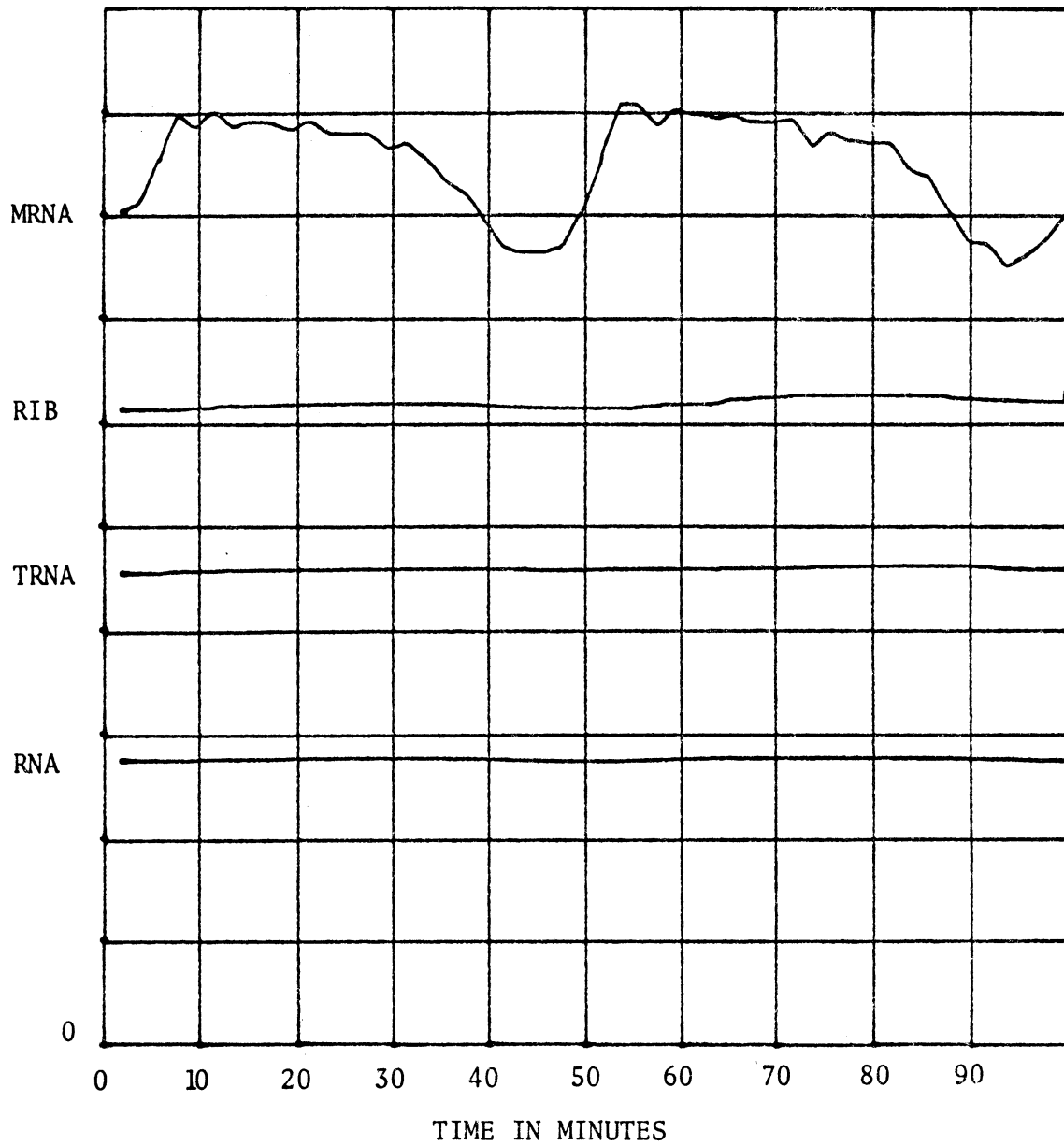
Figure 3.4: MRNA, RIB, TRNA, RNA in Simulated Growth in Environment One.
RNA is a weighted sum of MRNA, RIB, and TRNA. Scale is
linear, units for each pool are arbitrary. Sample points
are two minutes apart. Recording was begun after one
simulated cell division.

concentration provides an approximation; it is as if genes for each pool were "spread out" evenly over the entire chromosome. The shortcomings of this approximation are evident, but it is a necessary simplifying assumption at this stage of the model development.

The RIB curve is shown again here so that all the RNA fractions can be viewed together.

The TRNA pool is even steadier than RIB, because the model assumes that the effect of substrates on the rate of TRNA formation is less than first order. TRNA, unlike RIB, is present in lower *concentration* in rich environments (like casamino acids or broth) than it is in glucose minimal medium (Maaløe and Kjeldgaard, 1966, p. 90). It varies with growth rate much less than RIB does.

The curve given for total RNA does not represent a state variable--it is a weighted sum of the RNA components of the model. Its behavior is clearly dominated by the rRNA and tRNA, giving it the appearance of a constant value.

The graphs presented for growth in glucose minimal medium (Figures 3.1 through 3.4) show the behavior of the simulation during a brief period near the beginning of a run. In fact, the behavior shown in this period is representative of the behavior if growth is continued indefinitely. The longest runs made, allowing some thirty cell divisions, showed no trends toward higher or lower values; that is, after a few generations, the simulation can be seen to exhibit behavior with the one-generation time period, and some variations with a period of three or four generations, but no tendencies for the simulation to drift, or to oscillate unstably, have been observed.

(This is, of course, the end result of much work on the individual components of the model to coordinate them in such a fashion as to produce this stability.)

In casamino acid medium and broth medium the same sort of stable behavior has also been obtained. The concentrations of the major pools during growth in these media are presented in Appendix A (Figures A.1 through A.4 are for casamino acid medium (environment 2) and Figures A.5 through A.8 are for broth (environment 3). Most of the pools behave similarly, with the exception of those directly affected by transport of NUC and AA into the cell.

Shift-Up

The behavior of the ONECELL simulation can be tested by subjecting it to a change in the growth medium surrounding it. That is, the simulation is allowed to progress ("grow") in one medium for some period of time, then "shifted" to a new medium *without changing any variables but those representing the growth medium outside the cell.* This is designed to simulate a laboratory experiment in which a culture growing in glucose minimal medium is diluted into broth medium, for example. This shift-up (*up* because the new medium is richer in nutrients and permits faster growth) will be called a shift from environment 1 to environment 3 (broth).

The fact that the simulation can grow stably in environments 1 and 3 *does not* imply that the simulation will be able to adjust all of its variables in such a way as to attain the stable growth situation characteristic of environment 3. That is, the abrupt change in

conditions could cause one or more variables to grow without bound, "dragging" the other variables with it. Alternatively, a stable growth condition could be found in which the cell would remain, exhibiting periodic behavior, but different from the growth normally exhibited in environment three. Further, even if the simulation were able to attain the proper growth in environment three after the shift, its behavior during the adjustment period immediately following the shift might differ widely from the behavior observed in real cells during a shift-up.

None of the unpleasant possibilities just mentioned has occurred. That is, the simulation does properly shift from environment one to environment three. The agreement of the simulated shift results with those reported in Maaløe and Kjeldgaard (1966) is good.

The first attempts to shift the simulated cell from one environment to another did not work. After some investigation, however, it was discovered that the trouble lay in the mRNA synthesis equation. The difficulty was that a factor in the equation was being normalized by dividing it by MRNA0, the initial concentration of MRNA *in each environment*; that is, a different value was used in environment one than that used in environment three. This was easily correctable merely by eliminating the divisor entirely, since its presence was only to scale the concentration for convenience of comparison with the initial value. It should be clear that this error causes no difficulty until a shift is attempted, but in a shift gives rise to a completely incorrect result. The elimination of this term did not represent a refitting of the simulation parameters, but rather a "debugging" correction. After this bug was corrected, the shift progressed normally.

Figure 3.5 shows the behavior of several outputs of the simulation.

This figure is a logarithmic plot of the amounts (not concentrations)

of RNA, DNA and the total cellular mass as computed in the simulation.

The vertical axis is ruled in lines whose distance apart corresponds

to one doubling of amount. The units of the three variables plotted

are scaled so that the three lines coincide over the pre-shift period,

which represents about two doubling times in the pre-shift medium.

Each point is the sum of three values, each value generated by a

separate run of the ONECELL simulation. The runs were done so that

the time-step recording was begun in the first run   9 minutes after

a division, in the second run, 25 minutes, and in the third run, 41

minutes after a cell division. This procedure is an attempt to elimi-

nate the effect of cell age on the behavior, since it is desired to

compare the output with the growth of a culture which is not synchronized

The crudeness of representation (i.e., three cell-ages) is clear,

but the length of computer time required for each run is extensive

(on the order of an hour on a very slow machine), so it was decided

that three samples were an appropriate number. The time interval

between consecutive points plotted is four minutes in all simulated

shift graphs.

First it should be noted at the time of the shift, the three

lines begin to diverge, then gradually become parallel (but not

coincident). That they are parallel indicates that they are doubling

at the same relative rates (higher, of course, than the rates at

which they doubled in environment one prior to the shift). That they

are not coincident means that the relative proportions of RNA, total
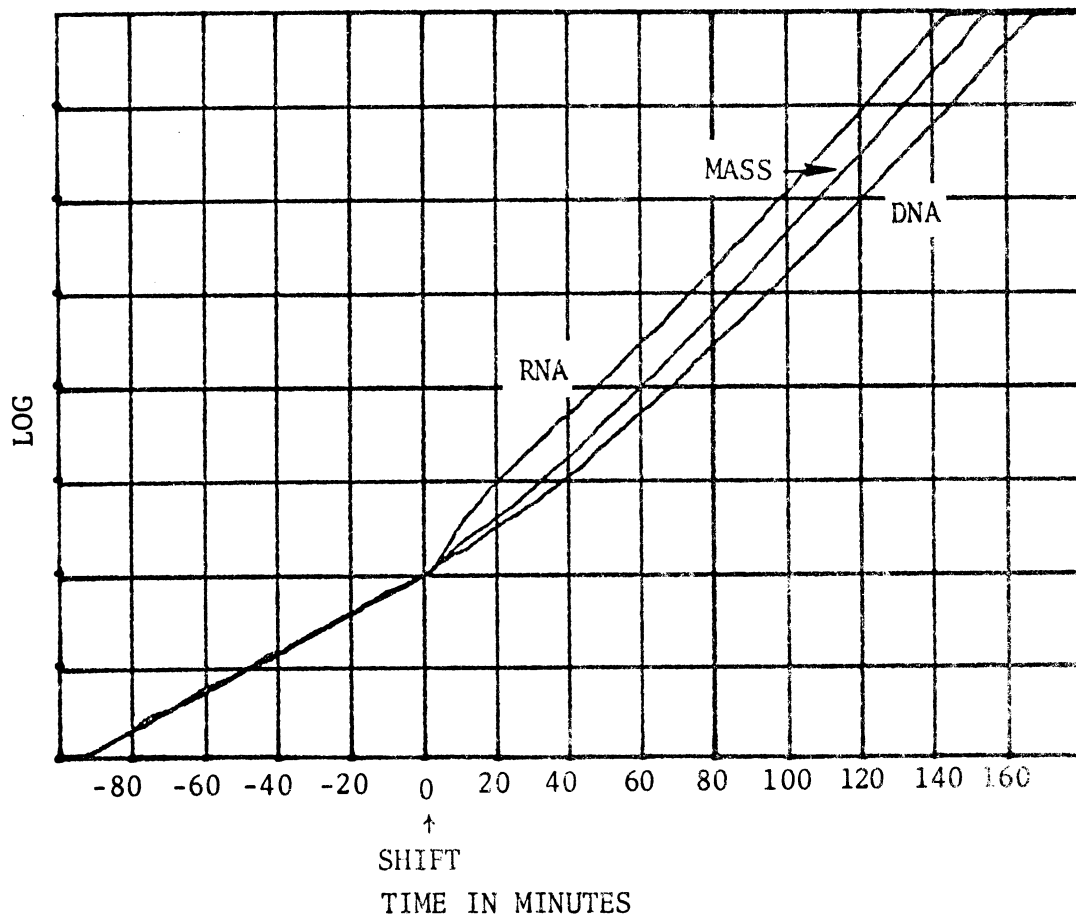
Figure 3.5: RNA, DNA, MASS in Simulated Shift-up from One to Three. Logarithmic plot of amounts - units adjusted so curves for balanced growth (pre-shift) in environment one coincide. Distance between horizontal lines corresponds to one doubling. Each point is a sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.

mass, and DNA in the cell in environment three are different from the proportions in environment one. The different shapes of the curves reflect the difference in the ways they respond to the change in environment.

To see the extent of agreement of the simulated shift results with actual shift experiments, Figure 3.5 should be compared with Figure 3.6. This latter figure is reproduced from Maaløe and Kjeldgaard (1966, p. 100). Their figure is the observed behavior of a colony of *Salmonella typhimurium*, and displays the characteristic pattern of changes which is common to many bacteria. In fact, Kjeldgaard (1967, p. 47) says of a reproduction of the same figure,

> "The observed pattern seems to be universal in character, and is exactly reproduced by *Escherichia Coli* strains both of the stringent and the relaxed type (see below), as well as by a Gram-positive organism like *Bacillus megaterium* (Sud and Schaechter, 1964)."

The agreement between the two figures (simulated growth versus real growth) appears to be good. In both, the RNA curve quickly attains a slope higher than that of either environments one or three, then falls to its eventual value in environment three. The mass curve in Figure 3.5 is directly comparable to the optical density curve of Figure 3.6, as indicated by Schaechter, Maaløe, and Kjeldgaard (1958, p. 594)

> "The values of mass/cell are expressed as the optical density at 450 mμ (1 cm. path) given by a suspension containing 10^7 cells/ml. The optical density was found to be proportional to the dry weight, irrespective of the cell size."

The DNA curves both remain close to the old rate for some period of time after the shift, then increase to approximately their ultimate
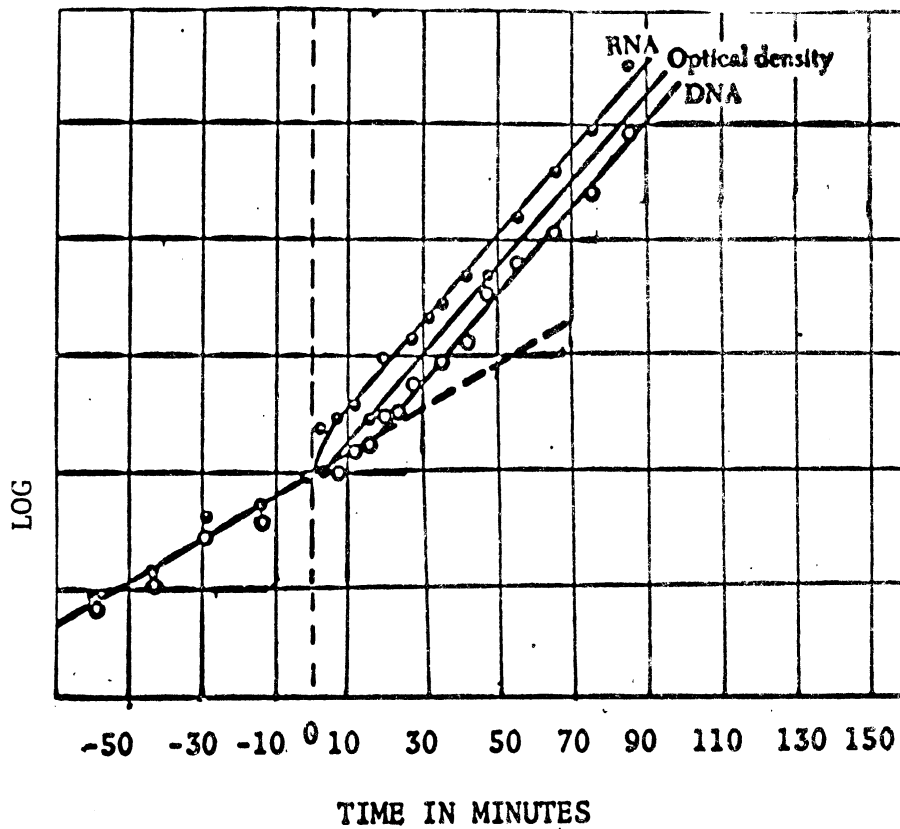
Figure 3.6:   RNA, DNA, Mass in Shift-up Experiment with *Salmonella
typhimurium*. Culture was shifted from glucose minimal
medium to broth (at 37°) at time 0. Optical density
(proportional to mass), RNA, and DNA content were fre-
quently determined. Logarithms of measured values are
plotted, transposed so as to make the curves representing
balanced growth in minimal medium coincide. The distance
between horizontal lines corresponds to one doubling.
(After Maaløe and Kjeldgaard (1966, p. 100)).

rates.

Figure 3.7 is the same type of plot as Figure 3.5, showing total cellular mass against the number of individuals cells in the simulated culture. The sawtooth nature of the curve is due to the fact that only three cells are used to represent the culture, so when one of these divides, there is an abrupt jump in the number of cells. The curve can be fit very well by two straight lines, intersecting at time 80 or 90, and should be viewed as a jagged approximation to these two lines. This figure should be compared to Figure 3.8, which displays mass (optical density) versus colony counts (an index of the number of viable cells). In both figures, the number of cells continues to increase at the old rate for a period in excess of an hour (about 70 minutes in Figure 3.8 versus about 80 or 90 in Figure 3.7). This time period represents the length of time necessary for an increased rate of initiation of DNA replication to have an effect on the rate of division.

The mass line and count line have not yet become parallel in Figure 3.7, but longer runs have shown that they become parallel within the next few generations.

Graphs of the behavior of other pools during the shift from glucose minimal medium to broth are included in Appendix A (Figures A.9 - A.13). In addition, a shift from environment one to environment two (casamino acids) was simulated, and the results are presented in Appendix A (Figures A.14 - A.20). The shift followed the same general pattern as the shift to broth, and the simulated cells achieved stable growth in environment two.
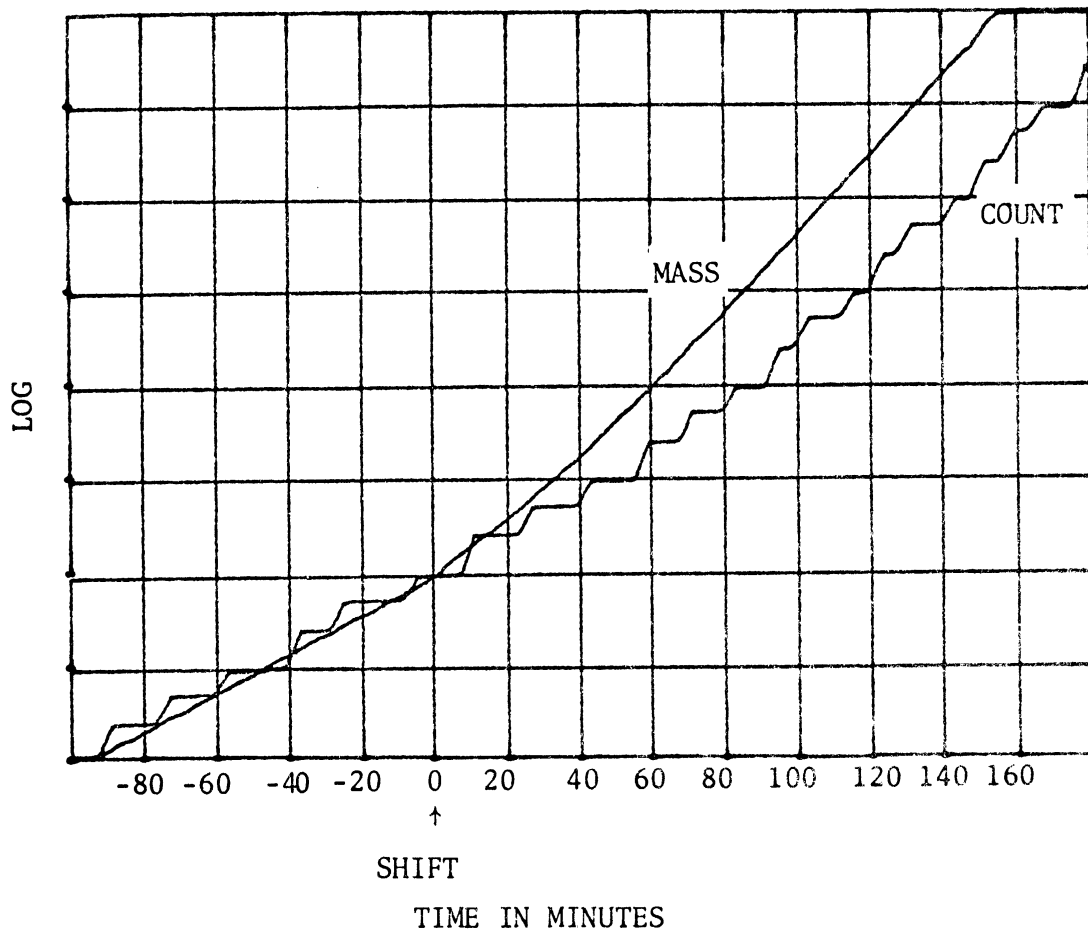
Figure 3.7: Mass and Cell Count in Simulated Shift-up from One to Three. Logarithmic plot of simulated total cellular mass and number of cells with units adjusted so curves for balanced growth (pre-shift) in environment one coincide. Distance between horizontal lines corresponds to one doubling. Each point is a sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.
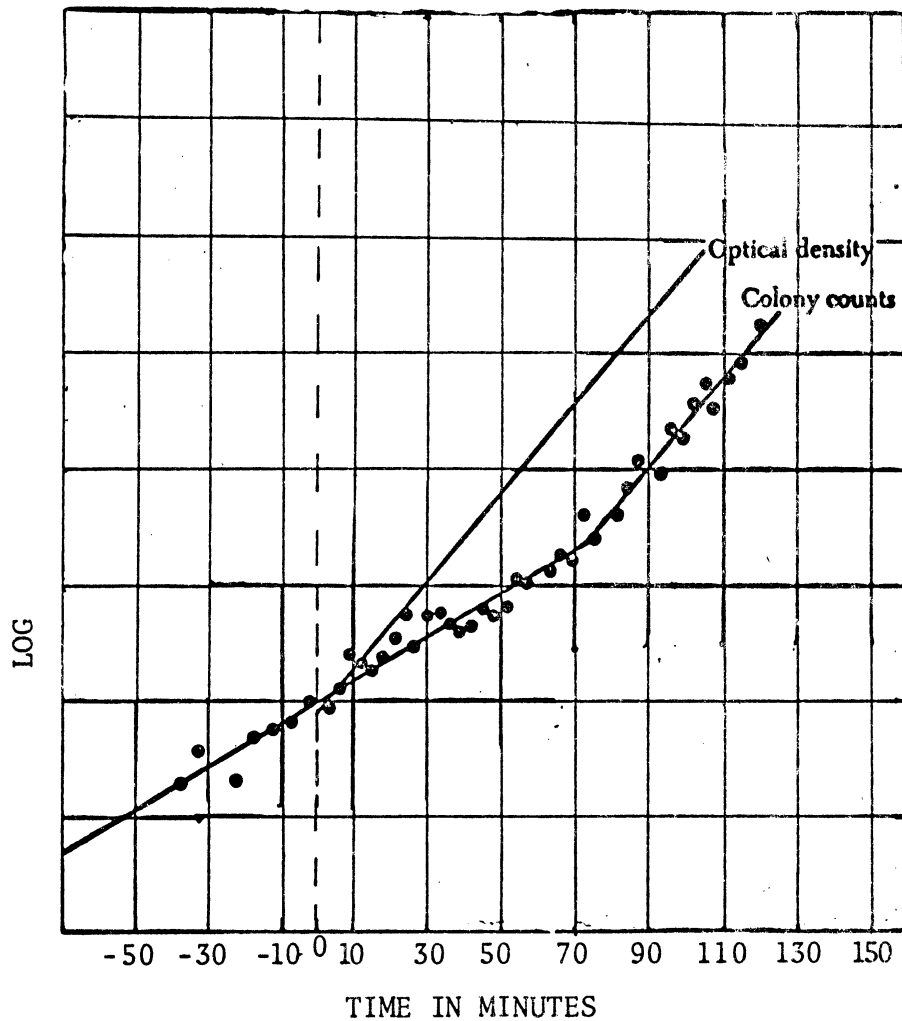
Figure 3.8: Mass and Viable Counts in Shift-up Experiment with
Salmonella typhimurium. Culture was shifted from glucose
minimal medium to broth (at 37°) at time 0. Optical
density (proportional to mass) and viable counts were
frequently determined. Logarithms of measured values
are plotted, transposed so as to make the curves
representing balanced growth in minimal medium coincide.
The distance between horizontal lines corresponds to
one doubling. .(After Maaløe and Kjeldgaard (1966, p. 100)).

Shift-Down

The simulation of two shifts down was attempted. The first is

a shift from broth to glucose-minimal medium. The results of this

shift are presented in a series of graphs of the same form as those

for the shift-up experiments. At this point a small additional sophis-

tication was introduced into the plotting routine which produced the

computer-generated figures shown here. Each plotted value represents

a *weighted* sum of the three values generated by the three ONECELL

runs (each run representing a cell shifted at a different age). The

weighting is done by dividing each cell by its volume at the time

recording of points (on the disk, for eventual plotting) is begun.

The effect of this is to assign larger weights to small cells (at

the start of recording), and to assign smaller ones to the (initially)

large cells. This weighting should compensate for the fact that in

an exponentially growing culture that is *not* synchronized, there are

more small cells than large ones at any instant of time. The weights

in effect cause the single small cell in the simulation run to repre-

sent more cells than the large cell does. Of course, the small cell

becomes large, and the large cell becomes two small ones, but this

does not (and should not) change the weighting which is done.

Figure 3.9 shows the responses of DNA, RNA, mass, and number of

cells to the change in environment. The number of cells can be seen

to continue to grow at the pre-shift rate for about an hour (an

implication of the DNA replication and cell division model employed).

The DNA rate falls off after about twenty minutes. The rate of increase

of mass decreases, and the rate of RNA formation drops sharply.

The behavior of the simulation can be compared with the behavior
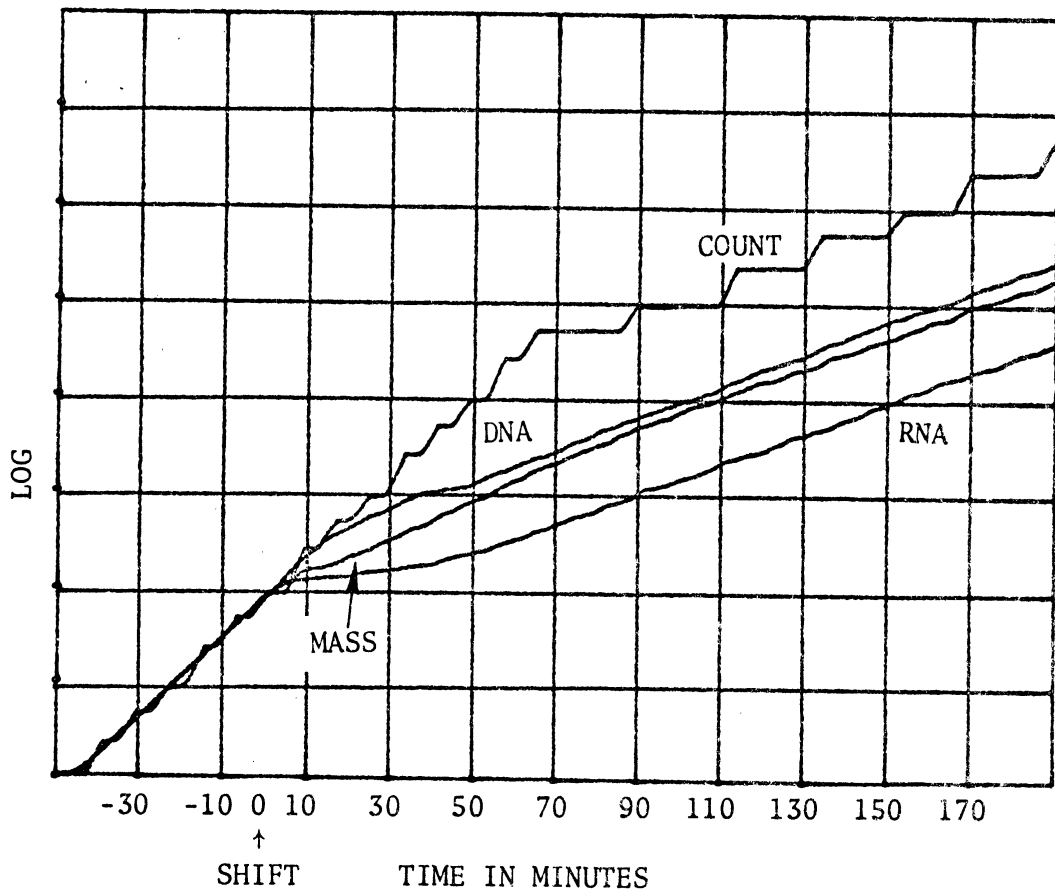
Figure 3.9: RNA, DNA, Mass, and Count in Simulated Shift-down from Three to One. Logarithmic plots of amounts - units adjusted so curves for balanced (pre-shift) growth in environment three coincide. Distance between horizontal lines corresponds to one doubling. Each point is a weighted sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.

of a real culture, as shown in Figure 3.10, reproduced from Maaløe
and Kjeldgaard (1966, p. 118). There is some general similarity, but
there are many differences between the behavior of the simulated culture
and that of the real one. The number of cells (colony counts) in both
figures continues at the old rate after the shift, but does so for a
longer period in the simulation. The DNA replication rate appears
to agree well with the experimentally-determined one. The rate of
mass increase (heavily influenced by protein) does not fall to zero
in the simulation as it appears to in the real culture. Also the
rate of increase of RNA, while falling to nearly zero, does not go
negative, as it does in the real culture.

Some explanations for these differences between the real culture
and the simulated one can be found upon examination of the model.
The overall picture appears to be that the simulated cells are not
"starved" of NUC and AA to a sufficient degree when the shift occurs.
That is, for example, net production of protein, which is virtually
eliminated at the time of the shift in the real culture, can continue
(though at a lower rate) in the simulated culture. This is evident
in Figure 3.11. The fact that AA drops to about 25% of its usual con-
centration in environment one does not cause a sufficient reduction
in the rate of protein synthesis. This may indicate one of several
possibilities: (1) the dependence of protein synthesis on amino acid
concentration has been made too small in the simulation, (2) the amino
acid pool does not fall low enough after the shift, or (3) some other
protein-synthesis rate-limiting factor, such as ATP, does not fall
as far as it should. It is likely that a combination of these factors
is responsible. The possibility that ATP (shown in Figure 3.12) is at
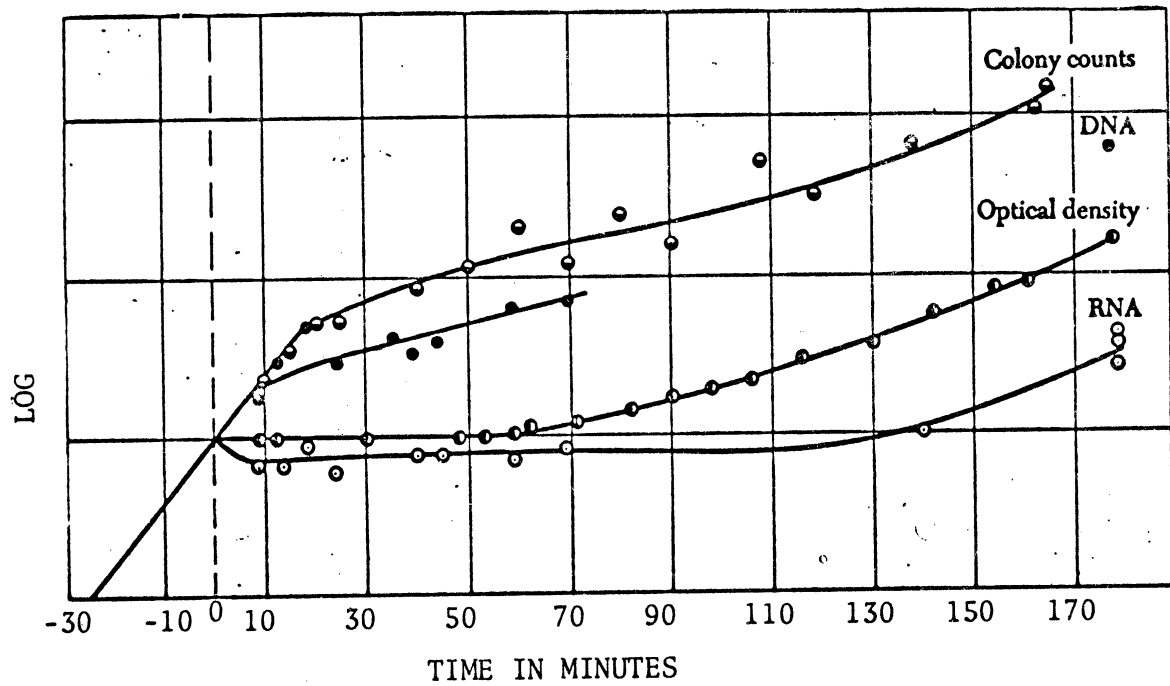
Figure 3.10: Shift-down Experiment with *Salmonella typhimurium*.
Culture was shifted (at 37°) from broth to
glucose-minimal medium at time 0. Optical density,
viable counts, DNA, and RNA were frequently determined.
The logarithms of the measured values are plotted
against time, and all values are transposed so as to
make the curves representing balanced growth in broth
coincide. The distance between horizontal lines corre-
sponds to one doubling. (After Maaløe and Kjeldgaard
(1966, p. 118)).
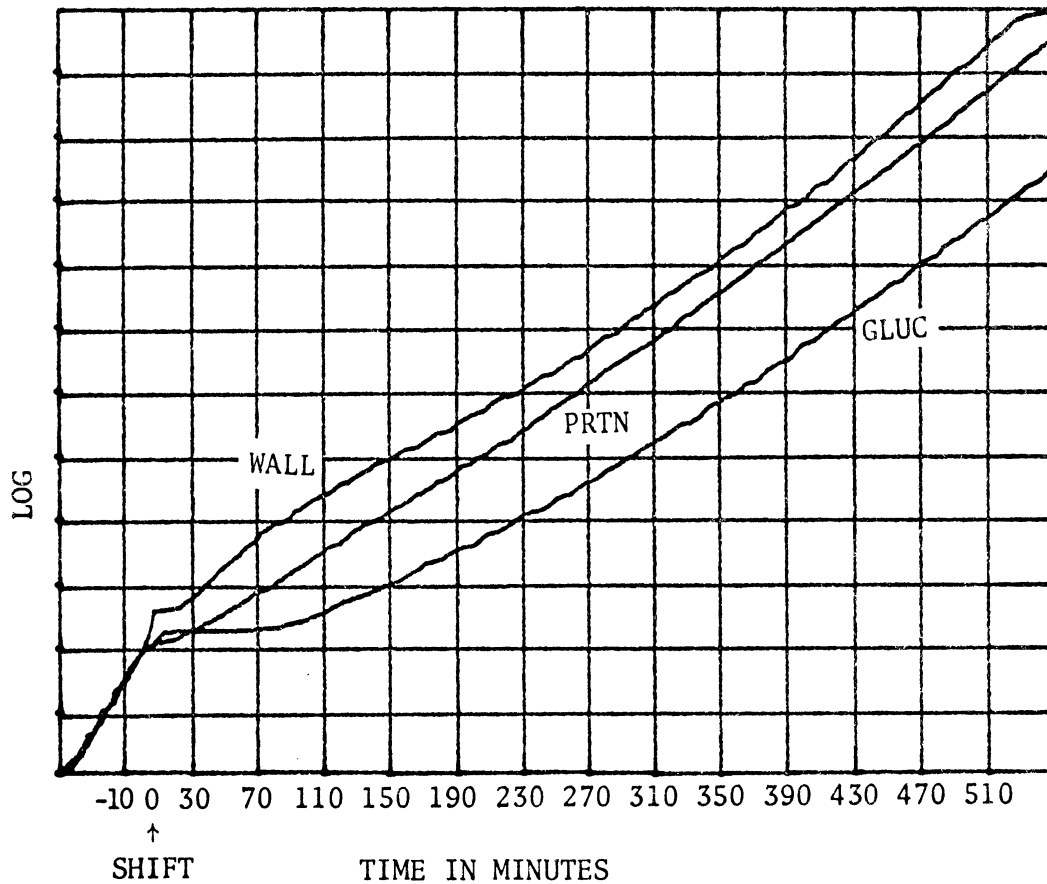
Figure 3.11: WALL, PRTN, and GLUC in Simulated Shift-down from
Three to One. Logarithmic plots of amounts - units
adjusted so curves for balanced (pre-shift) growth
in environment three coincide. Distance between
horizontal lines corresponds to one doubling.
Each point is a weighted sum of three values, one
calculated in each of three ONECELL runs. The runs
differ in the age of the cell at the time of the shift.
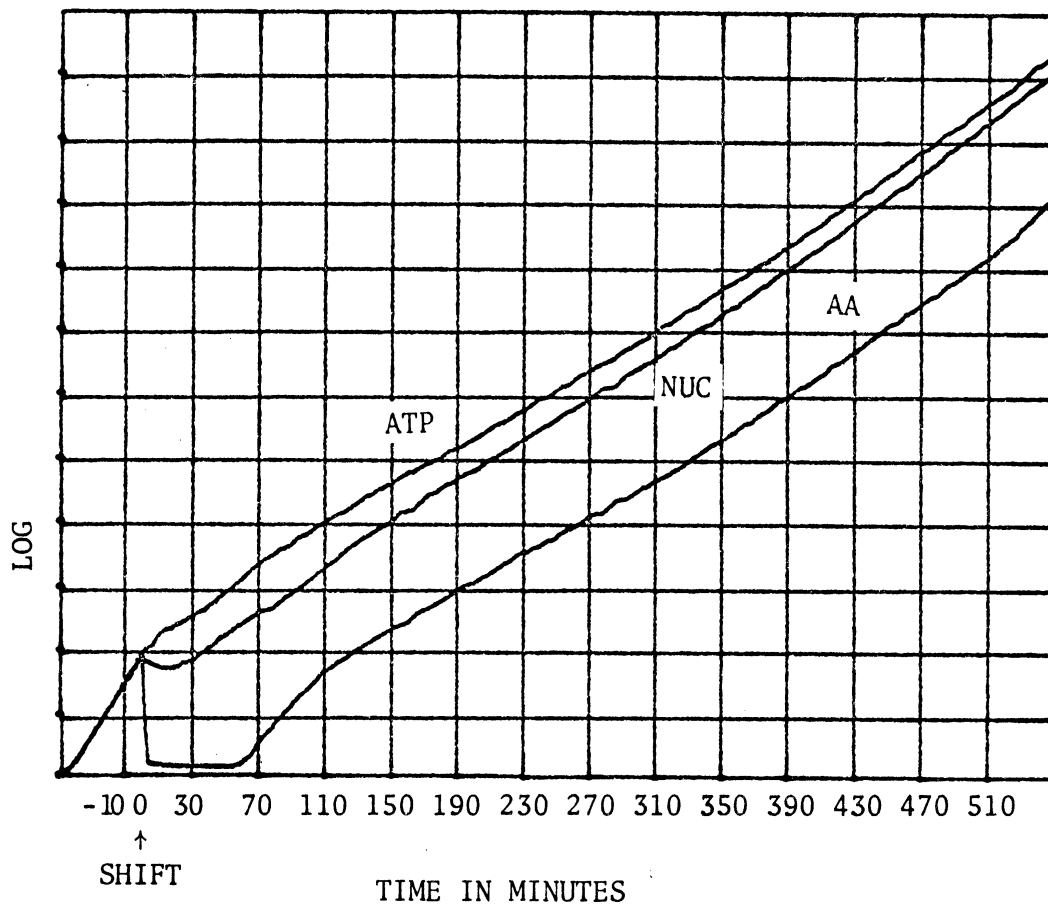
Figure 3.12: ATP, NUC, and AA in Simulated Shift-down from Three to
ne. Logarithmic plots of amounts - units adjusted so
curves for balanced (pre-shift) growth in environment
three coincide. Distance between horizontal lines
represents one doubling. The AA curve dips below the
x-axis, but is shown just above it for clarity.
Each value is a weighted sum of three values, one
calculated in each of three ONECELL runs. The runs
differ in the age of the cell at the time of the shift.

fault is a strong one (and a new model for ATP-ADP-NUC has been
developed and programmed for future introduction into the simulation).

The synthesis of RNA suffers from a similar problem; synthesis
goes on at a very low rate, but in the real culture, a net loss of
RNA is occurring instead. The inability of the simulated cells to
reproduce this phenomenon can be blamed on the form of the RNA models.
The MRNA equation allows for decay of mRNA when synthesis is slowed,
but neither the RIB nor TRNA equations do (see Figure 3.13). In the
real culture, some ribosomal or transfer RNA must be breaking down,
since breakdown of the total amount of mRNA would not be enough
to explain the dip observed in the RNA curve. Tuning of the simulation
could not eliminate the discrepancy; a better model of the rRNA and
tRNA pools is required. Some tentative work has been done on a new
rRNA model, but no final form for a new equation has been arrived at.

The number of cells (as determinable by colony counts) continues
to grow at the pre-shift rate for too long in the simulation (Figure
3.9 vs. Figure 3.10). The explanation for this appears to lie in the
way in which the DNA replication-cell division apparatus is connected
to the rest of the metabolism. As explained in Chapter 2, the rate of
DNA replication is postulated (by Helmstetter, et al. (1969)) to go
on at a fixed and maximum rate in both environments one and three.
In growth conditions where the cells grow even more slowly than in
environment one, in particular when the doubling time T is greater than
60 minutes, Helmstetter, et al., postulate that the rate of DNA repli-
cation falls off in such a way that the time to replicate the chromosome
is 2/3T, and the time to division after completion of a replication
round is 1/3T. Of course, the DNA replication rate for slow growth
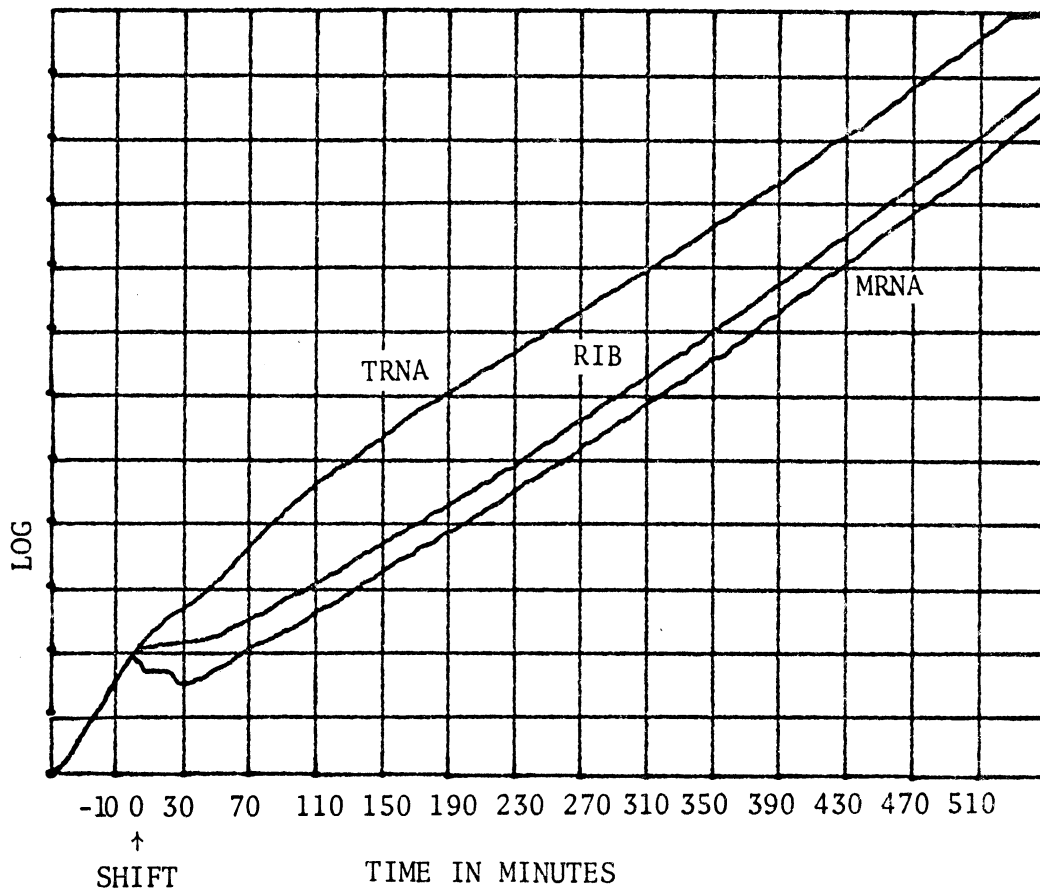
Figure 3.13: MRNA, RIB, TRNA in Simulated Shift-down from Three to One. Logithmic of measured values are poltted, transposed so as to make the curves representing balanced growth in minimal medium coincide. The distance between horizontal lines corresponds to one doubling. (After Maaløe and Kjeldgaard (1966, p. 100).

cannot be incorporated in the simulation in this form - a cell does
not "know" what its generation time T will be; rather, its DNA
replication rate determines what its generation time will be. The
assumption is therefore made that the rate of DNA replication is a
function of the level of the NUC pool, since NUC provides the precursors
for the replication of DNA. When NUC is high (i.e., its value in
environments one, two or three in steady-state growth), the DNA repli-
cation goes on at its maximal rate. However, when NUC is low (less
than 90% of its concentration in environment one) the DNA replication
rate (called DNART) is linearly proportional to NUC.

The problem with the number of cells can now be seen to be that
the simulated rate of DNA replication (DNART) does not fall off as it
must after the shift, so that the cell divisions continue at the full
rate until the effect of less frequent initiations comes into play
(one hour after the shift).

The causes for this might be that (1) the replication rate
is not adequately represented as a function of NUC, or must be a
nonlinear function of NUC, or (2) the level of NUC does not fall as
far as it should in the shift from broth to glucose minimal medium.
Figure 3.12 shows that NUC falls somewhat during the shift, but not
a great deal, and it recovers quickly. The fault is probably a com-
bination of both (1) and (2) above. There is experimental evidence
(Ward and Glaser, 1970) that the DNA replication rate is *not* inde-
pendent of protein synthesis, so that it is expected that an improved
model of the DNA replication mechanism will be required. This new
model should certainly be able to account for the observed reduction
in cell division rate after a shift-down.

In assessment of the behavior of the simulation in the shift from broth to glucose minimal medium, it can be said that the simulation was able to shift, that is, the new steady-state growth condition was attained, but that the dynamic behavior of the pools during the shift did not fit the data on real cultures nearly as well as was the case in the shifts up from glucose to casamino acids or broth.

A shift down from casamino acid medium to glucose medium was also simulated. The results of that shift were similar to those of the other shift down, and graphs of the simulated shift are given in Appendix A (Figures A.21 - A.25).

Shift to Lactose

The last type of shift experiment performed is a shift from glucose minimal medium to lactose minimal medium. This is a severe test of the simulation, since the simulated cell is essentially deprived of a carbon source (which is needed to provide both carbon skeletons and energy). The cell is initially unable to transport or utilize lactose. In the absence of glucose and the presence of lactose, synthesis of a set of enzymes for transport and utilization of lactose is induced.

The inducible lactose system is a widely studied system, as can be seen in the book *The Lactose Operon* (Beckwith and Zipser, 1970). The properties of the system which are modeled in ONECELL are only the simplest and among the first discovered. Insofar as the present form of the simulation is concerned, there are only a few properties of the lactose system that are important. First, there are about 3,000 molecules of β-galactosidase (one of the LAC enzymes, which

catalyzes the breakdown of lactose to glucose and galactose) when an
E. coli cell is growing exponentially in lactose medium (Watson, 1970,
p. 438), and only a few molecules per cell when growing in glucose
(unadapted to lactose). Second, the transport-facilitating enzyme
(β-galactoside permease) and the lactose-splitting enzyme (β-galacto-
sidase) are induced together (although not in the same amounts).
Third, the induction of the enzymes occurs when a culture is shifted
from glucose minimal medium to lactose minimal medium.

The "inducer exclusion" effect of glucose in preventing LAC
induction when glucose is present (Magasanik, 1970, p. 190) is
modeled in ONECELL, however, the transient repression by glucose
(Magasanik, p. 194) is not modeled. Thus the present model should
not be used to examine the growth of a lactose culture to which glucose
is added. Furthermore, catabolite repression by glucose, as discussed
in Knorre (1969, p. 229) is not modeled, so no oscillatory behavior
in the induction of β-galactosidase will be observed. The simple form
of the model for lactose induction and use makes it possible to solve
explicitly for an entry rate constant and a rate constant for conver-
sion of lactose to glucose and galactose, given a value for the internal
concentration of lactose in steady-state growth on lactose. It is
assumed that the galactose utilizing system (which is also inducible)
would not constitute a bottleneck, and that the internal glucose
concentration ultimately achieved in steady-state growth on lactose
was the same as is found in environment one (glucose minimal medium).
This latter assumption is based on the observation that fully adapted
cells grow at about the same rate on lactose minimal medium as on
glucose minimal medium (Dean and Hinshelwood, 1966, p. 81).

The test of the simulation is whether it can survive the "shock"
of glucose deprivation and recover to achieve normal growth in lactose
medium.

The shift to lactose progressed well for two of the three cell
ages shifted; however, the third revealed a situation which had not
been taken into account in the implementation of the DNA replication-cell
division model. When the simulated cell was shifted during the period
in which a round of DNA replication had just begun, but the cell
division which normally follows the completion of the last replication
round had not yet occurred, the simulation eventually encountered a
situation not considered in the formulation of the program logic.
The situation occurs when a round of replication is completed before
the division which follows the preceding round has occurred. The
program logic had only been designed to keep track of completion of
one replication round at a time. It employs a logical variable ITRM
which is set (to 1) when a round is completed, and reset (to 0) when
division occurs. It was a simple matter to add a new logical variable
ITRM2 which "tells" the program that the newest round has been completed,
and that the variable indicating a completion should not be reset
when the cell division occurs. This change is not a change in the
assumptions or form which the model is *intended* to have, but only
a correction to bring about the performance that the model was already
presumed to have. In this sense, it should not be viewed as a
refitting of the model, but as a correction to the program logic.
The correction of this situation brought the performance of the cell
shifted at the last stage of its division cycle into line with the
cells shifted in the two earlier stages. Figure 3.14 is a graph
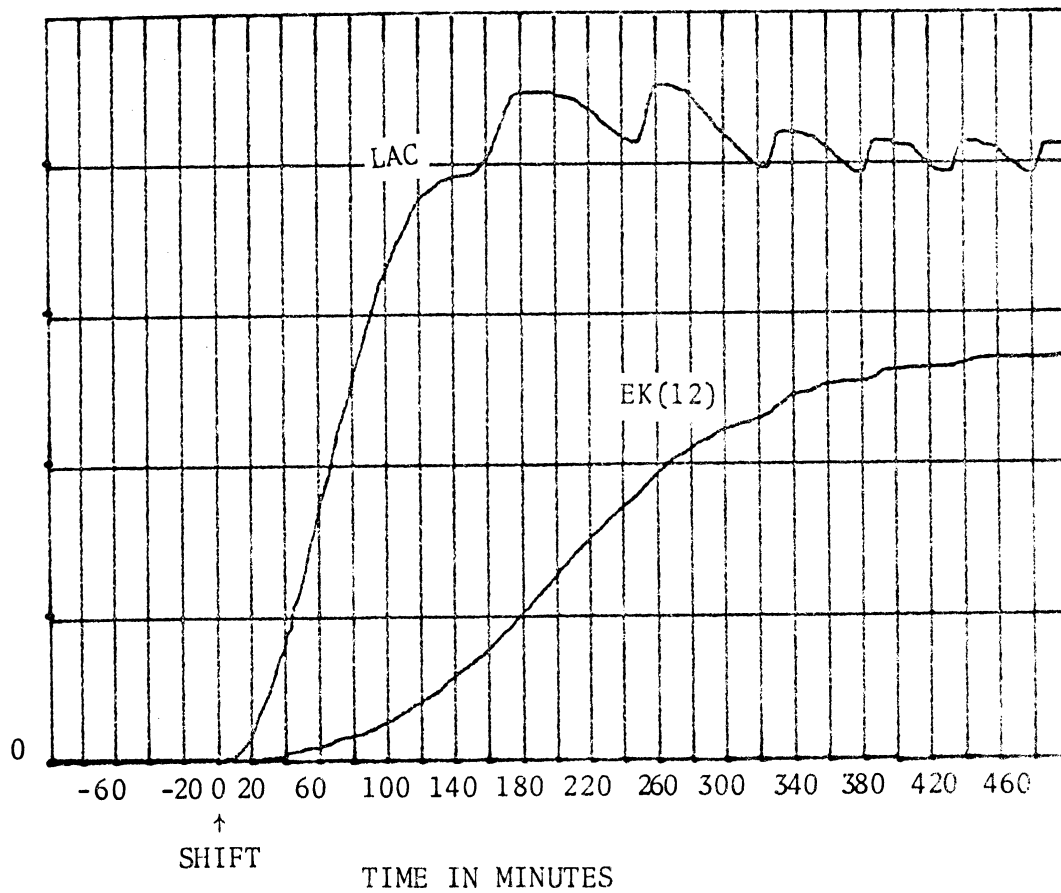
Figure 3.14: LAC and EK(12) in Simulated Shift to Lactose. At time 0 ONECELL is shifted from glucose minimal medium (environment one) to lactose minimal medium. EK(12) represents LAC operon enzymes induced by this shift. Scale is linear, units arbitrary. Sample points are from a ONECELL run shifted four minutes after cell division, and are four minutes apart.

(linear scales) of LAC and LAC enzymes (called EK(12)) versus time, for a single ONECELL shifted ten minutes after a cell division.

The shape of the EK(12) curve is as expected; that is, its slope is initially low, reflecting the slow rate at which the depressed cellular synthesis machinery is able to form the required mRNA and enzymes for lactose transport and utilization. The rate accelerates as the enzyme already synthesized enables the cell to use lactose and resume biosynthetic operations at a higher rate. Finally the curve levels off as the level of full induction is approached.

The LAC pool does not reach the neighborhood of its steady-state growth value for about two hours. In the meantime, GLUC has been at a very low level (see Figure 3.15), with a strong effect on the other pools. (GLUC is shown slightly above 0 in the graph, so that it is distinguishable from the axis, but its minimum value is actually 0.) The effect of the low GLUC on the AA and NUC pools is shown in Figure 3.16. The lowered concentrations of NUC and AA are reflected in the MRNA curve (Figure 3.15), and in the other pools (as shown in the logarithmic plots of Figures 3.17 - 3.21). These graphs use the weightings described for the shift-down graphs.

Figure 3.17 shows the different way in which the shift affects DNA and RNA. DNA synthesis does not fall off quite as fast, although it drops to zero eventually. RNA synthesis is never quite zero, but falls off quickly to a low value. (Of course, the differential rate of synthesis of mRNA specific for the LAC enzymes is very high.) The curves for DNA and RNA regain the proper slopes for growth at the 50-minute generation time expected in steady-state growth on lactose. The small distance between the DNA and RNA lines after about time
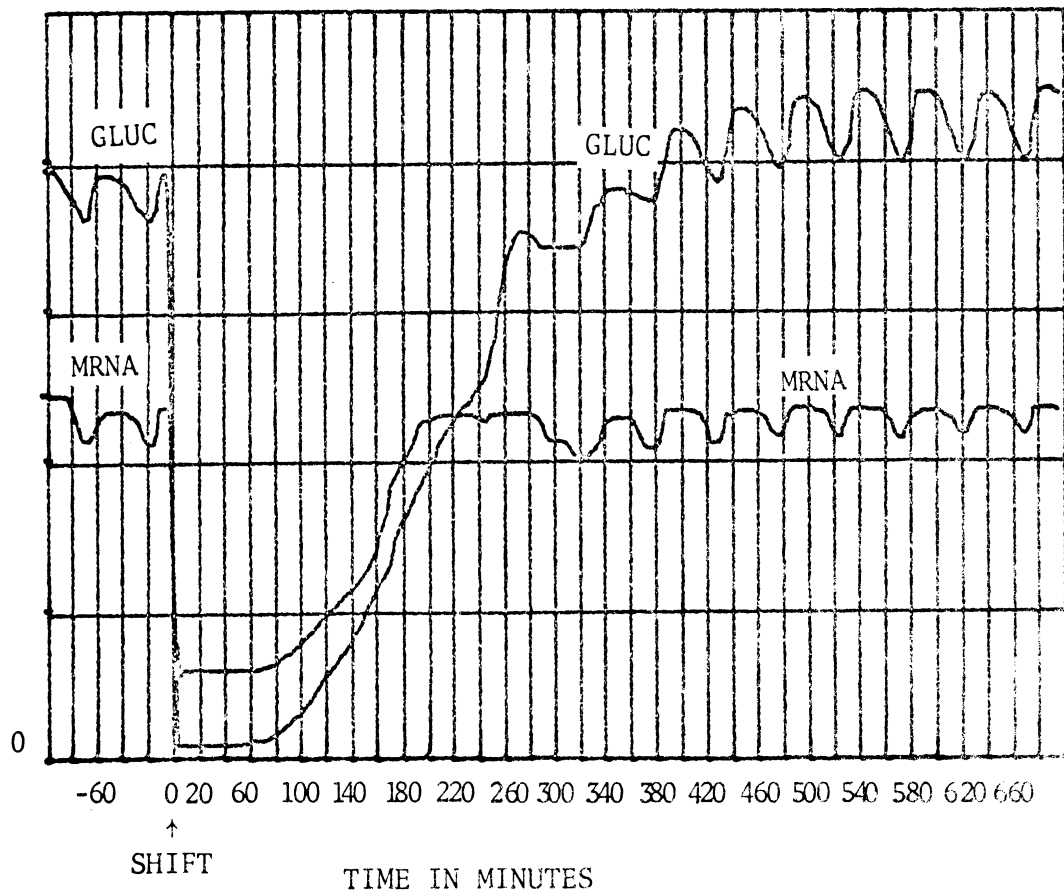
Figure 3.15: GLUC and MRNA in Simulated Shift to Lactose. At time
0 a shift from glucose minimal medium to lactose
minimal medium is simulated. Scale is linear, units
arbitrary. GLUC goes to 0, but is shown just above
the x-axis for clarity. Sample points are from a
ONECELL run shifted four minutes after cell division,
and are four minutes apart.
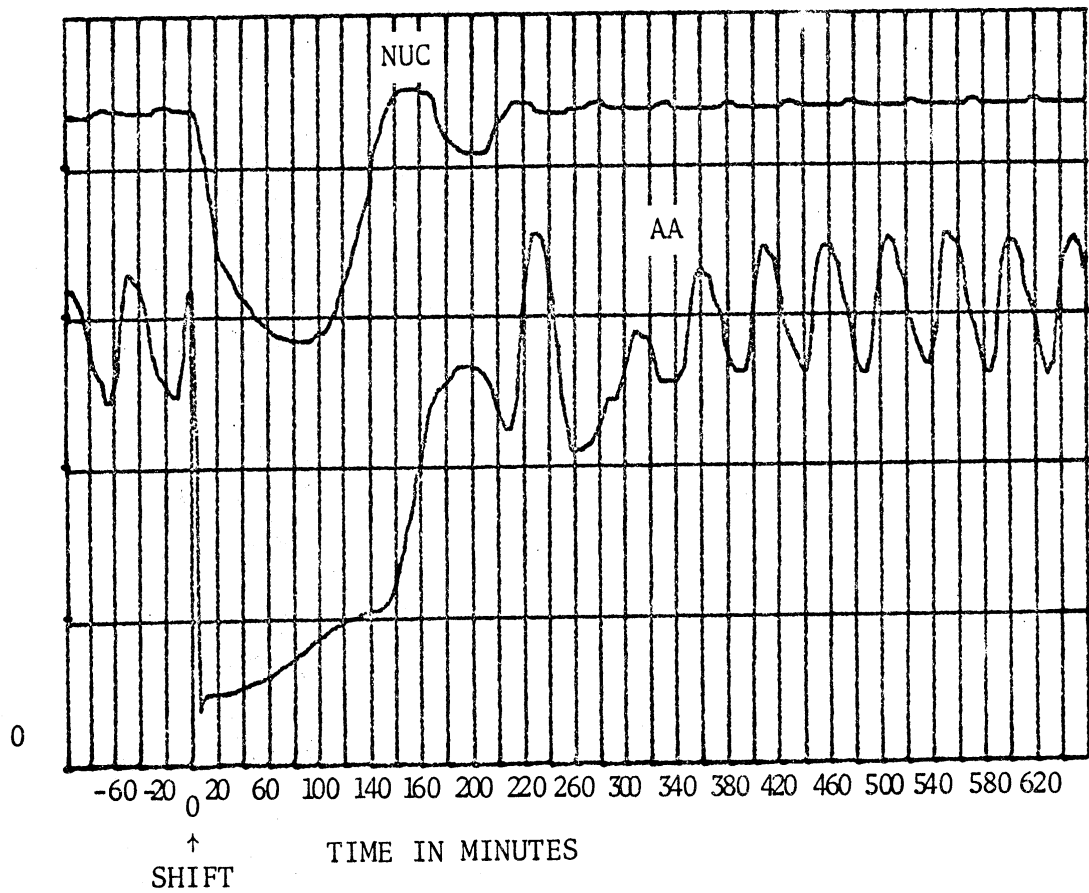
Figure 3.16: NUC, AA in Simulated Shift to Lactose. At time 0, a shift from glucose minimal medium to lactose minimal medium is simulated. Scale is linear, units arbitrary. Sample points are from a ONECELL run shifted four minutes after cell division, and are four minutes apart.
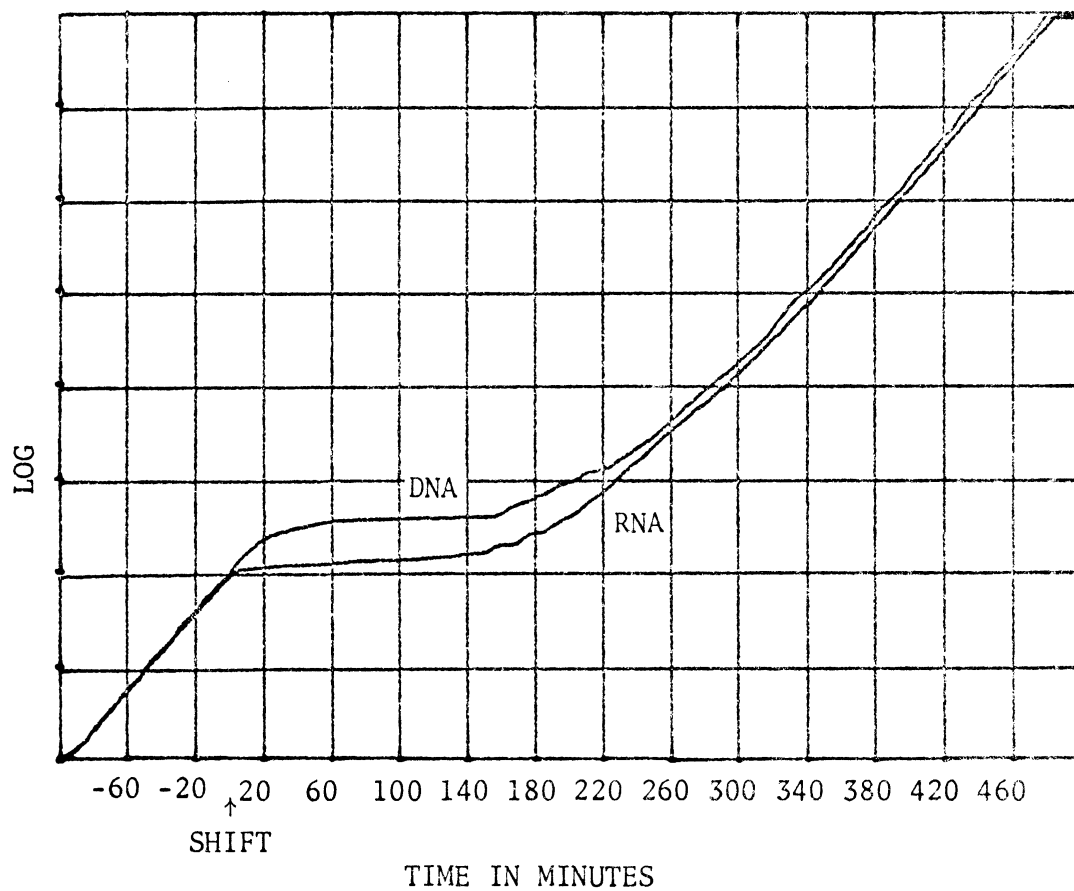
Figure 3.17: DNA, RNA in Simulated Shift to Lactose. Logarithmic
plots of amounts - units adjusted so curves for balanced
(pre-shift) growth in environment one (glucose minimal
medium) coincide. Distance between horizontal lines
represents one doubling. Each value is a weighted sum
of three values, one calculated in each of three ONECELL
runs. The runs differ in the age of the cell at the
time of the shift.

260 is attributable to the procedure used for scaling the two variables

so that they have a common slope in the pre-shift region. That is,

the scaling is done using the basis of the pre-shift performance,

and a small difference of the values used for scaling from their

long-term averages causes the larger divergence observed in the

right-hand portion of the graph. Examination of printed output shows

that DNA and RNA do in fact attain the proper ratios expected of

them.

Figure 3.18 shows the curves for cellular mass and cell counts.

Both curves become flat after the shift, until the internal GLUC

pool has begun to climb. The lagging of the mass curve below the

count curve for a short time shows that smaller-than-normal cells

are formed during that period.

The curves for ATP and WALL, both of which depend on glucose,

also flatten out after the shift (Figure 3.19). It is felt that the

fact that ATP does not decline is an indication that the ATP model is

probably not as sensitive as it should be. The reason it does not

is that the reactions which use ATP have come to a virtual standstill

due to substrate limitation, so the demand for ATP is small. The

improved ATP-ADP-NUC model which is being prepared will probably alle-

viate this difficulty.

The RIB and PRTN curves exhibit a flattening after the shift

(Figure 3.20). RIB then leads PRTN for a short period, after which

they attain their asymptotic rates and ratios.

The curve for NUC is repeated, this time as a logarithmic plot,

in Figure 3.21. Also shown there is the TRNA curve, which decreases

in slope after the shift, then attains the rate required for maintenance
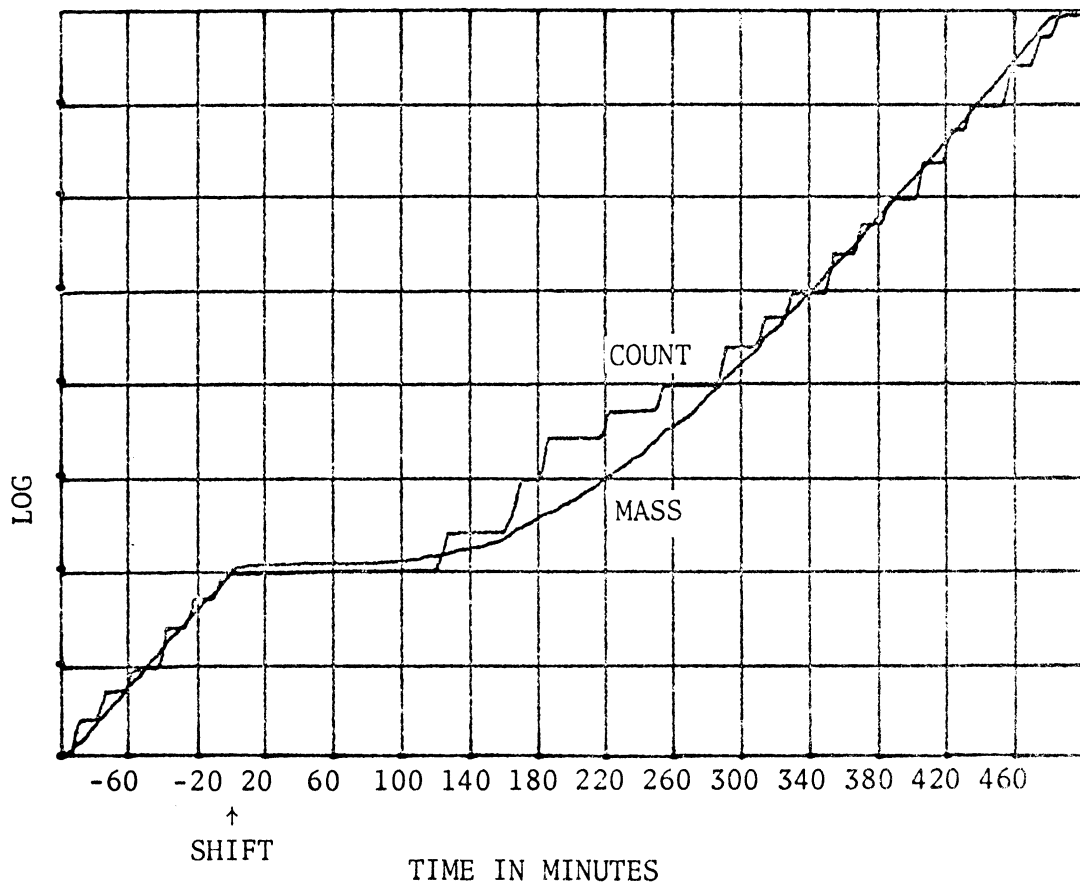
Figure 3.18: Mass and Count in Simulated Shift to Lactose. Logarithmic plots of total cellular mass and numbers of cells (viable counts). Units are adjusted so curves for balanced (pre-shift) growth in environment one (glucose minimal medium) coincide. Distance between horizontal lines represents one doubling. Each value is a weighted sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.

96
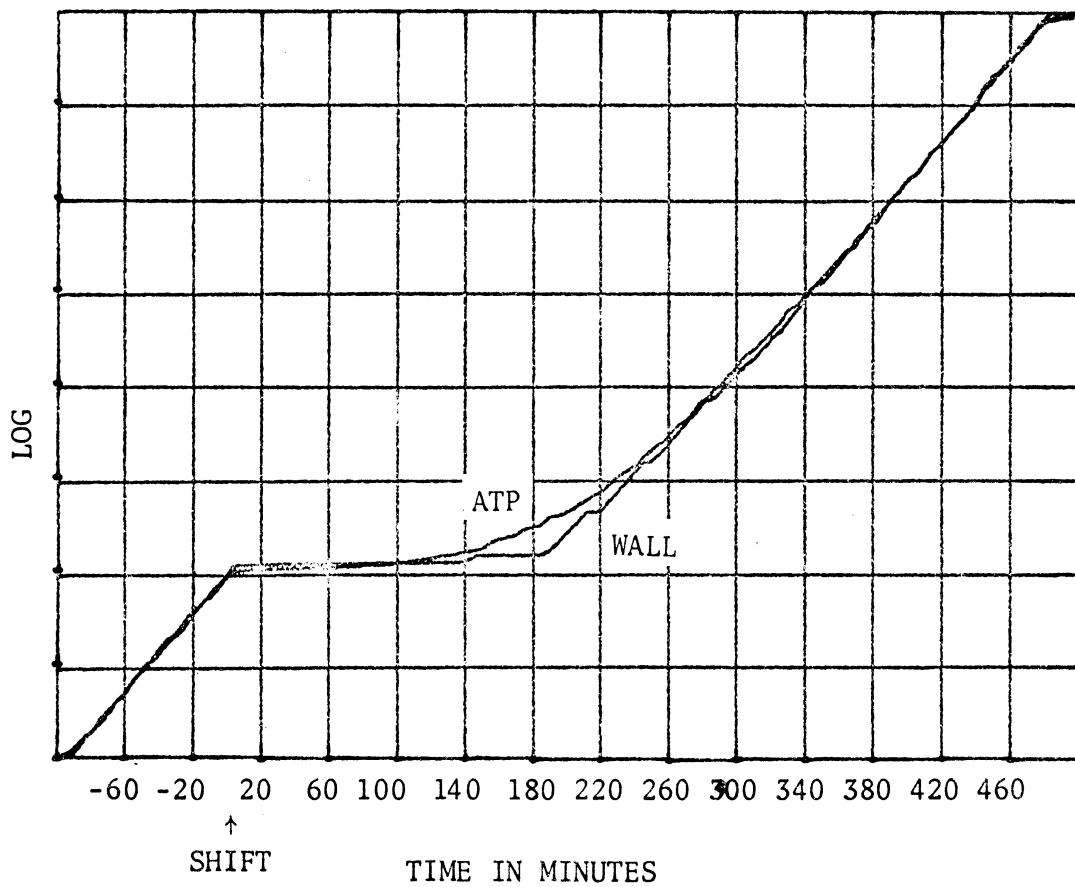
Figure 3.19: ATP and WALL in Simulated Shift to Lactose. Logarithmic plots of amounts - units are adjusted so curves for balanced (pre-shift) growth in environment one (glucose minimal medium) coincide. Distance between horizontal lines represents one doubling. Each value is a weighted sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.
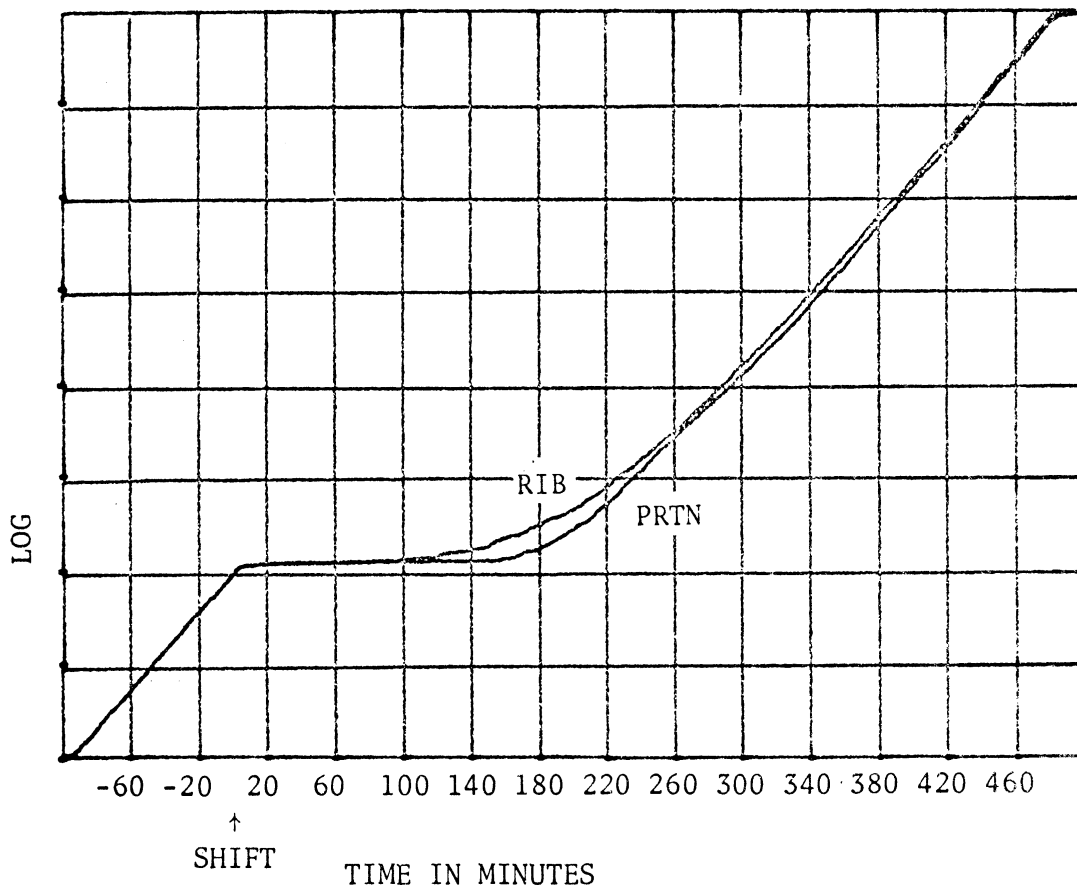
Figure 3.20: RIB and PRTN in Simulated Shift to Lactose. Logarithmic plots of amounts - units are adjusted so curves for balanced (pre-shift) growth in environment one (glucose minimal medium) coincide. Distance between horizontal lines represents one doubling. Each value is a weighted sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.
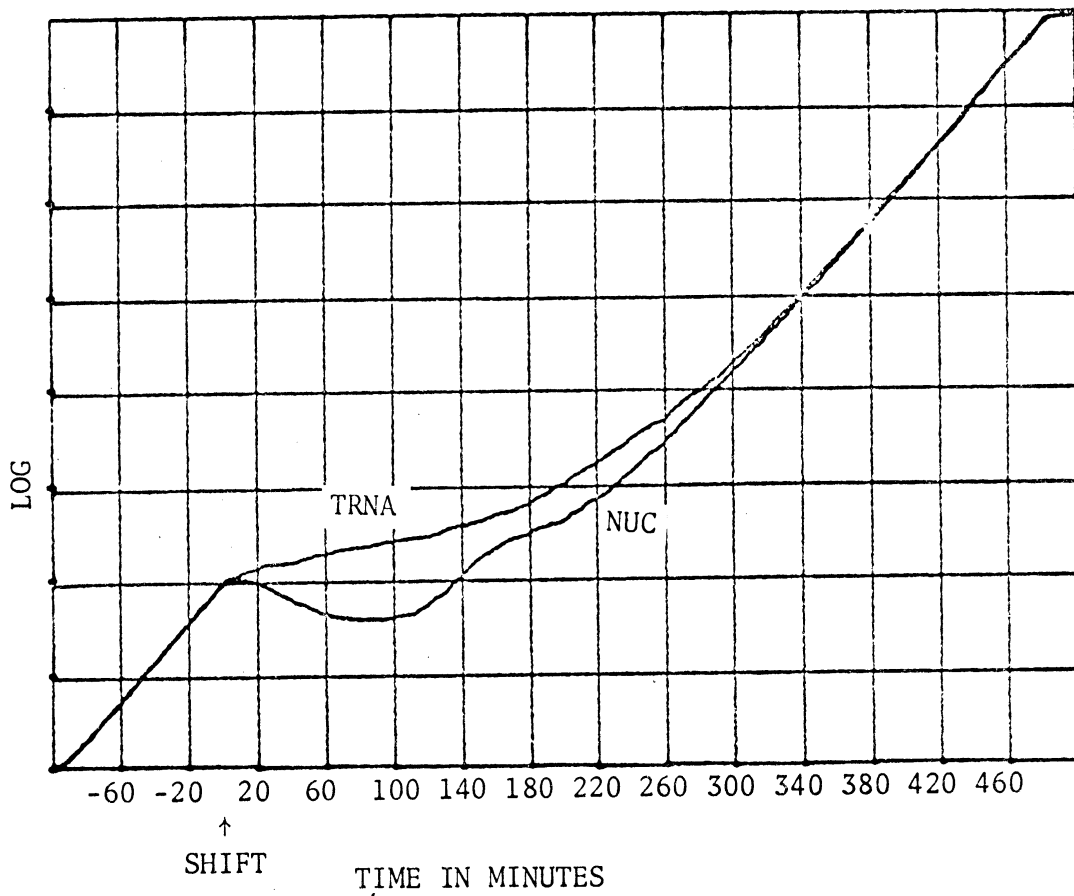
Figure 3.21: TRNA and NUC in Simulated Shift to Lactose. Logarithmic plots of amounts - units are adjusted so curves for balanced (pre-shift) growth in environment one (glucose minimal medium) coincide. Distance between horizontal lines represents one doubling. Each value is a weighted sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.

of the TRNA concentration in steady-state growth.

A general assessment of the shift to lactose indicates that the induction apparatus worked properly to the extent it could be expected to; oscillation of LAC enzyme production was not obtained because no catabolite repression was modeled. Furthermore, the simulated cell was able to withstand the stresses of carbon-source starvation and regain stable growth at the proper rate, in spite of large disruptions in the levels of its constituent pools.

CHAPTER 4

The preceding chapters have discussed the development and present performance of the ONECELL simulation. In this chapter the future directions which might be taken in continuing this research will be discussed.

There are several avenues of endeavor which may be pursued from the point of development of the simulation reported here. There are improvements to the ONECELL model which have already been programmed, and await only a refitting of parameters for inclusion in the ONECELL simulation. In addition, the performance of the ONECELL model in simulated shifts, as reported in Chapter 3, suggests further improvements. That is, it is possible to refine models and refit parameters using the information gained from observation of the shifts. This type of improvement should yield a model of greater power than the present one, which can be tested in still other simulated experimental situations. For example, the effects of treatment with certain drugs, such as chloramphenicol or nalidixic acid, could be simulated, and the performance of the simulation again compared with experimental data from the microbiological literature.

The interest of the author in development of a model of a bacterial colony has already been discussed. Preliminary work, as reported in Goodman, Weinberg, and Laing (1970) has shown that it is within current technical capabilities to implement a colony simulation based on the ONECELL model. The prototype of a colonial simulation system has already been implemented by the author at the Logic of Computers Group. The spatial representation and symmetry assumptions made in

the colony model facilitate simulation in the Cellular Space Simulation
System developed there. (The hardware and software used in both the
single cell and colony simulations are described in Goodman, Weinberg,
and Laing (1970).).

The process of developing a colony model based on ONECELL involves
two principal areas of modeling effort. One is the representation
of the spatial or geometric aspects of the colony; the other is the
relation of the rates at which various materials enter and leave cells
to the concentrations of these materials in the extracellular environment
(intercellular "spaces" and the agar medium on which the colony develops).
Of course, these two areas are not independent, and decisions in one
area impose constraints on the other.

The problem of geometric representation can best be viewed by
first considering the nature of a cellular space. A (two-dimensional)
cellular space may be roughly characterized as an infinite checkerboard
on which all squares are identical in the potential which they have
for assuming different states. Each square, or *cell* (italicized to
avoid confusion with a cell in the biological sense) can be viewed
as an automaton, or machine, or computer, which calculates its next
state using information about its present state and the present states
of its neighbors (which are generally considered to be other *cells*
adjacent to it, or to which it is directly connected). Inputs from
outside the space may also be used in determining the state transition.
A (two-dimensional) cellular space is easily seen to be useful in
representing the behavior of a surface which can be discretized into
a regular (statically arranged) array of components which are identical
in state set (not necessarily in state). It should be clear that a

bacterial colony is not well represented as such a surface, since

it exhibits significant variation in three dimensions, and is not

statically organized.

Two fundamental points must be made to demonstrate the utility

of a cellular space for representing a colony. The first is that

the colony exhibits a strong degree of radial symmetry (about an axis

through the center of the colony and perpendicular to the agar surface).

Thus, all behavior which does not violate this symmetry assumption

can be captured on a plane surface (a section through the axis of the

colony). The second point is that the expanding (dynamic) configur-

ation of cells in a real colony does not prohibit their representation

by an array of automata with a static neighborhood relationship. The

fundamental insight needed for this assertion is that a *cell* (of the

cellular space) need not represent one cell (i.e., one bacterial cell).

The difficulties which would be encountered in such a one-for-one

representation are immediately obvious if the situation is considered

in which two or more adjacent cells divide simultaneously. A fund-

amental property of cellular spaces concerns the rate at which infor-

mation (i.e., the effect of changes of state) can propagate through

the space, and this property implies the impossibility of handling

the above situation if a one *cell* for one cell model is attempted.

However, it is not necessary to represent each cell as a *cell*; the

alternative is to develop a procedure which determines, at each cell

division, what *cell* is to represent each daughter cell produced.

Thus each *cell* represents many cells (see Figure 4.1). In fact, the

criterion for determining how many *cells* are necessary to represent

a colony of a particular size should be the level of detail at which

the internal differences in the colony are to be represented; i.e., how many distinct micro-environments must be considered.
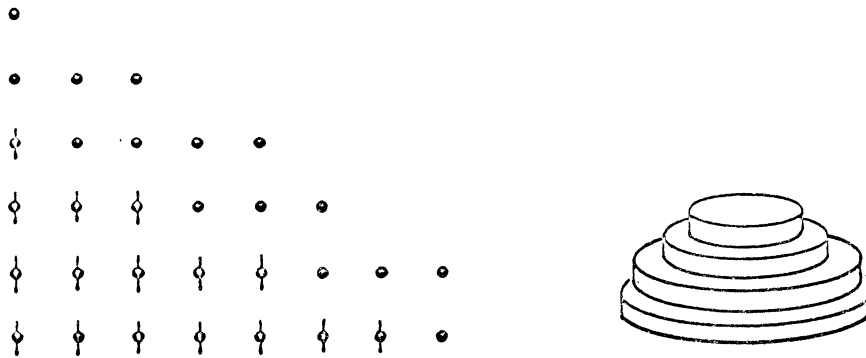
Figure 4.1: Representation of Colony Geometry. The left-hand side shows the general form used in the cellular space to represent a colony of the approximate shape shown on the right. The lines on the left side signify *cells* that are growing; the points, *cells* that do not yet represent any bacterial cells.

One difficulty which the above representation raises is the problem of translating a given cellular space configuration to the geometric situation it represents. A reasonable solution appears to be to maintain dynamically for each *cell* both its height from the agar surface and its distance from the center of the colony. The radial distance can be calculated using the radial distance of the *cell* to the left and the number and volume of the *cell* under consideration, provided packing density assumptions are made.

It may be evident from the discussion above that although a cellular space may be useful in modeling a colony, it is not an ideal formalism, due to the static relationships which it imposes on neighbors. However.

this very shortcoming is a boon in allowing ease of geometric inter-
pretation of configurations, as also discussed above. Other
formalisms for representing multi-cellular growth are the subjects
of research, for example that of Laing (1970), Lindenmayer (1968),
and others. It is also their experience that increasing the flexi-
bility of the relationships between cells increases the difficulty
of mapping from the model space back to a three-dimensional Euclidean
space.

There are areas in which it is desirable to modify the formalism
of a cellular space to accommodate a colony model in a simple fashion.
In particular, it is desirable to bound the space below with a layer
of distinguished *cells* (or an extra-spatial model) representing the
agar medium on which the colony is growing. Further, the radial
symmetry is most effectively modeled by using a left-hand boundary
as well. The finiteness of the automata (*cells*) is also
interesting - in fact they are finite (since the state of each cell
is stored as a finite number of bits of memory), but they are more
easily conceived of as consisting of a finite set of variables which
range over the (nonfinite) set of real numbers. It would be possible,
on the other hand, to discretize the ONECELL model variables to yield
a model with a much smaller state set than the present (practically
infinite) one. Its transition function, however, might be more complex,
and such a reduction of state-space size is of interest mainly from a
theoretical viewpoint.

The second area of modeling effort required to use ONECELL in
a colony model is to relate the flow of materials into and out from
a cell to the concentration of those materials outside the cell.

The interactions between ONECELL and the environment which are included in the state set of the prototype are the four types of nutrients (NUC,AA,GLUC, and LAC) and oxygen and wastes. The problems in connecting the flow rates with the concentrations are (1) specification of the volume of extracellular material which the nutrients, etc., are considered to be present in, and (2) specification of the rates of diffusion of the various substances through the extracellular space. These are both difficult problems, but can be made more tractable by means of some simple assumptions concerning the packing density of cells in the colony and the composition of the intercellular space.

The colony model is still in need of considerable development. It, like the ONECELL model, suffers from a scarcity of information concerning some of its variables. While some investigators have studied the development of colonies on solid (agar) medium, the measurements taken are often of macroscopic variables, such as radial growth rate, rather than of variables which display different levels in various parts of the colony. There is work which tries to connect the two levels of variables, such as that by Pirt (1967) and Whittaker and Drucker (1970), but this sort of work is still in an early stage. The approach taken by Pirt in describing the distribution of nutrients available to a colony growing on agar is a good indication that the colony model would be of utility in the study of bacterial growth, and that the questions that are formulatable within the model are biologically interesting ones.

One of the chief limitations in the development of the model to date has been severe shortage of both core memory and output

capabilities. With the forthcoming alleviation of these restrictions, in the form of a doubling of core memory and the addition of a line printer, future work on the simulation should be able to proceed at an accelerated pace.

The ONECELL model was originally designed by Weinberg to operate with certain fixed internal concentrations, receiving no inputs from its environment, and producing no changes in its environment. The new ONECELL model has as inputs concentrations of four different types of nutrients, and has available as outputs the amount of each nutrient consumed, allowing dynamic alteration of the environment of the cell. In addition to these inputs and outputs, it would be desirable to include in a future model additional environmental factors which interact with the cell's metabolism. Foremost is the supply of oxygen, determining whether the growth is aerobic or anaerobic. ONECELL now assumes that oxygen is plentiful. The structure of the model is such that it should not be difficult to introduce the effect of oxygen deprivation. Another new factor, which should probably also be introduced when oxygen is modeled, is wastes that are produced in (and may diffuse out from) the cell. Lactate, for example, will be generated in large quantities under anaerobic growth conditions, and should probably be considered in the model.

It is hoped that the ONECELL model, in one form or another, will continue to be developed. The author feels that in spite of the difficulties which will doubtless be encountered, the goal is a worthwhile one, and the model embodies a level of detail and an array of biological controls which should enable it to serve as a useful tool in the study of bacterial growth.

APPENDIX A
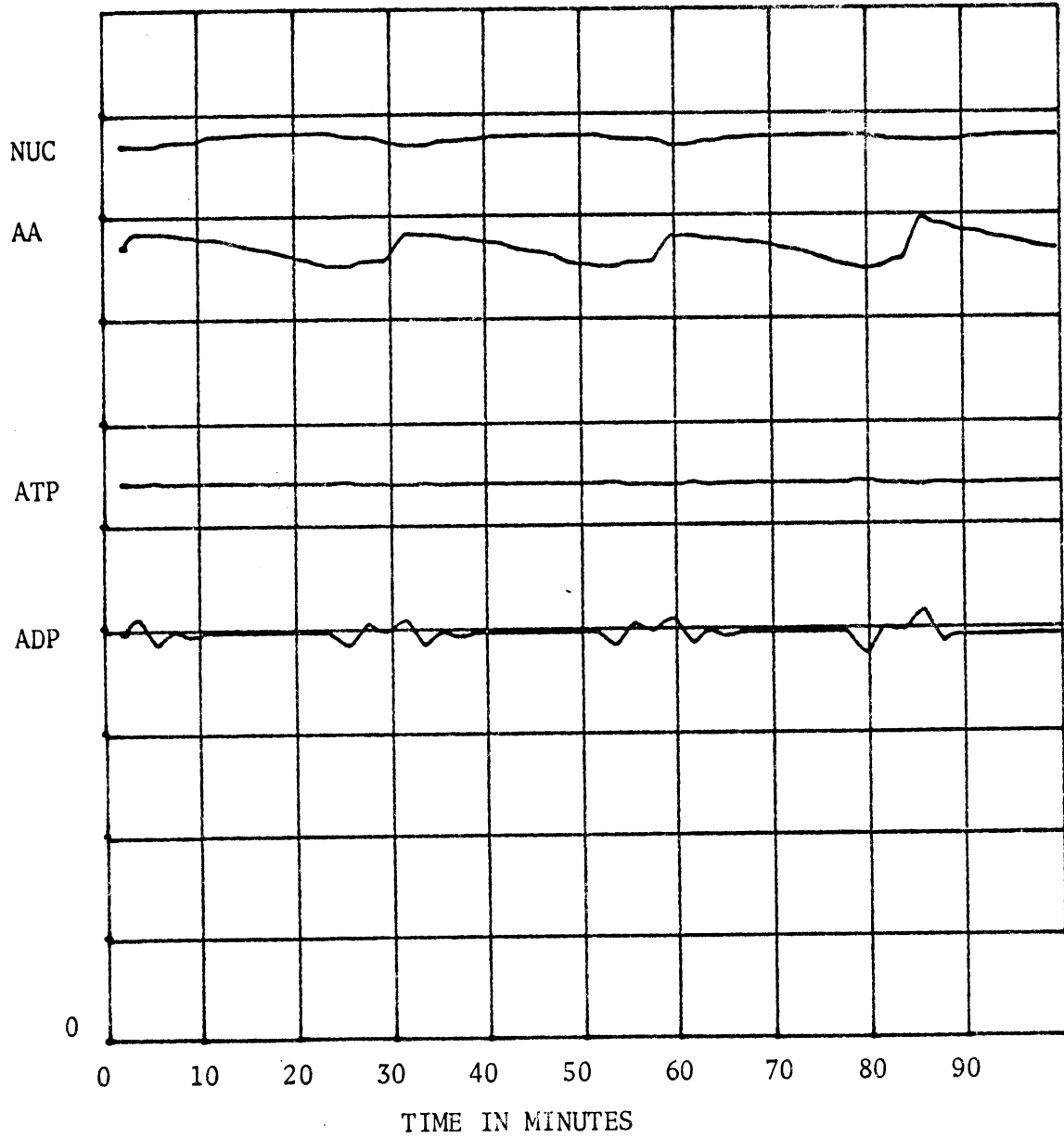
Figure A.1: NUC, AA, ATP, ADP in Simulated Growth in Environment
Two. Scale is linear, units for each pool are arbitrary.
Sample points are two minutes apart. Recording was
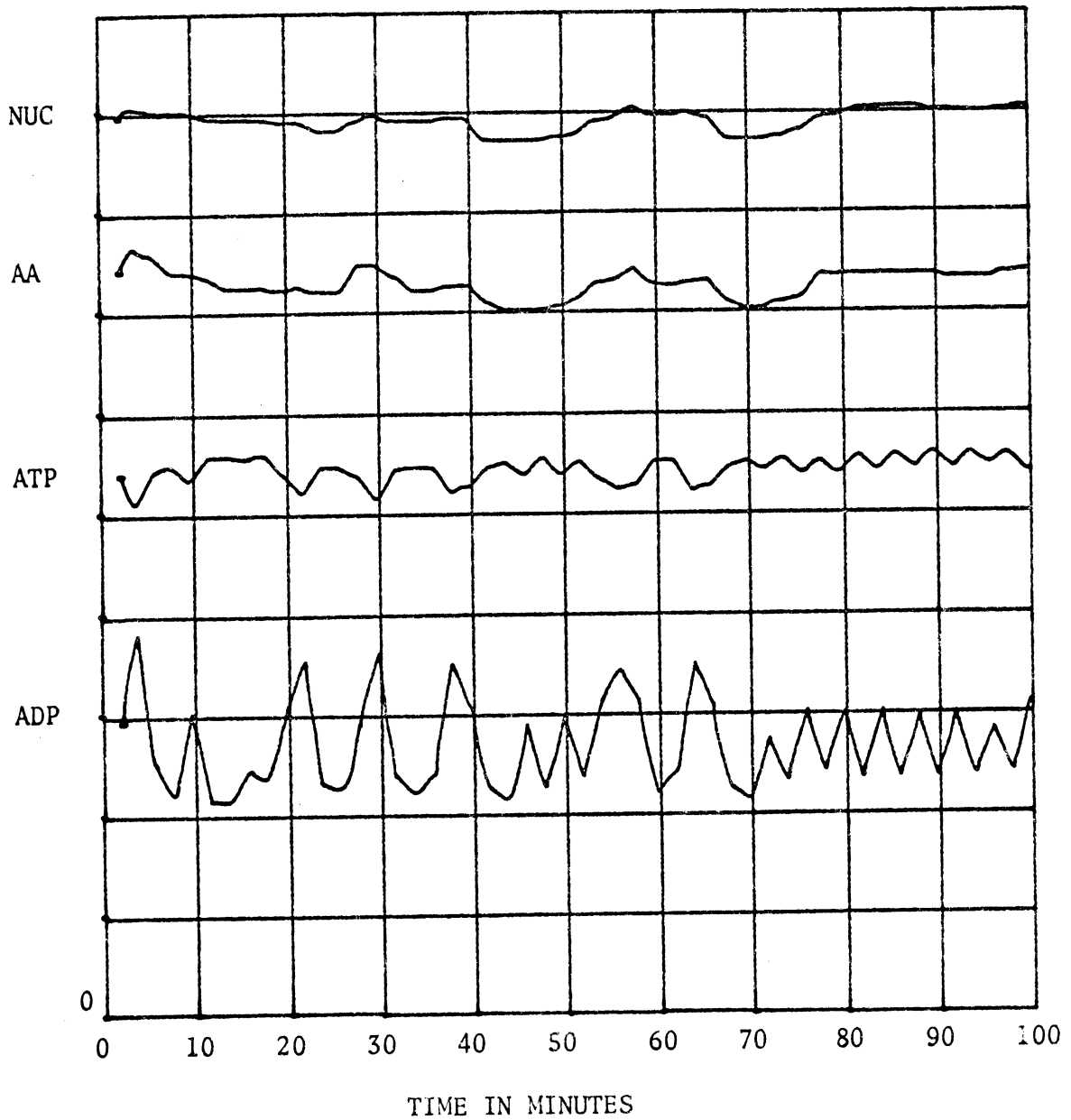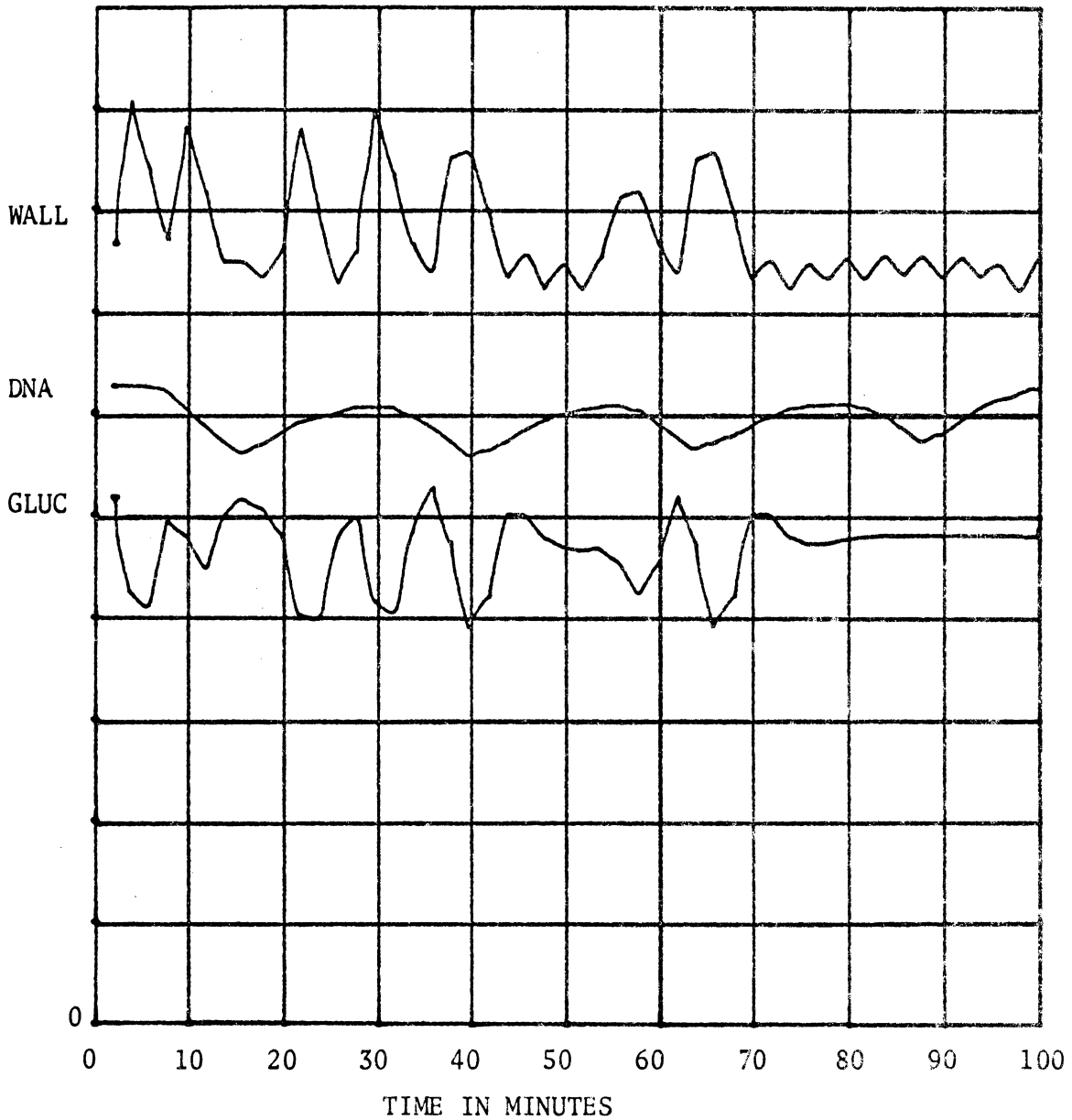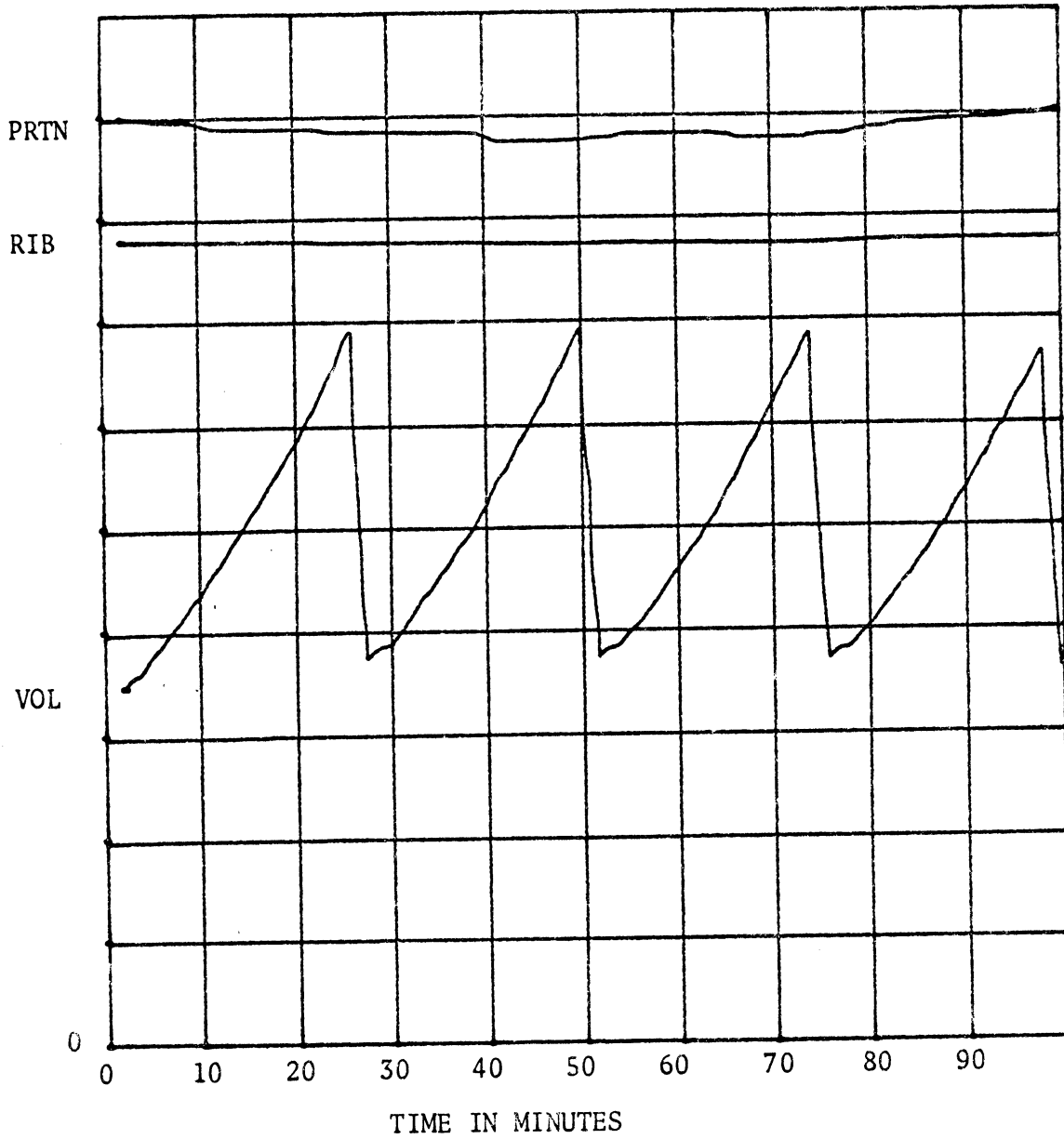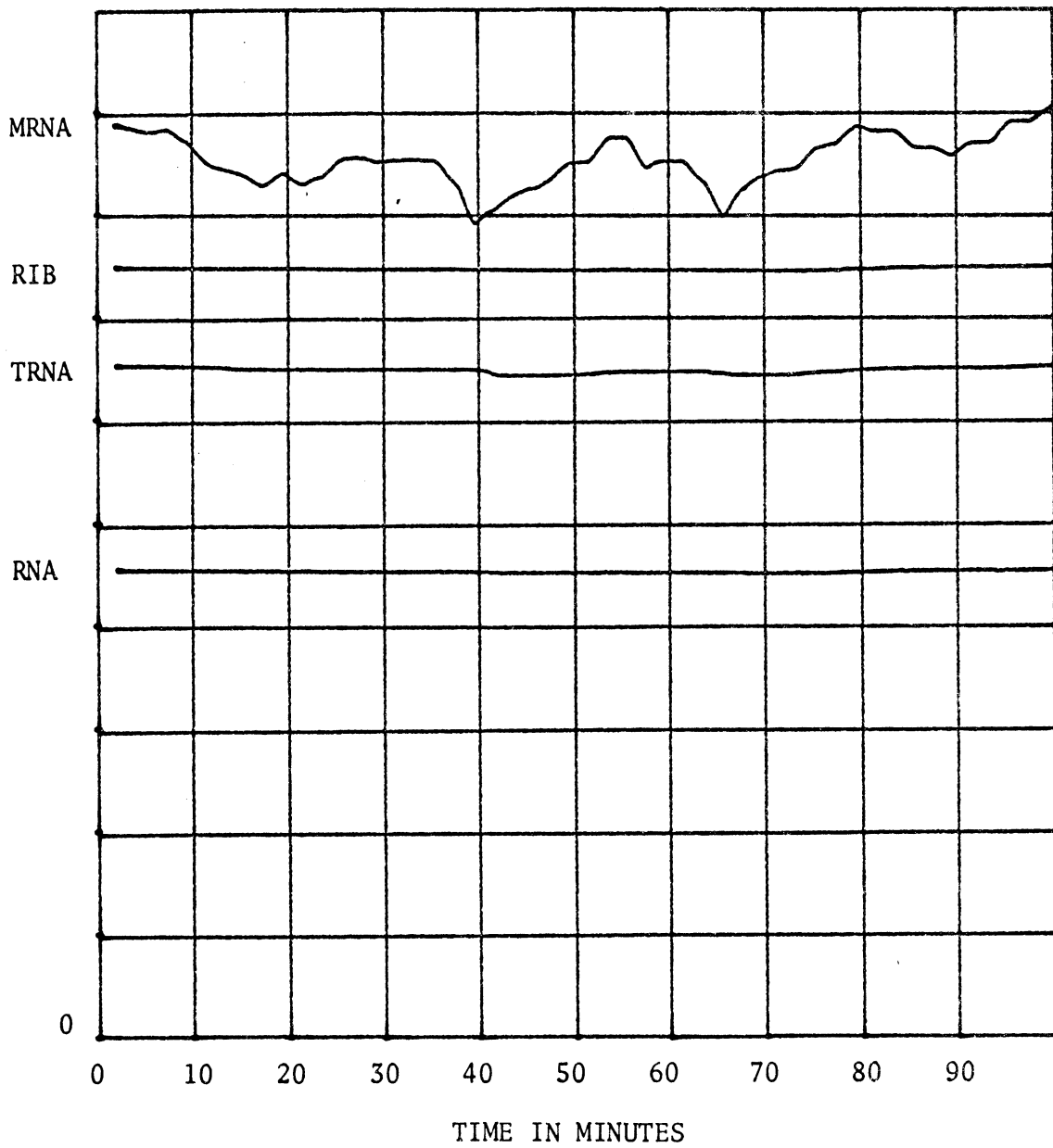begun after one simulated cell division.
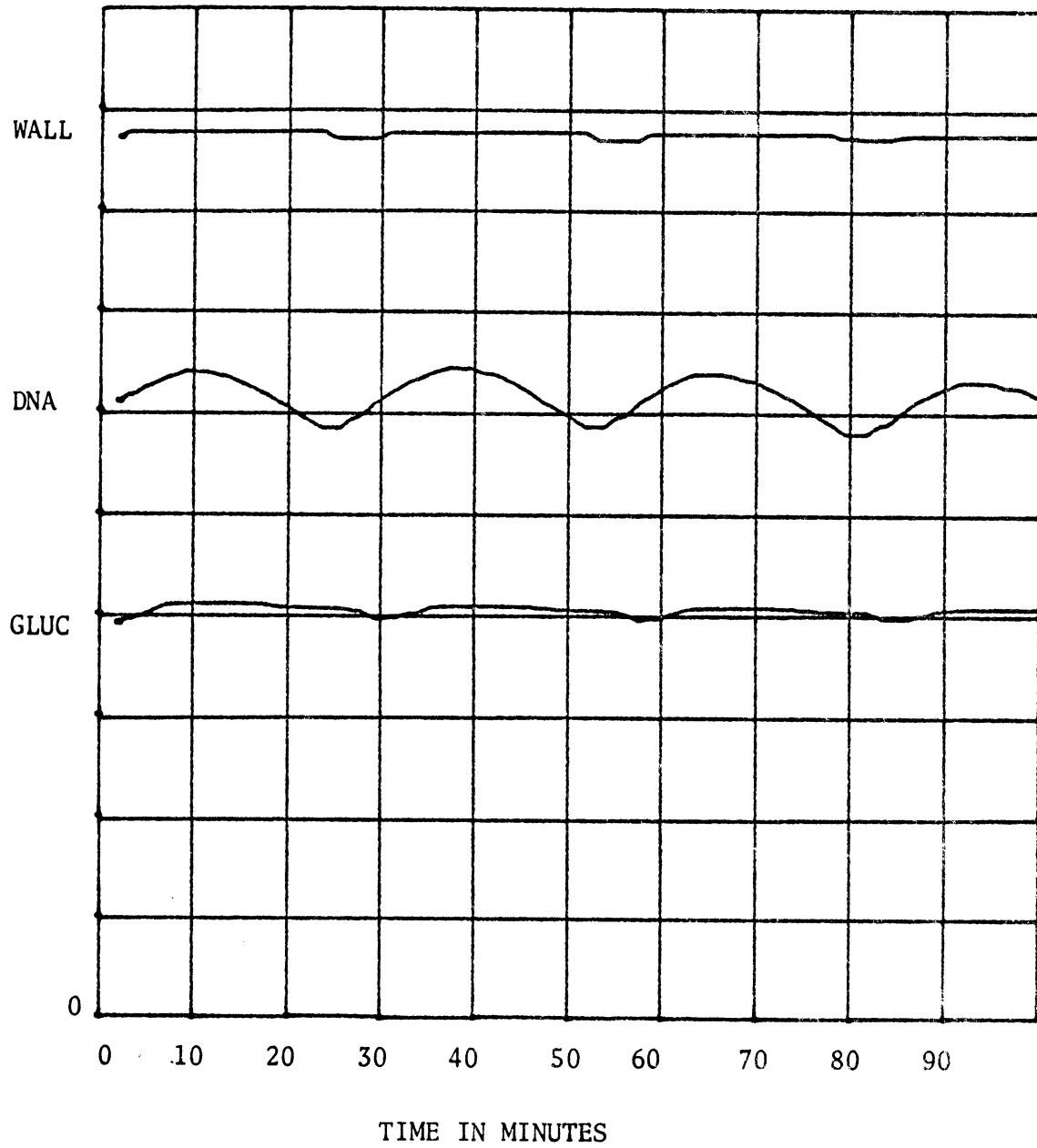
Figure A.2:   WALL, DNA,   GLUC in Simulated Growth in Environment Two.
Scale is linear, units for each pool are arbitrary.
Sample points are two minutes apart.   Recording was
begun after one simulated cell division.
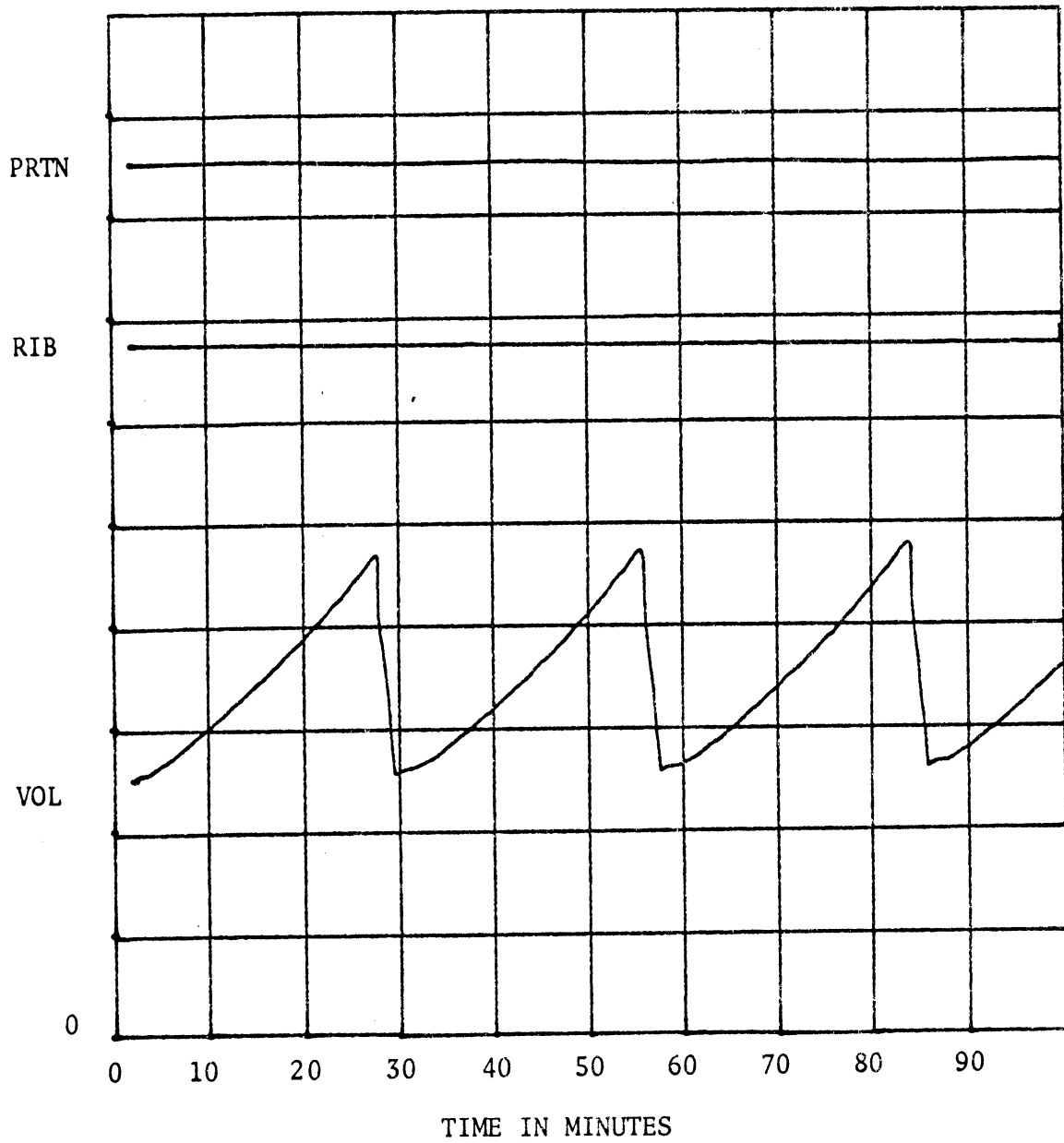
Figure A.3: PRTN, RIB, VOL in Simulated Growth in Environment Two.
Scale is linear, units for each pool are arbitrary.
Sample points are two minutes apart. Recording was
begun after one simulated cell division.
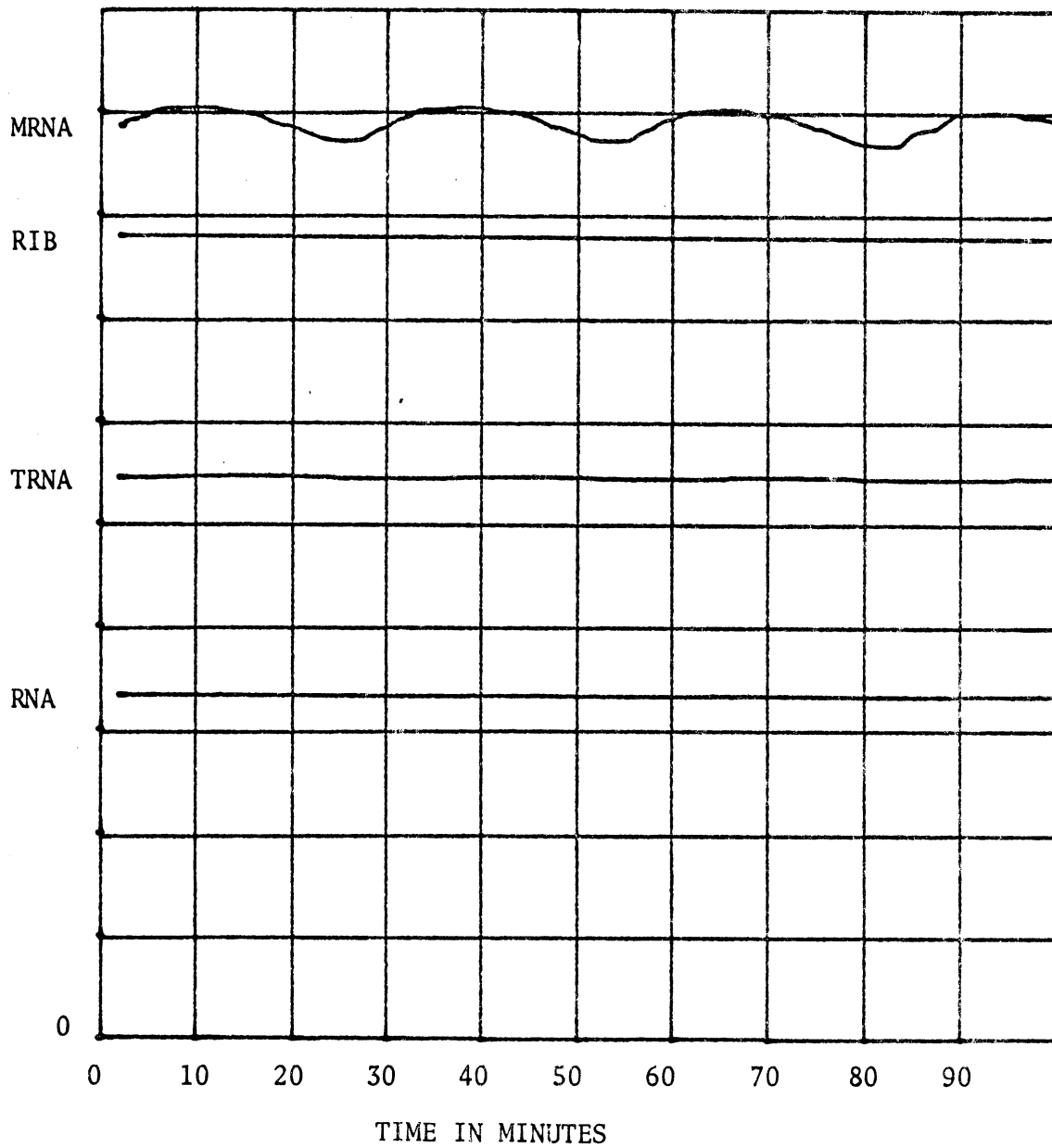
Figure A.4:   MRNA, RIB, TRNA, RNA in Simulated Growth in Environment Two.
RNA is a weighted sum of MRNA, RIB, and TRNA. Scale is
linear, units for each pool are arbitrary. Sample
points are two minutes apart. Recording was begun
after one simulated cell division.

Figure A.5: NUC, AA, ATP, ADP in Simulated Growth in Environment
Three. Scale is linear, units for each pool are
arbitrary. Sample points are two minutes apart.
Recording was begun after one simulated cell division.

Figure A.6:   WALL, DNA, GLUC in Simulated Growth in Environment Three.
             Scale is linear, units for each pool are arbitrary.
             Sample points are two minutes apart. Recording was
             begun after one simulated cell division.

Figure A.7: PRTN, RIB, VOL in Simulated Growth in Environment Three.
Scale is linear, units for each pool are arbitrary.
Sample points are two minutes apart. Recording was
begun after one simulated cell division.

Figure A.8: MRNA, RIB, TRNA, RNA in Simulated Growth in Environment
Three. RNA is a weighted sum of MRNA, RIB, and TRNA.
Scale is linear, units for each pool are arbitrary.
Sample points are two minutes apart. Recording was
begun after one simulated cell division.
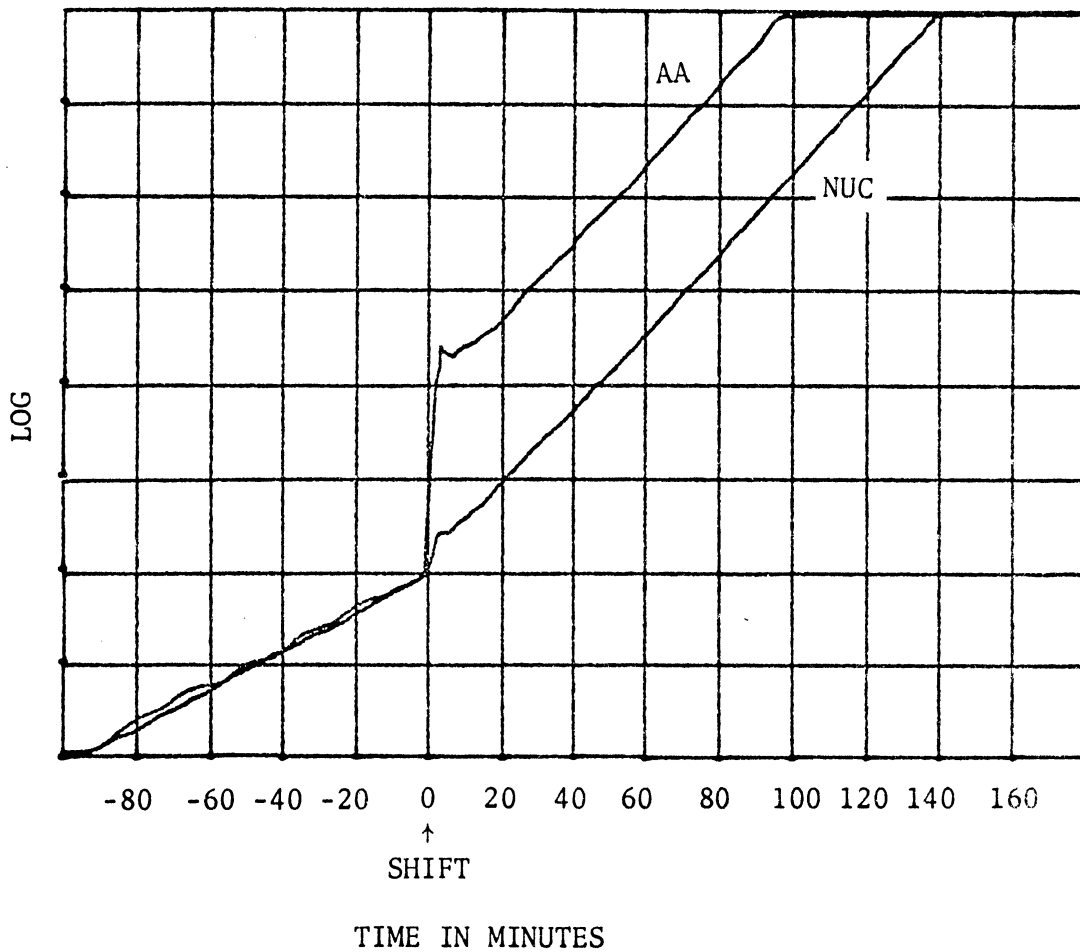
Figure A.9: AA and NUC in Simulated Shift-up from One to Three
Logarithmic plot of amounts - units adjusted so curves
for balanced growth (pre-shift) in environment one coincide.
Distance between horizontal lines corresponds to one
doubling. Each point is a sum of three values, one cal-
culated in each of three ONECELL runs. The runs differ
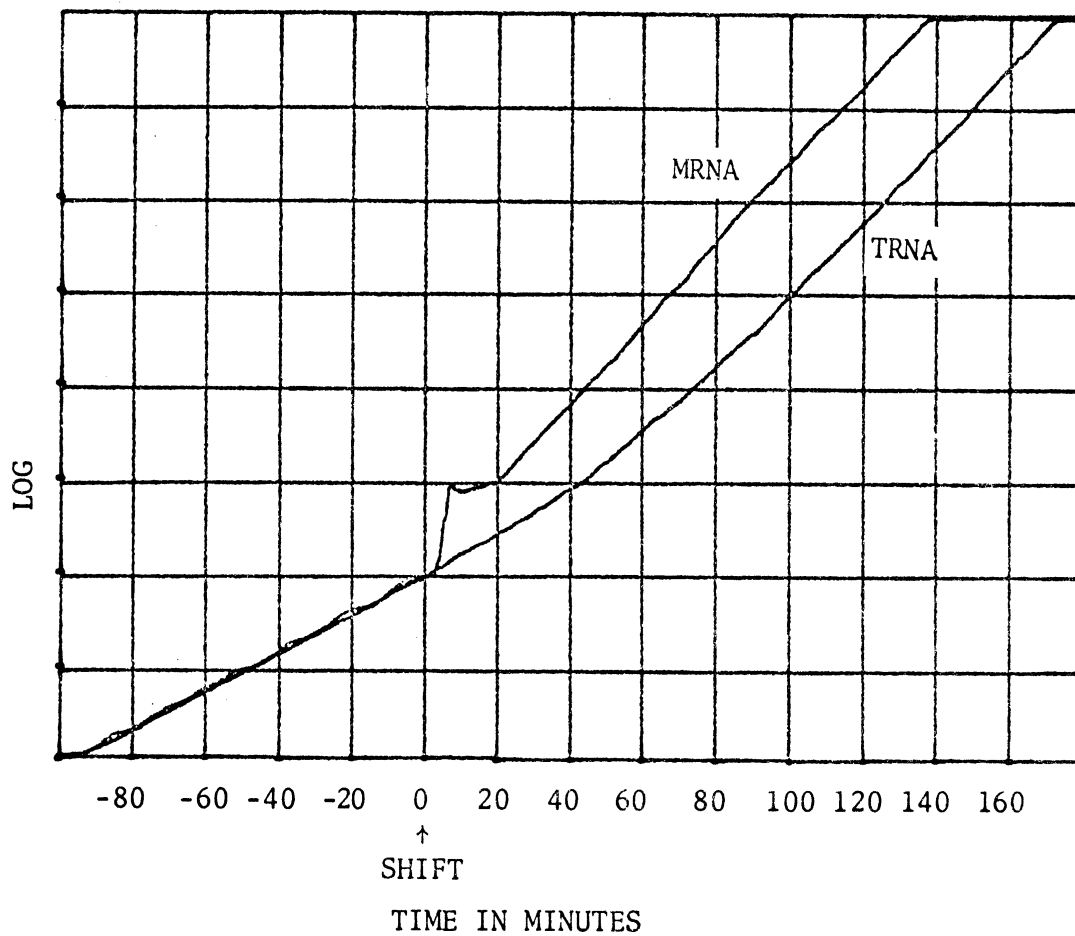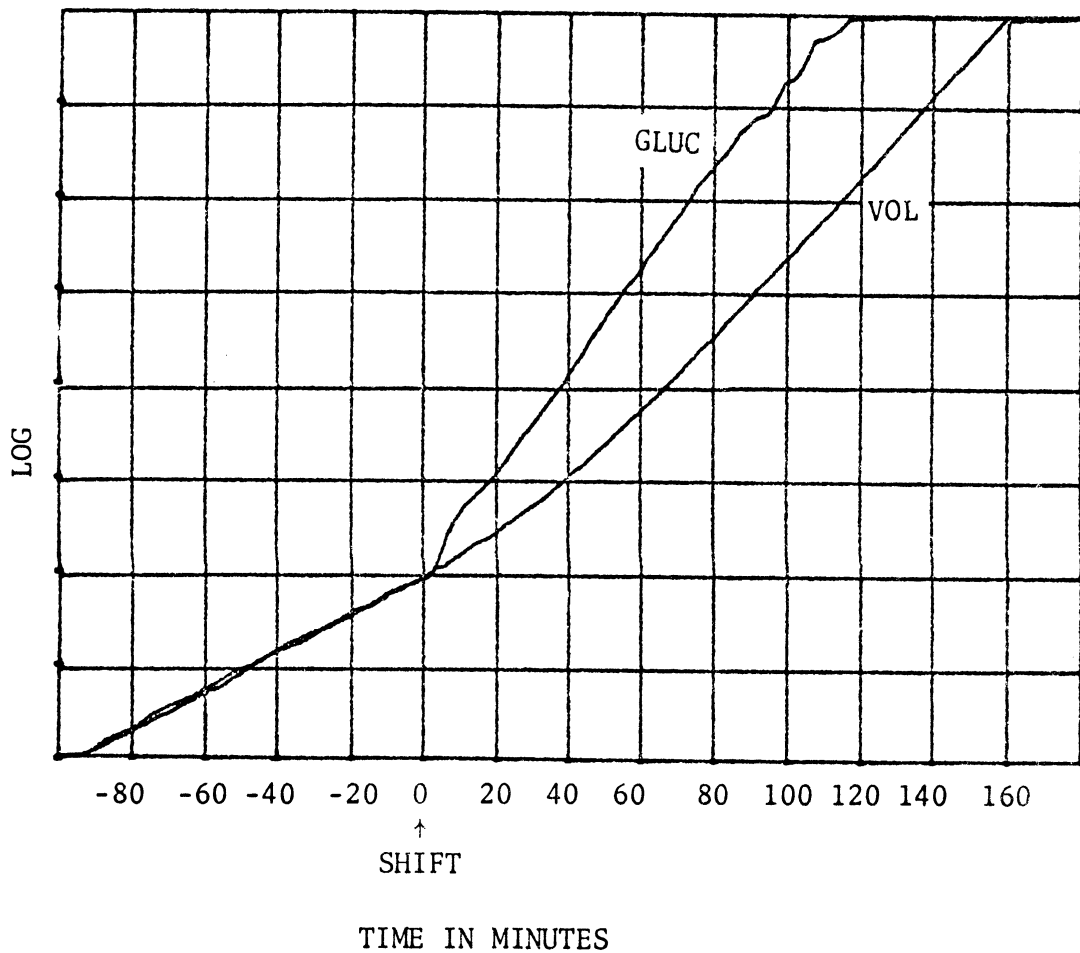in the age of the cell at the time of the shift.
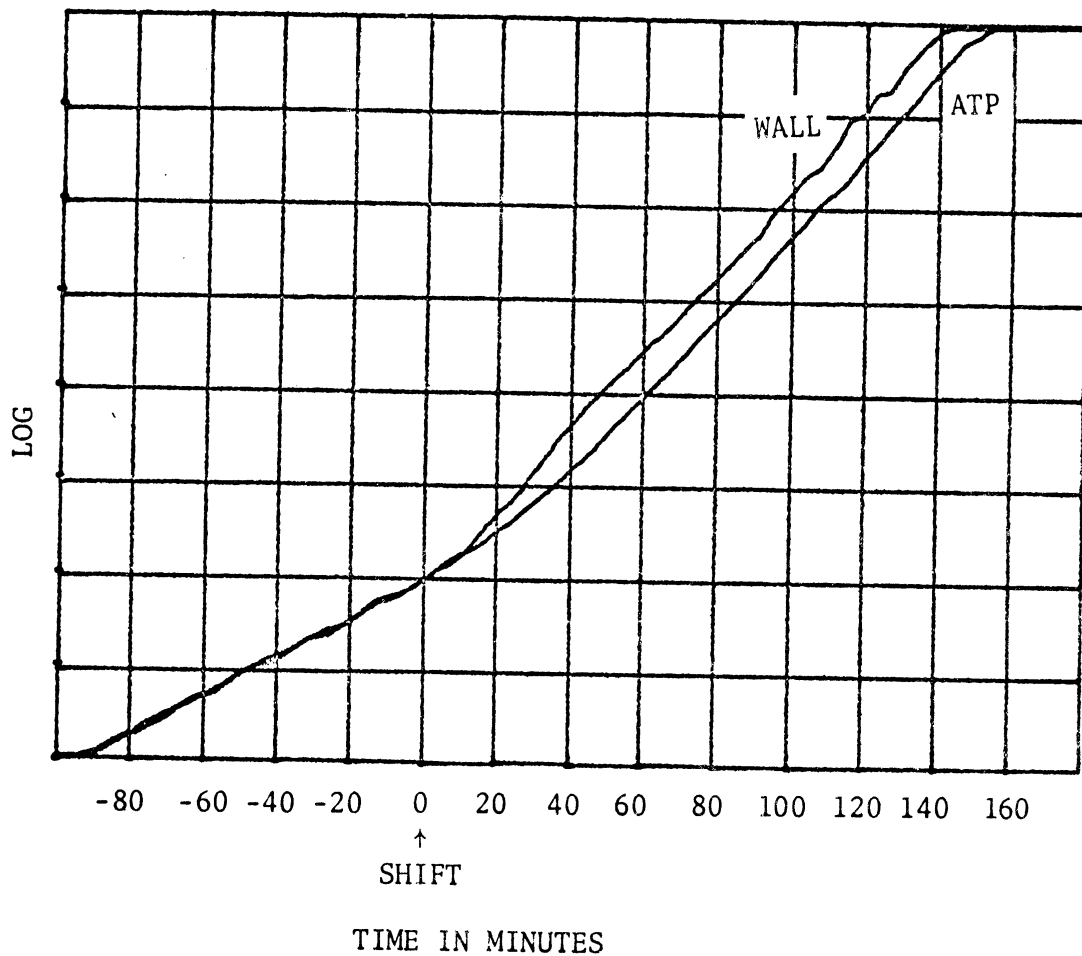
Figure A.10: WALL and ATP in Simulated Shift-up from One to Three. Logarithmic plot of amounts - units adjusted so curves for balanced growth (pre-shift) in environment one coincide. Distance between horizontal lines corresponds to one doubling. Each point is a sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.
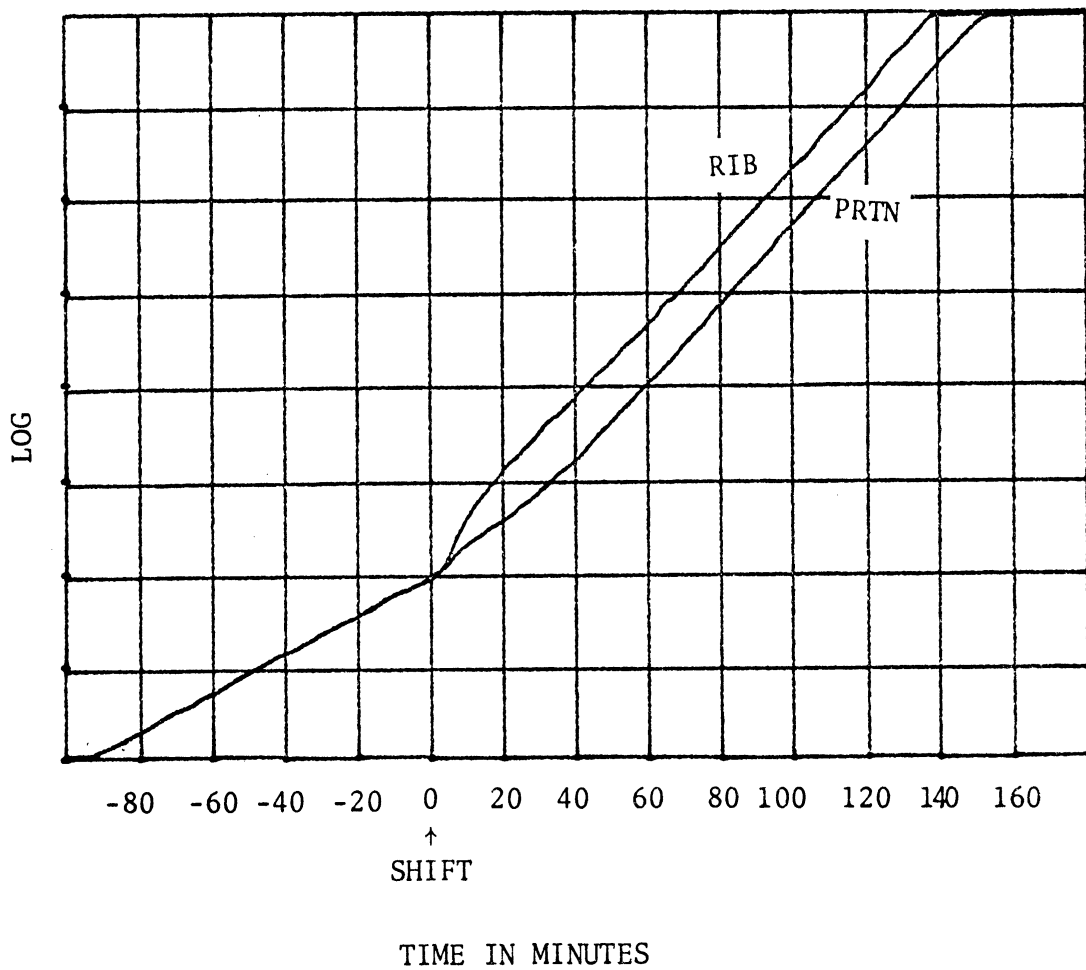
Figure A.11: RIB and PRTN in Simulated Shift-up from One to Three. Logarithmic plot of amounts - units adjusted so curves for balanced growth (pre-shift) in environment one coincide. Distance between horizontal lines corresponds to one doubling. Each point is a sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.
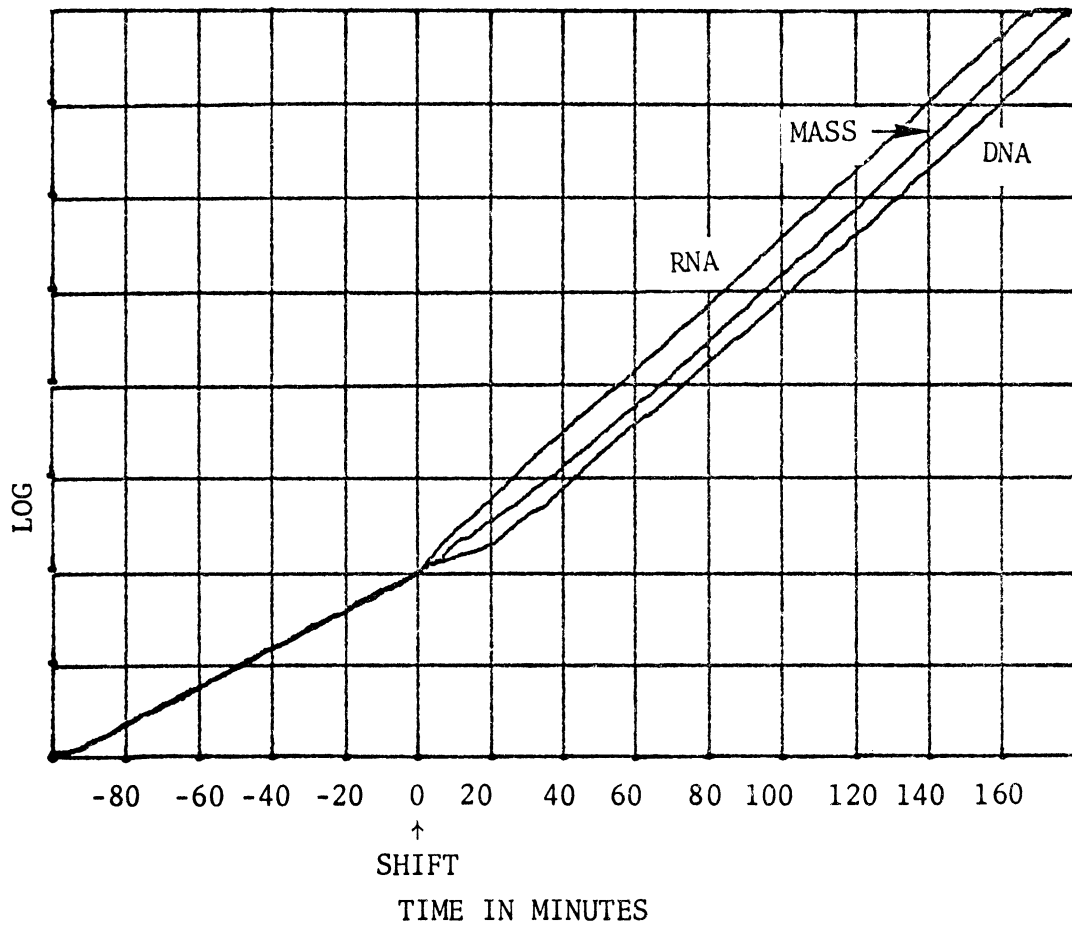
Figure A.12: MRNA and TRNA in Simulated Shift-up from One to Three. Logarithmic plot of amounts - units adjusted so curves for balanced growth (pre-shift) in environment one coincide. Distance between horizontal lines corresponds to one doubling. Each point is a sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.
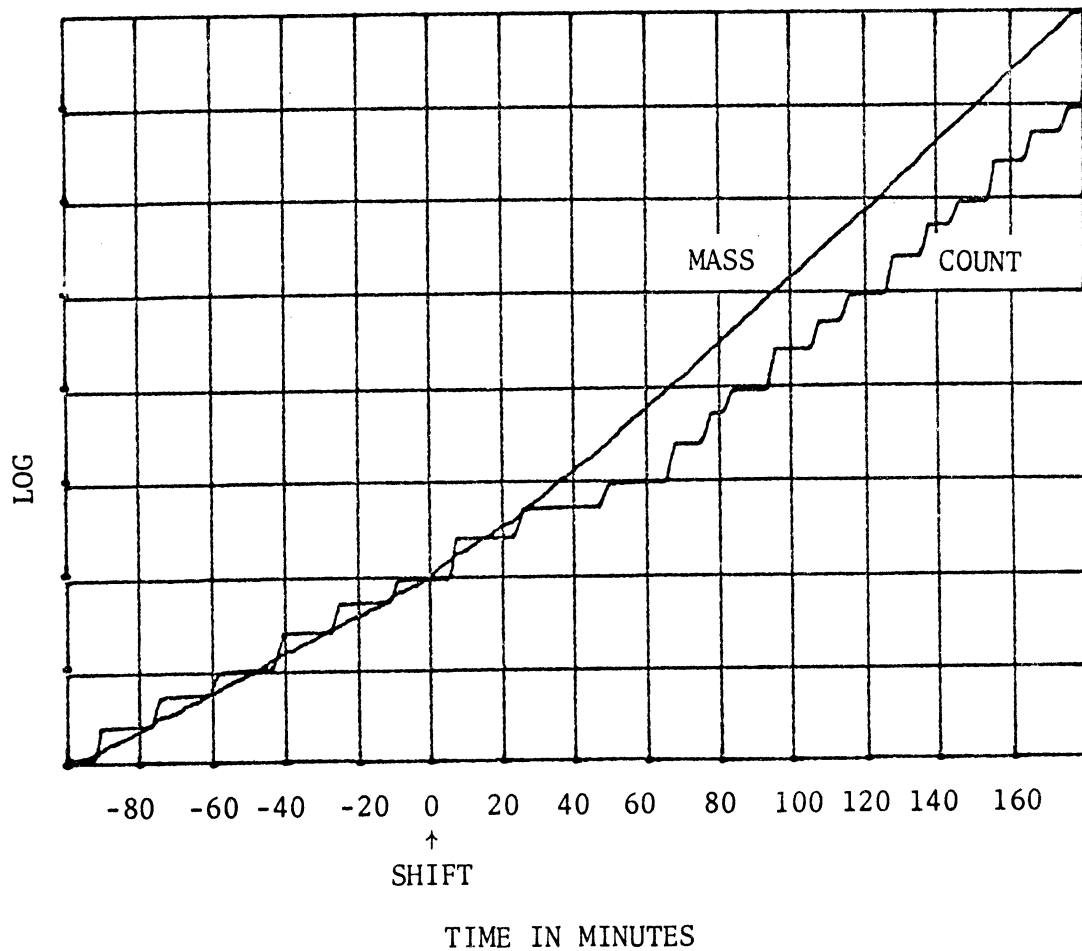
Figure A.13: GLUC and VOL in Simulated Shift-up from One to Three. Logarithmic plots of intracellular volume and amount of intracellular GLUC pool - units adjusted so curves for balanced growth (pre-shift) in environment one coincide. Distance between horizontal lines corresponds to one doubling. Each point is a sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.
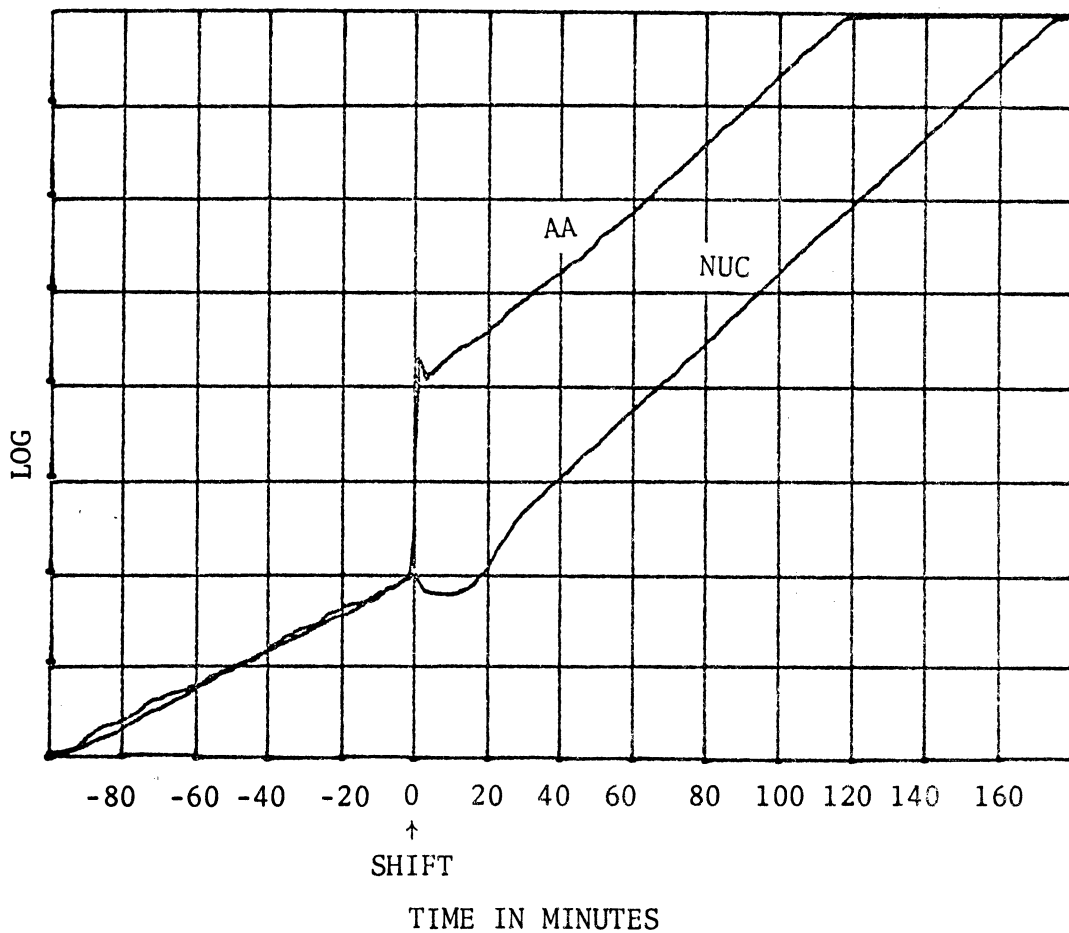
Figure A.14: RNA, DNA, MASS in Simulated Shift-up from One to Two.
Logarithmic plot of amounts - units adjusted so curves
for balanced growth (pre-shift) in environment one
coincide. Distance between horizontal lines corresponds
to one doubling. Each point is a sum of three values,
one calculated in each of three ONECELL runs. The
runs differ in the age of the cell at the time of the
shift.

Figure A.15: Mass and Cell Count in Simulated Shift-up from One to Two. Logarithmic plot of simulated total cellular mass and number of cells with units adjusted so curves for balanced growth (pre-shift) in environment one coincide. Distance between horizontal lines corresponds to one doubling. Each point is a sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.
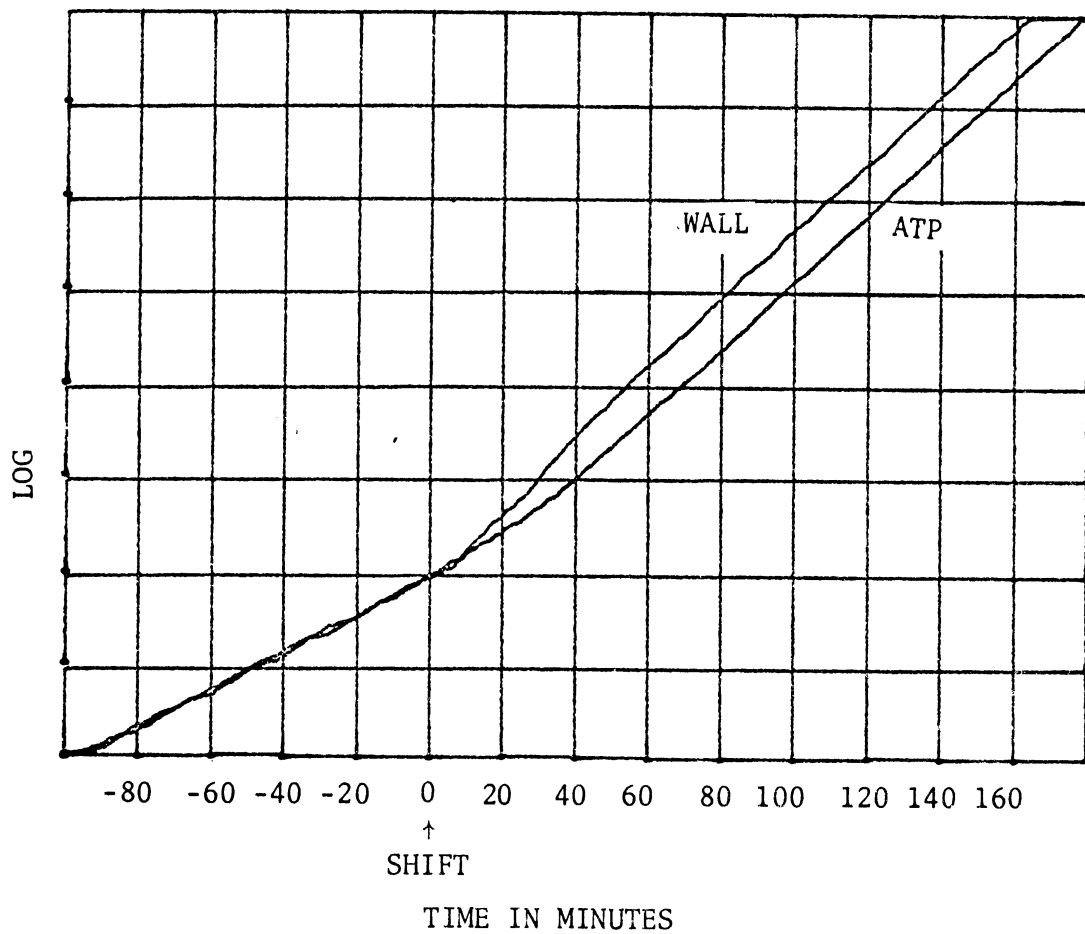
Figure A.16: AA and NUC im Simulated Shift-up from One to Two.
Logarithmic plot of amounts - units adjusted
so curves for balanced growth (pre-shift) in
environment one coincide. Distance between horizontal
lines corresponds to one doubling. Each point is a
sum of three values, one calculated in each of three
ONECELL runs. The runs differ in the age of the cell
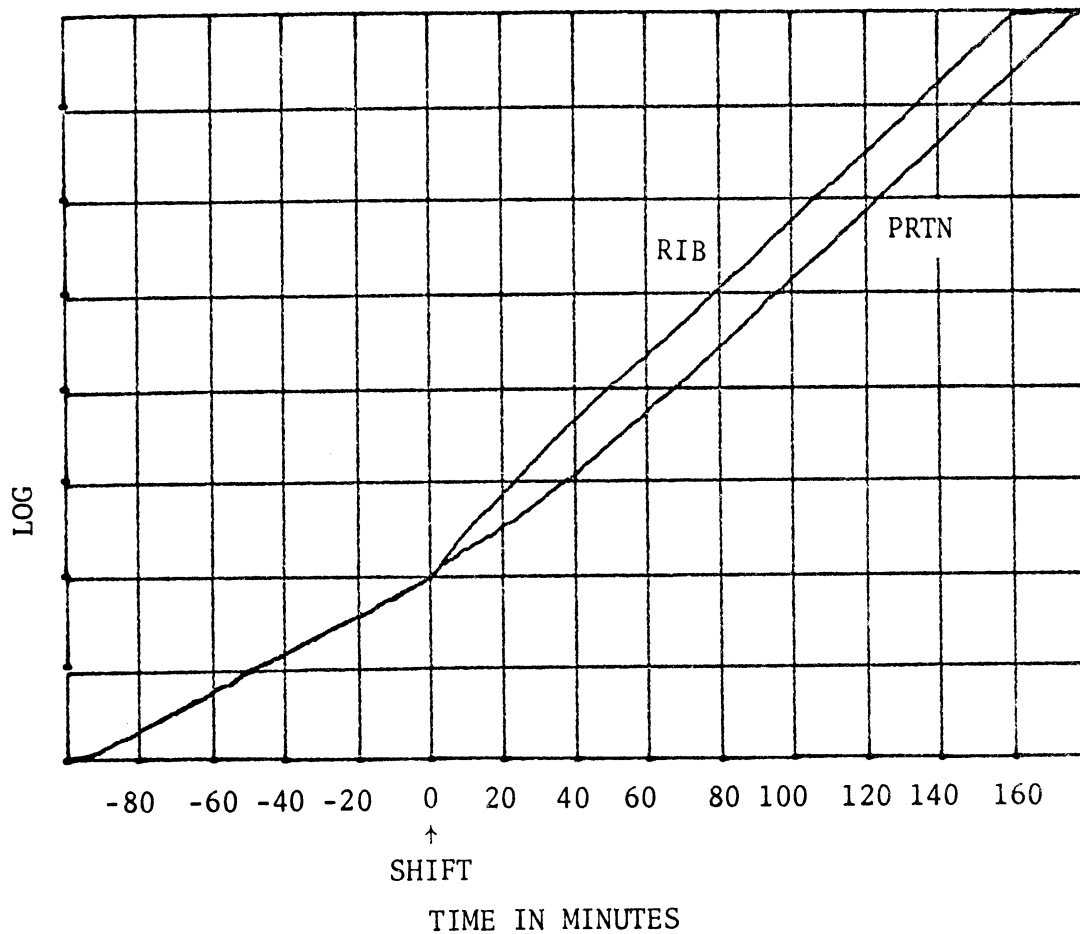at the time of the shift.

Figure A.17: WALL and ATP in Simulated Shift-up from One to Two.
Logarithmic plot of amounts - units adjusted so curves
for balanced growth (pre-shift) in environment one coincide.
Distance between horizontal lines corresponds to one doub-
ling. Each point is a sum of three values, one calculated
in each of three ONECELL runs. The runs differ in the
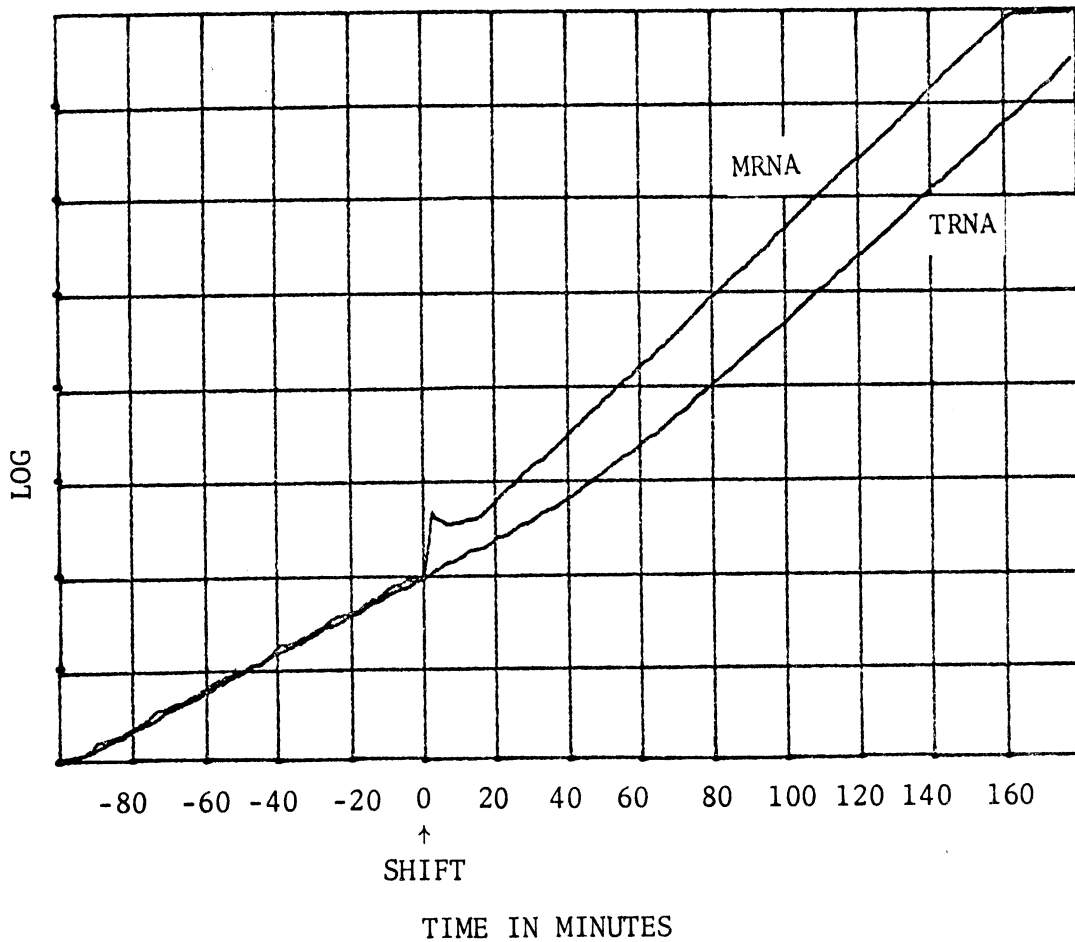age of the cell at the time of the shift.

Figure A.18: RIB and PRTN in Simulated Shift-up from One to Two.
Logarithmic plot of amounts - units adjusted so curves
for balanced growth (pre-shift) in environment one coincide.
Distance between horizontal lines corresponds to one
doubling. Each point is a sum of three values, one cal-
culated in each of three ONECELL runs. The runs differ
in the age of the cell at the time of the shift.

Figure A.19: MRNA and TRNA in Simulated Shift-up from One to Two.
Logarithmic plot of amounts - units adjusted so curves
for balanced growth (pre-shift) in environment one coincide.
Distance between horizontal lines corresponds to one
doubling. Each point is a sum of three values, one cal-
culated in each of three ONECELL runs. The runs differ
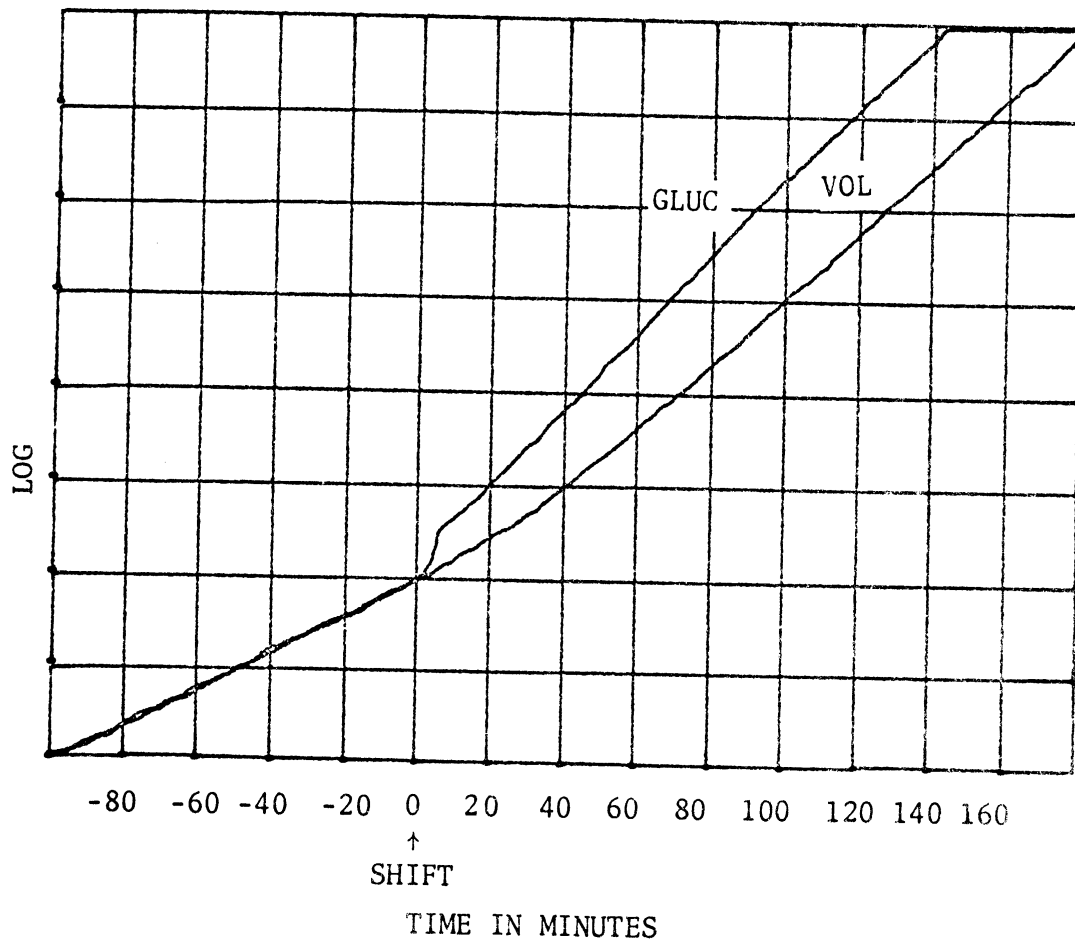in the age of the cell at the time of the shift.

Figure A.20: GLUC and VOL in Simulated Shift-up from One to Two.
Logarithmic plots of intracellular volume and amount
of intracellular GLUC pool - units adjusted so curves
for balanced growth (pre-shift) in environment one
coincide. Distance between horizontal lines corresponds
to one doubling. Each point is a sum of three values,
one calculated in each of three ONECELL runs. The runs
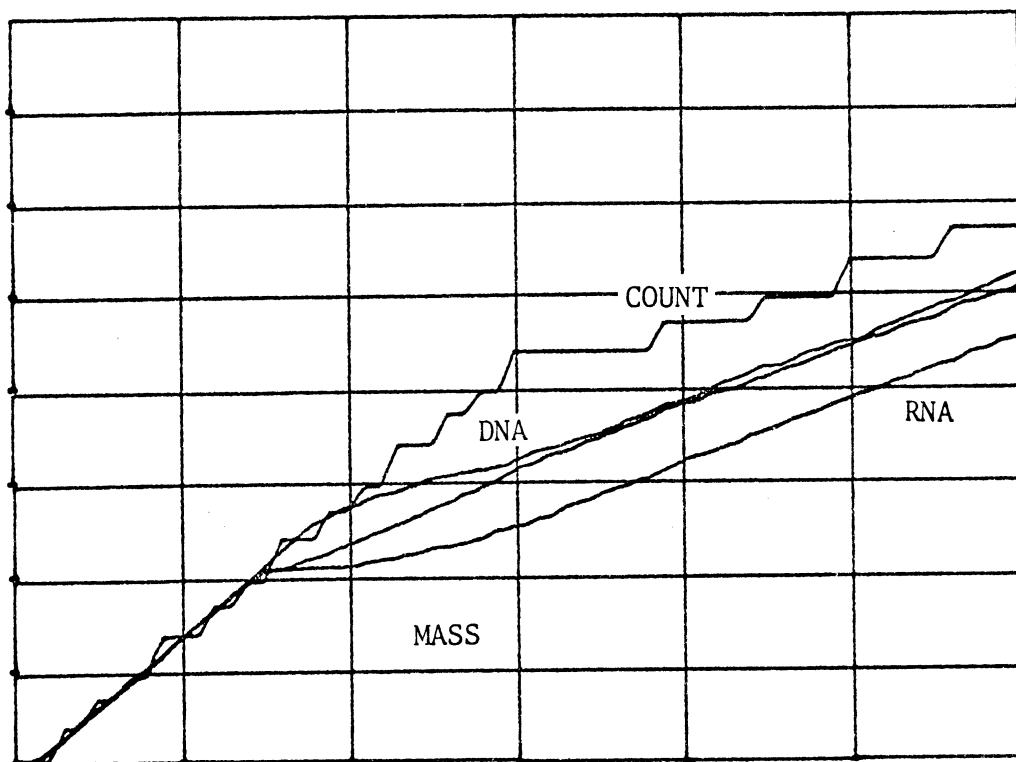differ in the age of the cell at the time of the shift.

Figure A.21: RNA, DNA, Mass, and Count in Simulated Shift-down from
Two to One. Logarithmic plots of amounts - units adjusted
so curves for balanced (pre-shift) growth in environment
Two coincide. Distance between horizontal lines corresponds
to one doubling. Each point is a weighted sum of three
values, one calculated in each of three ONECELL runs.
The runs differ in the age of the cell at the time of the
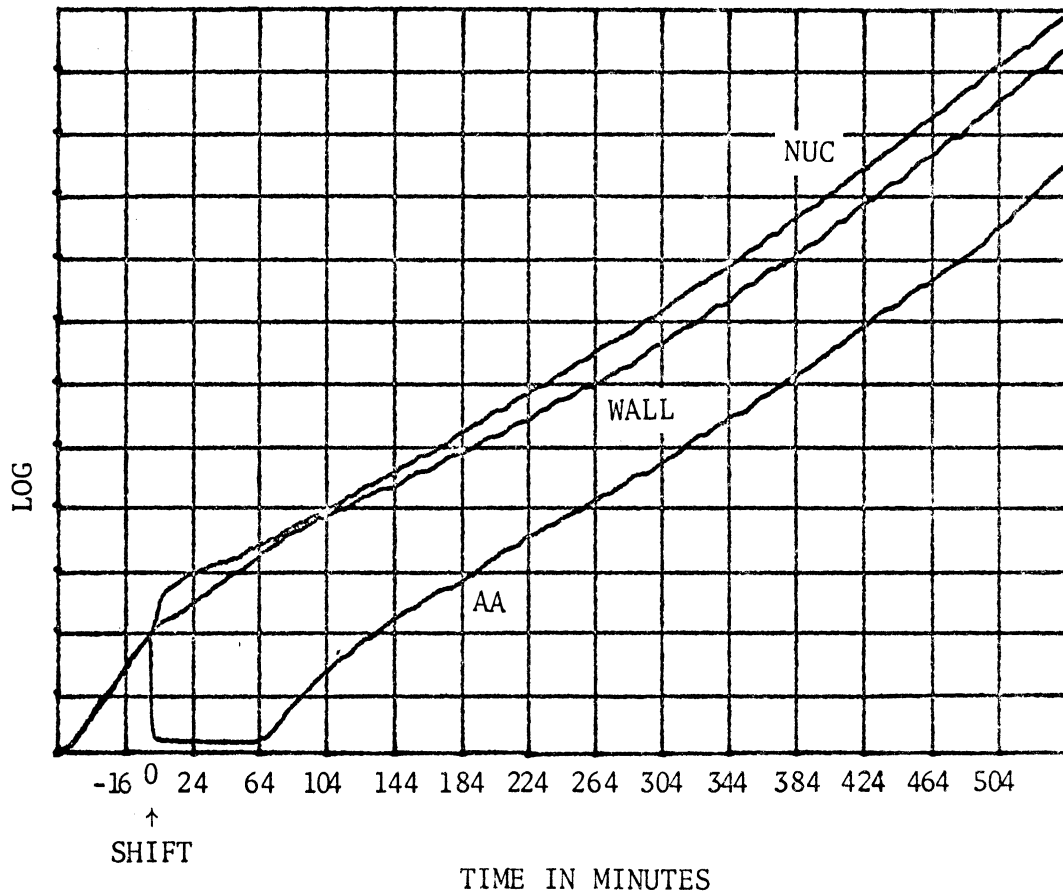shift.

Figure A.22: NUC, WALL, and AA   in Simulated Shift-down from Two to
One. Logarithmic plots of amounts - units adjusted so
curves for balanced (pre-shift) growth in environment
three coincide. Distance between horizontal lines corre-
sponds to one doubling. Each point is a weighted sum
of three values, one calculated in each of three ONECELL
runs. The runs differ in the age of the cell at the
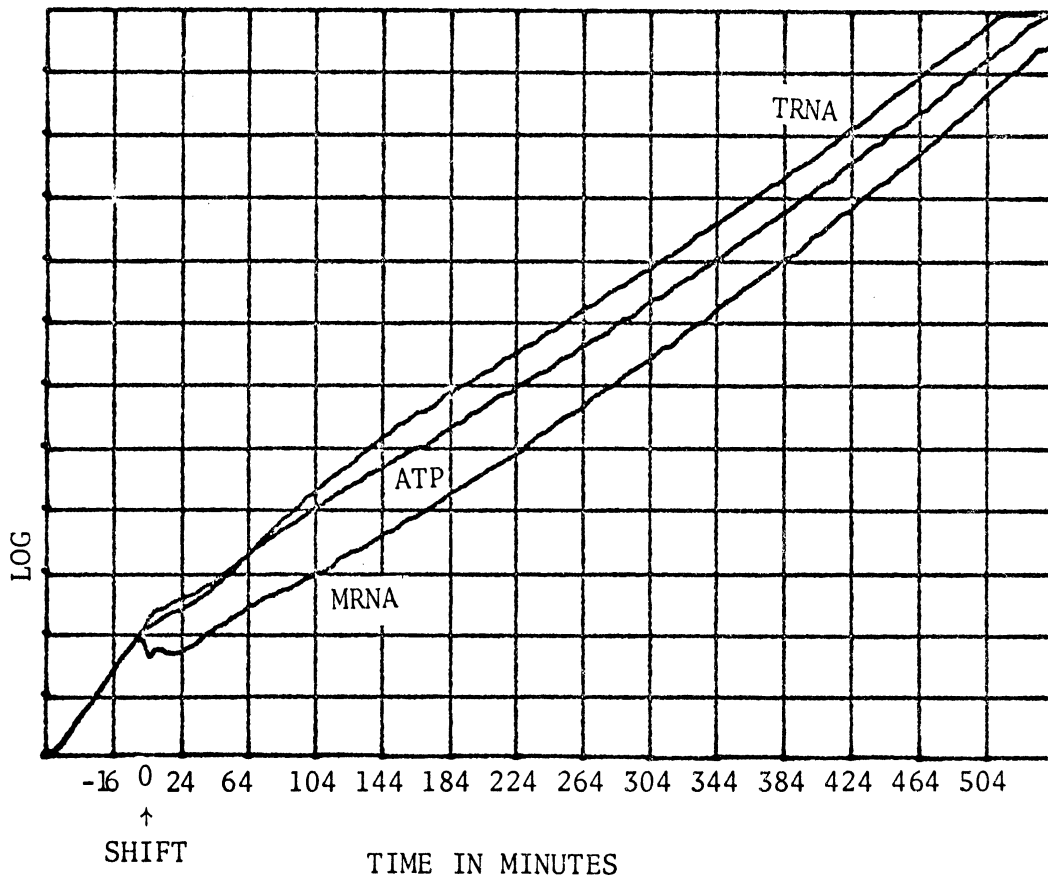time of the shift.

Figure A.23: TRNA, ATP, and MRNA in simulated Shift-down from Two to One. Logarithmic plots of amounts - units adjusted so curves for balanced (pre-shift) growth in environment three coincide. Distance between horizontal lines corresponds to one doubling. Each point is a weighted sum of three values, one calculated in each of three ONECELL runs. The runs differ in the age of the cell at the time of the shift.
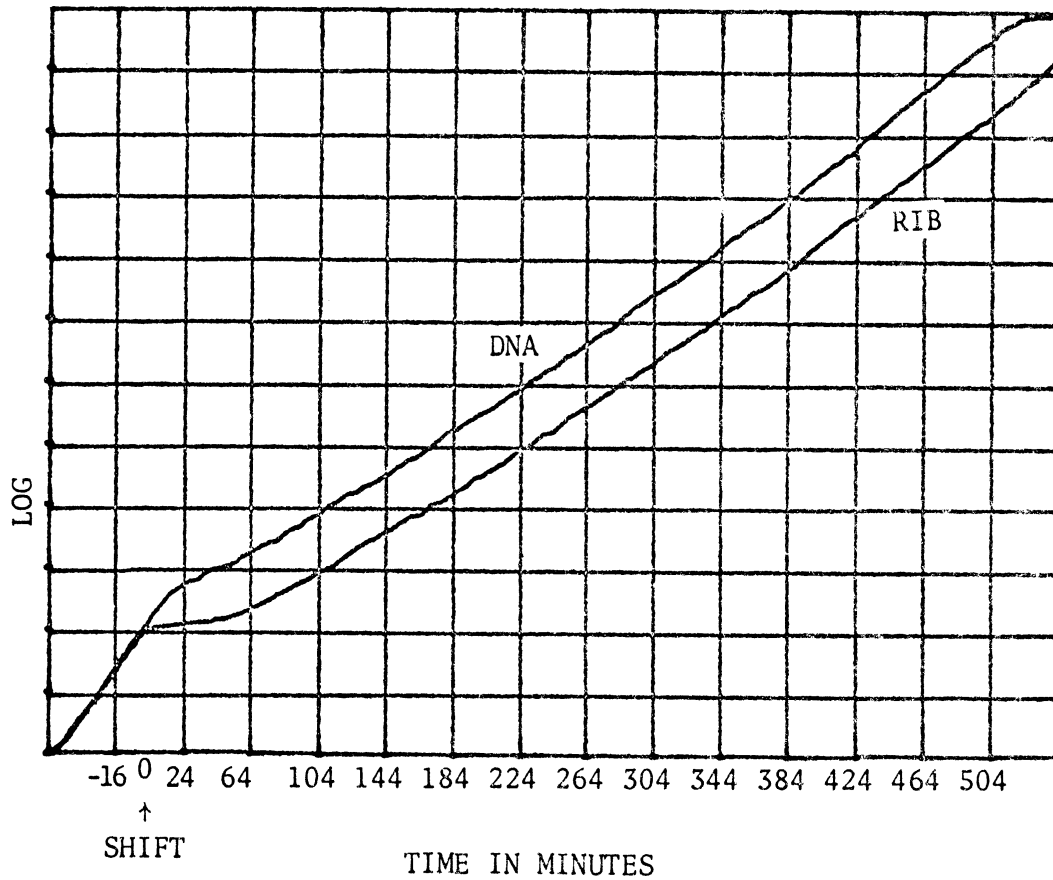
Figure A.24: DNA and RIB in Simulated Shift-down from Two to One.
Logarithmic plots of amounts - units adjusted so curves
for balanced (pre-shift) growth in environment three
coincide. Distance between horizontal lines corresponds
to one doubling. Each point is a weighted sum of three
values, one calculated in each of three ONECELL runs.
The runs differ in the age of the cell at the time of
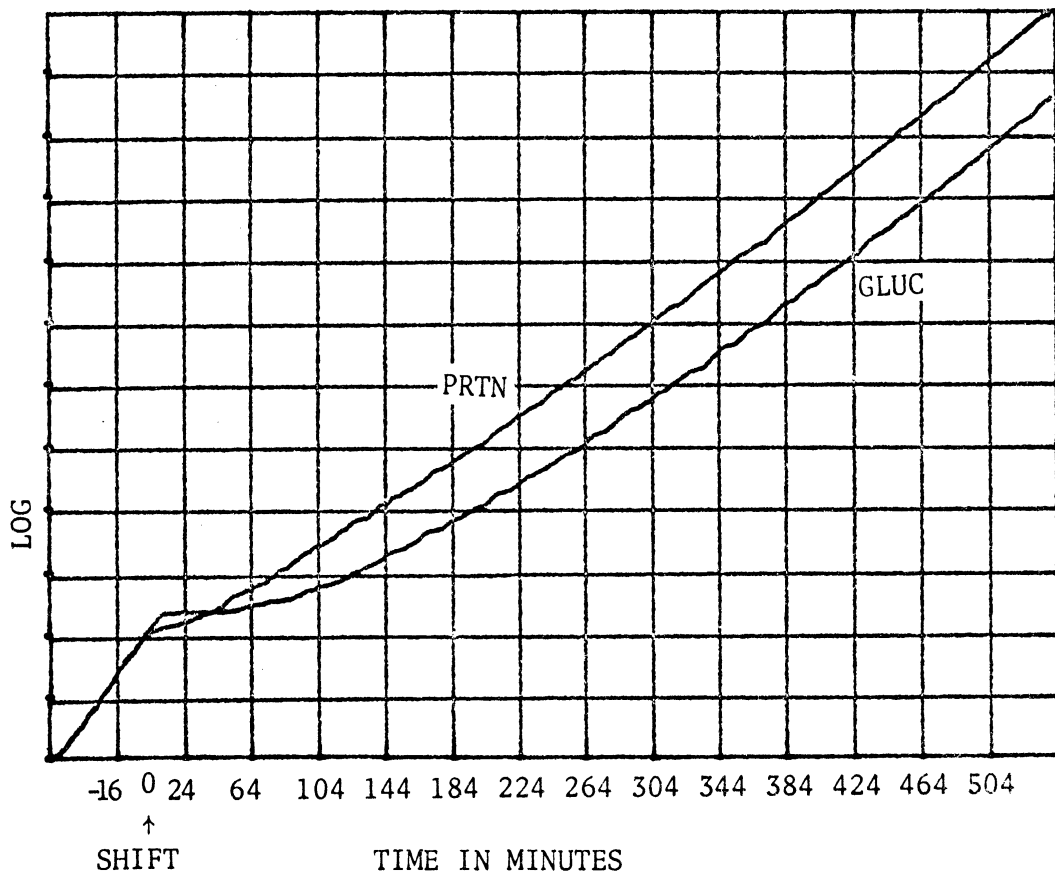the shift.

Figure A.25: PRTN and GLUC in Simulated Shift-down from Two to One.
Logarithmic plots of amounts - units adjusted so curves
for balanced (pre-shift) growth in environment three
coincide. Distance between horizontal lines corresponds
to one doubling. Each point is a weighted sum of three
values, one calculated in each of three ONECELL runs.
The runs differ in the age of the cell at the time of
the shift.

BIBLIOGRAPHY

Aiba, S.; Nagai, S.; Endo, I.; and Nishizawa, Y. *A.I.Ch.E.J.*, *15*, p. 624, 1969.

Beckwith, J.R. and Zipser, D. (eds.) *The Lactose Operon*, Cold Spring Harbor, New York, Cold Spring Harbor Laboratory, 1970.

Caswell, H.; Koenig, H.E.; Resh, J.; and Ross, Q. (unpublished manuscript), 1971.

Chance, B.; Garfinkel, D.; Higgins, J.; and Hess, B. *J. Biol. Chem.*, *235*, p. 2426, 1960.

Dean, A.C.R. and Hinshelwood, C. *Growth, Function and Regulation in Bacterial Cells*, Oxford, Clarendon Press, 1966.

Donachie, W.D. and Begg, K.J. *Nature*, *227*, p. 1220, 1970.

Eakman, J.M.; Fredrickson, A.G.; and Tsuchiya, H.M. *Chem. Engr. Prog. Symp. Series*, *69*, pp. 37-49, 1966.

Fredrickson, A.G.; Ramkrishna, D.; and Tsuchiya, H.M. *Mathematical Biosciences*, *1*, pp. 327-374, 1967.

Goodwin, B.C. *Temporal Organization in Cells*, Academic Press, New York, 1963.

Gordon, G. *System Simulation*, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

Gunsalus, I.C. and Shuster, C.W. In *The Bacteria*, *II*, Gunsalus, I.C. and Stanier, R.Y. (eds.) Academic Press, New York, 1961.

Heinmets, F. *Analysis of Normal and Abnormal Cell Growth*, Plenum, New York, 1966.

Helmstetter, C.; Cooper, S.; Pierucci, O.; and Revelas E. *Cold Spring Harbor Symposia on Quantitative Biology*, *XXXIII*, pp. 809-822, 1968.

Kelley, W.S. and Schaechter, M. *Adv. in Microbial Physiol.*, *2*, pp. 89-142, 1968.

Kepes, A. and Cohen, G.N. In *The Bacteria*, *IV*, Gunsalus, I.C. and Stanier, R.Y. (eds.) Academic Press, New York, 1962.

Knorre, W.A. In *Continuous Cultivation of Microorganisms*, Málek, I.; Bacon, K.; Fencl, Z.; Munk, V.; Ricica, J.; and Smrcková, H. (eds.) Akademia, Prague, 1969.

Kjeldgaard, N.O.   In *Advances in Microbial Physiology*, *I*, Rose, A.H. and Wilkinson, J.F. (eds.) Academic Press, London, pp. 39-96, 1967.

Laing, R.A. Technical Report 012520-3-T, The University of Michigan, 1970.

Langer, R.; Krohn, K.; and Rhodes, J. In *Systems Theory and Biology*, Mesarović, M.D. (ed.) Springer-Verlag, New York, pp. 130-140, 1968.

Lindenmayer, A. *J. Theoret. Biol.*, *18*, pp. 280-299, 1968.

Luria, S.E. In *The Bacteria*, *I*, Gunsalus, I.C. and Stanier, R.Y. (eds.) Academic Press, New York, pp. 1-34, 1960.

Maaløe, O. and Kjeldgaard, N.O. *Control of Macromolecular Synthesis*, Benjamin, New York, 1966.

Magasanik, B. In *The Lactose Operon*, Beckwith, J.R. and Zipser, D. (eds.) Cold Spring Harbor, New York, Cold Spring Harbor Laboratory, 1970.

Mandelstamm, J. and McQuillen, K. (eds.) *Biochemistry of Bacterial Growth*, Wiley, New York, 1968.

Marr, A.G.; Painter, P.R.; and Nilson, E.H. *Microbial Growth*, Nineteenth Symposium of the Society for General Microbiology. Cambridge University Press, pp. 237-262, 1969.

Mesarović, M.D. *Systems Theory and Biology*, Mesarović, M.D. (ed.) Springer-Verlag, New York, pp. 59-87, 1968.

Pirt, S.J. *J. Gen. Microbiol.*, *47*, pp. 181-197, 1967.

Powell, E.O. In *Microbial Growth and Continuous Culture*. *Proceedings of the Third International Symposium*. Her Majesty's Stationary Office, London, p. 34, 1967.

Ramkrishna, D.; Fredrickson, A.G.; and Tsuchiya, H.M. *Biotech. and Bioeng.*, *IX*, pp. 129-170, 1967.

Rashevsky, N. *Mathematical Biophysics*, Dover, New York, 1960.

Rosen, R. *International Review of Cytology*, *23*, pp. 25-88, 1968.

____, *Dynamical System Theory in Biology*, *I*, Wiley-Interscience, New York, 1970.

Sanwal, B.D. *Bacteriological Reviews*, *34*, 1, pp. 20-39, 1970.

Schaechter, M.; Maaløe, O. and Kjeldgaard, N.O.   *J. Gen. Microbiol.*, *19*, pp. 592-606, 1958.

Simon, Z.   "Bacterial Cell Model.   Cell Cycle Parameters and Macromolecular Synthesis" preprint, 1970.

Stahl, W.R.   *J. Theoret. Biol.*, *9*, p. 371, 1965.

Tempest, D.W. and Ellwood, D.C.   *Biotech. and Bioeng.*, *11*, pp. 775-783, 1969.

Ward, C.B. and Glaser, D.A,   *Proc. Nat. Acad. Sci. USA*, *67*, P. 225, 1970.

___, *Proc. Nat. Acad. Sci, USA*, *68*, 5, pp. 1061-1064, 1971.

Watson, J.D.   *Molecular Biology of the Gene*, Second Edition, Benjamin, New York, 1970.

Weinberg, R. and Berkus, M.   Technical Report 01252-2-T, The University of Michigan, 1969.

Weinberg, R. and Zeigler, B.P.   *J. of Cybernetics*, *1*, 2, 1971.

Whittaker, D.K. and Drucker, D.B.   *J. Bacteriol.*, *104*, pp. 902-909, 1970.

Williams, L.S. and Neidhart, F.C.   *J. Mol. Biol.*, *43*, pp. 529-550, 1969.

Yates, F.E.; Brennan, R.D.; Urquhart, J.; Li, C.C.; and Halpern, W. In *Systems Theory and Biology*, Mesarović, M.D. (ed.) Springer-Verlag, New York, pp. 141-184, 1968.

Zeigler, B.P. and Weinberg, R.   *J. Theoret. Biol.*, *29*, pp. 35-56, 1970.