Dynamic Simulations of Single-Molecule Enzyme Networks

Dieter Armbruster,[*,†,‡] **John D. Nagy,**[§] **E. A. F. van de Rijt,**[†] **and J. E. Rooda**[†]

Department of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513, NL-5600 MB, Eindhoven, The Netherlands; Department of Mathematics and Statistics, Arizona State University, Tempe, Arizona 85287-1804; and Department of Life Sciences, Scottsdale Community College, 9000 E. Chaparral Rd., Scottsdale, Arizona 85256

Along with the growth of technologies allowing accurate visualization of biochemical reactions to the scale of individual molecules has arisen an appreciation of the role of statistical fluctuations in intracellular biochemistry. The stochastic nature of metabolism can no longer be ignored. It can be probed empirically, and theoretical studies have established its importance. Traditional methods for modeling stochastic biochemistry are derived from an elegant and physically satisfying theory developed by Gillespie. However, although Gillespie's algorithm and its derivatives efficiently model small-scale systems, complex networks are harder to manage on easily available computer systems. Here we present a novel method of simulating stochastic biochemical networks using discrete events simulation techniques borrowed from manufacturing production systems. The method is very general and can be mapped to an arbitrarily complex network. As an illustration, we apply the technique to the glucose phosphorylation steps of the Embden−Meyerhof−Parnas pathway in *E. coli*. We show that a deterministic version of the discrete event simulation reproduces the behavior of an analogous deterministic differential equation model. The stochastic version of the same model predicts that catastrophic bottlenecks in the system are more likely than one would expect from deterministic theory.

1. Introduction

Following advances in fluorescence correlation spectroscopy[1] and other optical methods,[2] it is now possible to study enzymatic reactions at the level of single molecules. At this scale, the usual description of such reactions via rate equations breaks down, so more appropriate stochastic models of single-molecule Michaelis−Menten kinetics have been developed recently.[3,4] In parallel with this line of research, interest in statistical fluctuations and their effects on intracellular biochemical dynamics, including calcium fluctuations,[5−11] genetic regulatory pathways, and glucose catabolism,[12] has recently blossomed. Theoretical work has demonstrated that random fluctuations can enrich biochemical reaction dynamics in nontrivial ways, especially in systems of modest size (10^2 to 10^6 interacting molecules). For example, small-amplitude stochastic fluctuations can generate sustained, large-amplitude oscillations even when the most appropriate deterministic description has no limit cycle.[9,12,13] This observation has relevance to intracellular biochemical dynamics since many such systems fit within this size range.[12] Therefore, one should not ignore randomness in either intracellular or experimental domains. This point was also strongly made by Rao et al.[14] in their review article.

The recent article by Kou et al.[3] gives a detailed discussion of the relationship between traditional Michaelis−Menten kinetics and the stochastic processes underlying any catalytic reaction. Kou et al. formulate their model to describe the dynamics of the probability distribution of waiting times for completion of a reaction catalyzed by a single enzyme over many events and a long period of time. In their simplest model

they make two key assumptions: enzyme availability is the rate-limiting factor, and the concentration of substrate is unaffected by the action of a single enzyme. In a similar model, Qian and Elson[4] show that, for certain substrate concentrations, such single-molecule enzyme reactions may become oscillatory.

In this paper, we build on those two approaches[3,4] and propose a new approach to simulations of *biochemical networks* consisting of a small number of enzymes and substrate molecules. We apply simulation techniques developed in industrial and production systems engineering to biochemical reaction pathways and show how they may be useful in simulating highly interconnected, nonlinear, nonlocal and nonstationary reaction pathways, which offer no hope of an analytic solution of the underlying evolution equations for the probability densities of individual reactions.

Typically, randomly fluctuating intracellular biochemical pathways are modeled using the Gillespie algorithm,[15,16,19] essentially a Markov chain stochastic process in continuous time. Such systems can be analyzed exactly to first order using the van Kampen expansion approach[9,13,20] regardless of the reaction system's details. However, treating the network dynamics as a Markov chain ignores the details of the single-molecule Michaelis−Menten reaction discussed by Kou et al. and Qian and Elson.[3,4] In addition, the Gillespie algorithm, despite advances in efficiency,[21] becomes prohibitively tedious on easily available computers for large systems like glycolysis and most other biochemical and genetic pathways. This has led to several accelerated discrete stochastic or multiscale stochastic simulation packages and to hybrid models. A state of the art review is presented in ref 17. All of these approaches, to our knowledge, are based on the Markov assumption.

The technique we introduce here represents a departure from standard models of biochemical reaction systems, although it is well-known in models of manufacturing operations.[22] The

---

* To whom correspondence should be addressed. E-mail: armbruster@asu.edu.
† Eindhoven University of Technology.
‡ Arizona State University.
§ Scottsdale Community College.

approach is to model a biochemical or genetic reaction system as a collection of concurrent or parallel processes connected by interacting channels. In that way, a biochemical network becomes a supply chain or a factory whose nodes (the machines) represent enzymatic processes and whose links represent possible pathways. We will show that we can easily give these nodes sufficiently detailed process structure to model single-enzyme Michaelis–Menten kinetics as well as higher order processes involving more than one substrate and more than one product. The resulting stochastic simulation can be done in standard discrete event simulators, allowing efficient computation of large, multistep pathways while retaining the core biophysical properties of the system.

Our approach is closely related to the paper by Kühl and Jobmann.[18] They argue that receptor–agonist interactions are best described by service-theoretic methods based on queueing theory. They generate different dose–response curves by changing the stochastic properties of the queueing models, emphasizing especially the influence of non-Markovian stochastics on the shape of these curves. We extend their approach in two ways: (i) by employing simulation tools derived from production systems engineering; and (ii) by generalizing from single ligand–receptor events to entire biochemical pathways involving a cascade of such events, feedbacks from one part of the network to another, and capacity constraints on individual enzymes (machines).

The remainder of our paper is organized as follows. We begin in section 2 by introducing the fundamentals of production systems simulation via multiple parallel processes and their application to biochemical networks. In section 3 we illustrate our approach by simulating a transient experiment on a portion of the Embden–Meyerhof–Parnas pathway in *E. coli* (glycolysis in eucaryotes). We discuss the discrete event simulation setup and detailed modeling of specific enzymatic reactions involved before moving on to a validation of the model by comparison with a standard system of rate equations. In section 4 we perform the data analysis of a large set of experiments and determine the resulting changes of the transient experiment due to fluctuations ignored in the ODE model. We show that one can simulate the pathway in several different modes of operation and determine the stability of the system as a function of the biochemical fluxes. We conclude in section 5 with an outlook on other biological networks that process a small number of molecules—in particular, gene regulation and signal transduction networks—that may also be studied with stochastic discrete event simulations.

2. Simulating Production Systems

Typical manufacturing systems or supply chains consist of many different parallel processes along with channels that feed information and/or material to the processes. Those channels create a network with inflows and outflows. A system is characterized at a given time t by its number of concurrent processes and their states at time t. State variables usually change instantaneously at separate points in time[22] leading to the notion of discrete event models (DEMs). Time in DEMs is represented by a variable called the "simulation clock". In many DEMs, similarly to the Gillespie algorithm, each "tick" of the simulation clock represents the time at which the next event occurs. For instance, suppose that at time t a machine begins working on a part (an enzyme molecule binds to a molecule of its substrate), and that it will complete the product at time $t + \tau$, where the choice of τ depends on some probability density. The simulation clock then advances τ time units, and the next event—starting
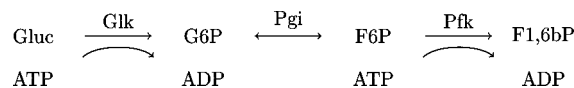
Figure 1. Glucose phosphorylation steps of the Embden–Meyerhof–Parnas pathway. See Table 1 for definitions of abbreviated names.

another product, setting up for a production process, breaking down or idling, for example—is chosen. In contrast, a time-based simulation typically describes the continuous evolution of a state variable using a differential equation. Such a simulation's computational cost is the product of a function of the number of time steps and a function of the number of concurrently evolving state variables. In particular, for multiscale parallel systems the resolution of the fastest time scale usually determines the time step of the simulation algorithm on the whole network. Discrete event models, on the other hand, scale with the total number of events in the system, which in multiscale parallel systems typically results in far less computational cost than that of the analogous continuous description.

Discrete event simulations are a special case of a process algebra,[23] which is the study of the behavior of parallel or distributed systems by algebraic means. Process algebras offer means to describe or *specify* such systems, and thus they have means to specify parallel composition. In addition, they can usually specify alternative composition (put things in a choice) and sequential composition (sequencing, put things one after the other). Most importantly, it is possible to reason about such systems using algebra, i.e., equational reasoning. By means of this equational reasoning, *verification* becomes possible; i.e., it can be established that a system satisfies a certain property.

DEMs exist at all levels of detail since the notion of a process and an event is general enough to cover any input–output behavior of an entire factory as well as the characteristic processes of a specific production step on a single machine. So, DEMs have the advantage of easily incorporating stochastic behavior into a variety of simulation processes. It is completely natural that the occurrence of an event like the failure of a machine is governed by a sample from a stochastic process, making DEMs a model of choice for stochastic simulations. The highly efficient nature of such simulations then allows many sample runs, which facilitates statistical evaluations of stochastic experiments.

Many different simulation platforms exist for discrete event simulations. In this paper we use the χ formalism,[24] and a simulation tool for discrete event systems based on χ.[25] This tool has been successfully applied to a large number of industrial cases, such as integrated circuit manufacturing plants, breweries, and process industry plants.

3. Transient Simulations for a Simple Metabolic Network

Escherichia coli catabolizes glucose in part using the Embden–Meyerhof–Parnas (EMP) pathway[26] to generate ATP. Here we focus only on the glucose phosphorylation steps of the EMP pathway, from glucose to fructose-1,6-bisphosphate (Figure 1). Table 1 shows all abbreviations used for enzymes and intermediate products of this section of the EMP pathway.

We choose this pathway because it is well-known, central to the metabolism of nearly every cell, and sufficiently complex to demonstrate the technique's usefulness. However, to keep things simple, we assume that each step in the pathway is controlled by a single enzyme. Of course, such is not the case in a single *E. coli*. We emphasize that the goal of this paper is to introduce a useful simulation paradigm in systems biology,

TABLE 1: Abbreviations Used for Enzyme and Substrate Names

(a) enzyme names		(b) substrate names	
	full name		full name
Glk	glucokinase (hexokinase IV)	Gluc	glucose
Pgi	phosphoglucoisomerase	G6P	glucose-6-phosphate
Pfk	phosphofructokinase	F6P	fructose-6-phosphate
		F1,6bP	fructose-1,6-bisphosphate
		ATP	adenosine triphosphate
		ADP	adenosine diphosphate

not to study the EMP pathway in detail, although this can be done with the technique we are illustrating.

3.1. ODE Model. Before entering the stochastic realm, we lay certain foundations for comparison by first developing the deterministic theory. Ishii et al.[27] studied essentially the same pathway we do but with deterministic ordinary differential equations. They then compare the numerical solutions to experimental data. In their model, Ishii et al. assume a well-mixed volume in thermal equilibrium. Their initial [Gluc] = 1.0 mM and [ATP] = 2.0 mM in a volume of $(\pi/2) \times 10^{-15}$ L, corresponding to the volume of one *E. coli* cell. Further, they assumed that the concentrations of each enzyme, Glk, Pgi, and Pfk, were 0.05 μM. Concentrations of all intermediate products, namely G6P, F6P, ADP, and F1,6bP, were initially zero. The experiment focused on the time evolution of glucose consumption and the eventual buildup of F1,6bP.

Adopting the same assumptions and goals as Ishii et al., we study the following model, which is a slight modification of theirs:

$$\frac{dC_{\text{Gluc}}}{dt} = -v_{\text{glk}} \tag{1a}$$

$$\frac{dC_{\text{G6P}}}{dt} = v_{\text{glk}} - v_{\text{pgi}} \tag{1b}$$

$$\frac{dC_{\text{F6P}}}{dt} = v_{\text{pgi}} - v_{\text{pfk}} \tag{1c}$$

$$\frac{dC_{\text{F1,6bP}}}{dt} = v_{\text{pfk}} \tag{1d}$$

$$\frac{dC_{\text{ATP}}}{dt} = -v_{\text{glk}} - v_{\text{pfk}} \tag{1e}$$

$$\frac{dC_{\text{ADP}}}{dt} = v_{\text{glk}} + v_{\text{pfk}} \tag{1f}$$

where the fluxes are given by

$$v_{\text{Glk}} = \frac{V_{\text{Glk}}C_{\text{Gluc}}C_{\text{ATP}}}{(K_{\text{Gluc}} + C_{\text{Gluc}})(K_{\text{ATP}_1} + C_{\text{ATP}})} \tag{2}$$

$$v_{\text{Pgi}} = \frac{V_{\text{Pgi}}C_{\text{G6P}}}{K_{\text{G6P}}C_{\text{G6P}}} \tag{3}$$

$$v_{\text{Pfk}} = \frac{V_{\text{Pfk}}C_{\text{F6P}}^{n}C_{\text{ATP}}}{(K_{\text{Pfk}}^{n} + C_{\text{F6P}}^{n})(K_{\text{ATP}_2} + C_{\text{ATP}})} \tag{4}$$

where C_i represents the concentration of compound i at (implicit) time t, and all V_{max} and K_{m} are constants.

Model (1, 2−4) is identical to Ishii et al.'s except that here we ignore the backward reaction from F6P to G6P. Parameter values used by ref 27 can be found in Table 2. The experiment thus becomes a standard initial value problem which can be solved numerically. Figure 2 shows the resulting dynamics, characterized by a decrease in [Gluc] and the subsequent increase and slow decrease of the intermediate products as they "bleed off" into the downstream portion of the pathway. These results are in general qualitative agreement with Ishii et al.'s. We will proceed now to build the analogous manufacturing system description of model (1, 2−4).

3.2. EMP Pathway as a Manufacturing System. Conceptually, we map a metabolic network such as the EMP pathway, which is a series of enzymatic reactions, to a production system in manufacturing, which comprises a series of processes that convert raw products into final products. For example, the section of the EMP pathway studied in the previous section consists of a sequence of processes (enzymatic reactions, analogous to machines in the plant) that convert glucose to F1,6bP. The fundamental unit for such a manufacturing network is a queue, which comprises a stochastic arrival process, determined by the stochastic behavior of upstream processes, of intermediate products into a "buffer," and a machine (an enzyme), which takes items from its buffer and handles them as described by a stochastic production process. In production simulation terminology, the relevant quantities used to describe such a system are throughput, cycle time, arrival rate, and work in progress (WIP). The associated biological terms are flux, turnover time for a single-molecule experiment,[3] incoming flux, and concentration of an intermediate product. The relationship between throughput F and WIP W is central in both realms—in manufacturing systems this is generally called the clearing function $F(W)$. For a queue with a steady-state length W and a machine with a Markovian production rate μ, F is parametrized by μ as follows:

$$F = \frac{1}{\mu}\frac{W}{1 + W} \tag{5}$$

More general clearing function models involve two parameters, e.g.

$$F = \frac{1}{\mu}\frac{W}{\alpha + W} \tag{6}$$

or

$$F = \frac{1}{\mu}\frac{W}{1 - e^{-\beta W}} \tag{7}$$

Obviously eq 6 is the Michaelis−Menten description of a single-molecule enzyme with $v_{\text{max}} = 1/\mu$ and α as the Michalis−Menten parameter.

TABLE 2: Parameter Values Used in ODE Numerical Solutions, Obtained from Ref 27 and Based on *E. coli*

enzyme	parameter	value	
Glk	$C_{Glk}(0)$	0.05	μM
	V_{glk}	225	$\mu mol \cdot min^{-1} \cdot mg^{-1}$
	K_{Gluc}	0.12	mM
	K_{ATP_1}	0.50	mM
Pgi	$C_{Pgi}(0)$	0.05	μM
	V_{Pgi}	1511	$\mu mol \cdot min^{-1} \cdot mg^{-1}$
	K_{G6P}	3.0	mM
Pfk	$C_{Pfk}(0)$	0.05	μM
	V_{Pfk}	145	$\mu mol \cdot min^{-1} \cdot mg^{-1}$
	K_{F6P}	0.46	mM
	K_{ATP_2}	0.04	mM
N/A	n (forF6P)	1.9	

Molecules traveling through the network are modeled as intermediate products which form a queue in front of a machine. Molecules awaiting conversion by enzymes are called *substrate* corresponding to a queue of products in front of a machine.

While the Gillespie algorithm[15,16] treats an enzyme-catalyzed reaction as a single Markov process with exponentially distributed turnover times, Kou et al.[3] explicitly argue that, even in the simplest case, at least two stochastic processes are involved. Specifically, they remark that "[at] low substrate concentration the binding of the enzyme to the substrate is the rate-limiting step in the reaction, so [the distribution of turnover times] *f(t)* reflects the statistics of this Poissonian step...governed by an exponential distribution. At high substrate concentrations, the dissociation of the enzyme−substrate complex to product is the rate-limiting step...so *f(t)*... is no longer Poissonian."

We will call the two independent stochastic processes that play a role in enzyme kinetics the search process and reconfiguration process. During the search process, substrate and enzyme molecules diffuse randomly until they collide. The search process is characterized by a search time τ_s which depends on the concentrations of substrate and enzyme molecules. In general, the search time τ_s is inversely related to substrate concentration, and this relationship can be modeled in any way, so various hypotheses on the correct spatial physics can be investigated. At a machine level, the search time can be modeled as the setup time of a machine, which most simply depends inversely on the queue length.

The reconfiguration process is modeled by a reconfiguration time, τ_r, which is defined as the time an enzyme needs to handle
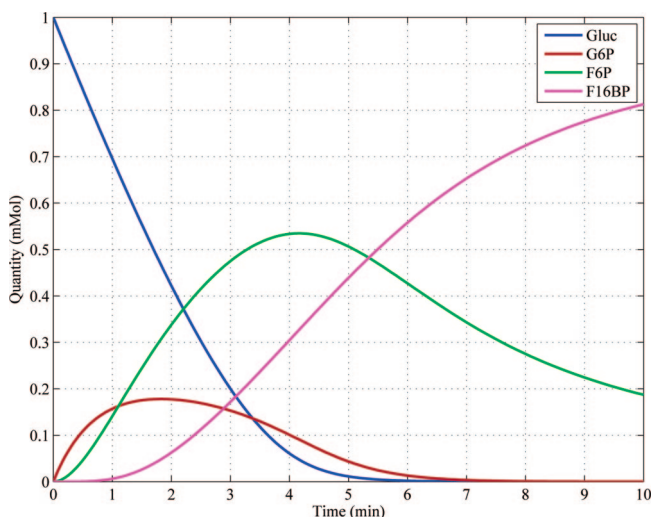
Figure 2. Numerical solution of the ODE model (eqs 1−4). See text for initial conditions.

the substrate and release the product(s). At high substrate concentration (high arrival rate of the product) all enzymes will be saturated (all machines are busy), and hence the stochastic behavior of the reaction no longer depends on substrate concentration but only on the reconfiguration process and the number of enzymes involved. Therefore, assuming a Michaelis−Menten model with a limiting maximum flux of $V_{max}[E]_t$ where V_{max} is the enzyme's "maximum velocity" and $[E]_t$ the total concentration of that enzyme, we can relate the reconfiguration time to measurable quantities by the expression

$$\tau_r = \frac{1}{V_{max}[E]_t V_{E.coli}} \tag{8}$$

where $V_{E.coli}$ is the volume of a single *E. coli*.

At low concentrations of the substrate (low arrival rates) the search process (setup or allocation process) dominates the stochastic behavior of the production step. Note that $1/(\tau_s + \tau_r)$ is the production rate of the combined two processes, i.e., the flux generated by one enzyme. Hence the specific enzyme activities v_{enzyme} of the enzymes Glk, Pgi, and Pfk described in eqs 2, 3, and 4 are related to the search time by the equation

$$\tau_s + \tau_r = \frac{1}{v_{enzyme}[E]_t V_{E.coli}} \tag{9}$$

Solving eqs 8 and 9 for τ_s results in

$$\tau_s = \left(\frac{V_{max}}{v_{enzyme}} - 1 \right) \tau_r \tag{10}$$

Equation 8 allows us to determine the reconfiguration time from measured values of V_{max} (Table 2), while eq 10 gives us the search time as a function of the substrate concentration, V_{max}, and the Michaelis−Menten parameters of this experiment listed in the same table.

In part, because of lack of detailed information about single-molecule reactions and in part to maintain simplicity, we make certain assumptions that cannot necessarily be justified by well-established physical first principles. These are the following:

• All molecules have the same size and one molecule corresponds to 1.0 μM. The time scale is chosen to be in minutes.

• Since search time depends on the length of the queue, wherever queue length changes, search time will be recalculated.

• While enzyme−substrate coupling is modeled via a concentration-dependent search process, ATP, required by Glk and Pfk, is allocated in a different way. Since it is needed in two of the production steps there is a competition for ATP among the machines (enzymes). The substrate−enzyme couple with the shortest search time wins and receives an ATP molecule when available. If ATP is depleted, the enzyme has to wait until a new ATP becomes available.

• Although the simulation language χ allows one to choose samples for the search and reconfiguration times out of many different precoded probability distributions, as well as user-defined distributions, we choose to model the search process via an exponential distribution with a mean determined from the experimental data.

• The reconfiguration time τ_r is assumed to be Γ distributed with a coefficient of variation of 3.0, corresponding to a heavier tail in the distribution. This assumption roughly mimics the

Single-Molecule Enzyme Networks

J. Phys. Chem. B, Vol. 113, No. 16, 2009 **5541**
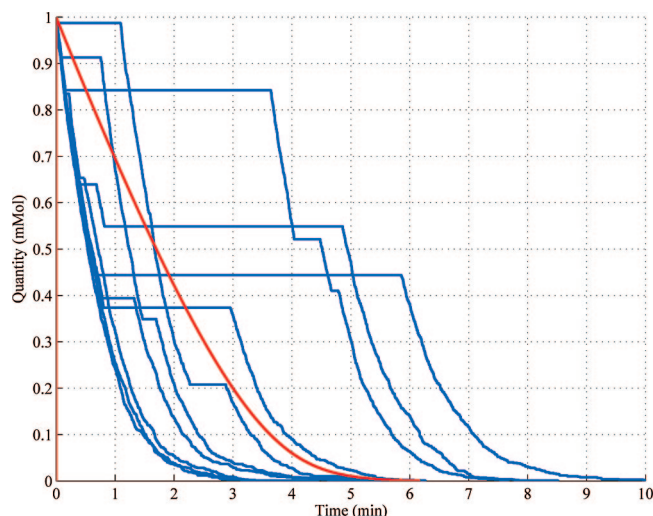
Figure 3. Comparison of 10 stochastic (blue) versus deterministic (red) time series of [Gluc] in transient experiments as described in the text.

distribution of turnover times ($f(t)$) described by ref 3 and references therein.

Note that certain biophysical implausibilities in these somewhat arbitrary assumptions can easily be corrected, with more detailed information or hypotheses about the actual reaction process, without damaging the ease of coding or performance of the simulation.

Despite the oversimplifications we make here, the model still performs well compared to the ODE approach. In particular, to compare the DES described above to the ODE results in section 3.1, we remove the stochastic nature of the DES by fixing all stochastic parameters at their mean values. We set the initial concentrations of glucose and ATP in the DES to 1000 and 2000, respectively, to preserve the stoichiometry assumed in the ODE. The resulting time evolutions of the DES and ODE models are essentially identical (Figure 2); in particular, the DES simulations on the time scale of Figure 2 lie exactly on the continuous curves generated from the ODEs. The only exception concerns the initialization of the DES system which is done via an arrival process leading to glucose buildup on a very fast time scale.

4. Results

4.1. Stochastic Transients. Now that we have shown the equivalence of the deterministic DES and ODE models, we repeat the transient glucose processing experiment with stochastic parameters. Figure 3 shows 10 sample paths for the glucose concentration dynamics together with the deterministic time series. Note that since there is no influx of glucose from an outside source, its concentration decays over time. The wide plateaus in the stochastic simulations are caused by extremely long search and/or reconfiguration times for Glk. The large variation in sample paths indicates that stochastic behavior plays a significant role.

4.2. Stochastic Analysis of a Continuous Flow System. The transient simulations stop after all glucose has been processed or when ATP has been depleted. Even with an exogenous supply of glucose, the system will eventually stop since there is no ATP regeneration mechanism in the model. Therefore, to study a continuous flow system we introduce a simple representation of downstream regeneration of ATP from ADP, modeled as a conversion process *CV* (see Figure 4). Conversion of one

molecule of ADP into one molecule of ATP is assumed to be exponentially distributed with mean process time t_c. This conversion process can represent ATP regeneration from downstream stages of the EMP pathway or oxidative phosphorylation, but here we ignore adenylate kinase and other more subtle reactions involving adenylate. In addition, we model glucose influx with a mean interarrival time of t_a, also assumed to be exponentially distributed.

Note that here we have introduced another simplification. The early steps of the EMP pathway, which hydrolyze ATP, are required before subsequent steps in the pathway, downstream of our model, can regenerate ATP. We ignore regulation of ATP regeneration via feedbacks on EMP from the downstream ATP-generating processes. For example, we ignore the well-known inhibition of Pfk by ATP. In general, both positive and negative feedbacks depend locally on the availability of an enzyme's substrates or products, or globally on up- or downstream intermediate products of the pathway. The conversion process we introduce here largely ignores the details of such feedback, but they can be incorporated into more realistic simulations by describing the parameters of the controlling probability distributions as functions of the concentration of relevant intermediates. In any case, incorporating general feedback presents no modeling or simulation issues beyond the addition of the conversion process shown in Figure 4.

We simulate the system for 2×10^3 min with 100 molecules ATP and ADP, respectively, and monitor average substrate concentrations, outflux, and cycle time.

Outflux is determined by the number of F1,6bP molecules produced per time interval averaged over 9 equivalent time intervals and 20 stochastic sample runs. We introduce the dimensionless variable Δ, which represents a scaling of the average outflux relative to the mean influx $1/t_a$, as follows:

$$\Delta = \frac{\text{outflux}}{1/t_a} = \text{outflux} \cdot t_a \qquad (11)$$

For $\Delta = 1$ the system successfully processes all incoming glucose molecules, and all buffers stay finite. Such a system is therefore called *stable*. For $\Delta < 1$ the arrival rate exceeds the production capacity of the system, and at least one buffer blows up. The amount of time a molecule spends in transit from glucose to F1,6bP is called the *cycle time*. A natural, nondimensional measure of cycle time is the quotient of the average "actual" cycle time and the sum of all mean reconfiguration times, denoted by the dimensionless parameter Φ. Intuitively, Φ measures the amount of waiting in the system due to both searches for enzyme substrate coupling and the stochasticity of the reconfiguration process.

Table 3 compares the stochastic and deterministic results for three different sets of mean interarrival times (t_a) and mean conversion times (t_c). Note that, although these data do not include variances, they do describe the influence of stochasticity on mean cycle time and substrate distribution.

In the first simulation, with $t_a = 0.009$ and $t_c = 0.001$, ATP is rapidly regenerated. Therefore, stochastic and deterministic time evolutions for adenylate are essentially identical. However, in the stochastic results we see a large increase in concentrations of Gluc and F6P because Glk and Pfk create production bottlenecks (see Table 2). Cycle time also increases significantly due to stochasticity, outflux does not change, consistent with Little's law.[28]

As t_c increases, ATP is regenerated more slowly, starving Glk and Pfk of ATP. These results suggest that, on average,
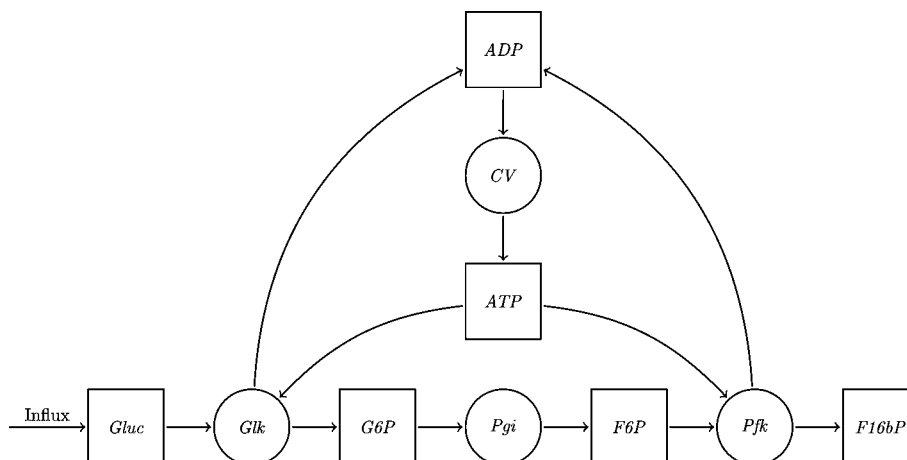
Figure 4. Discrete event model representation of glucose phosphorylation in EMP pathway with ATP regeneration.

TABLE 3: Comparison between Stochastic and Deterministic Discrete Events Simulations of EMP Glucose Phosphorylation for Various Values of Glucose Interarrival Time t_a and Mean Times between ATP Regeneration t_c

t_a	t_c	Gluc	G6P	F6P	ATP	ADP	Φ	Δ
0.009	0.001							
stochastic		1816.6	59.4	1041.3	198.2	1.5	4141	1.00
deterministic		872.0	71.9	486.1	199.1	1.0	2005	1.00
difference (%)		−108.3	17.4	−114.2	0.5	−50.2	−106.5	0.0
0.016	0.007							
stochastic		705.0	29.7	510.9	129.3	69.6	3165	1.00
deterministic		117.8	39.7	291.8	198.8	0.8	1120	1.00
difference (%)		−498.6	25.1	−75.1	35.0	−8270	−182.5	0.2
0.016	0.008							
stochastic		4396.8	29.0	633.1	90.5	108.2	12742	0.93
deterministic		117.8	39.7	291.8	198.6	0.8	1120	1.00
difference (%)		−3633	27.1	−117.0	54.4	−13862	−1038	6.7

when a significant amount of ADP is generated, the cycle time increases, and Glk creates a bottleneck leading to a blowup in its glucose buffer while depleting all downstream buffers. If t_a is held fixed at 0.016, then the stability boundary between regions in which ATP is depleted and regions in which it is not lies somewhere between $t_c = 0.007$ and $t_c = 0.008$. This observation suggests that the cell has a distinctly different distributions of intermediate products, and therefore suffers a sudden shift in metabolic states, as its energy charge (a weighted average of "high energy" phosphate per adenylate molecule) decreases.
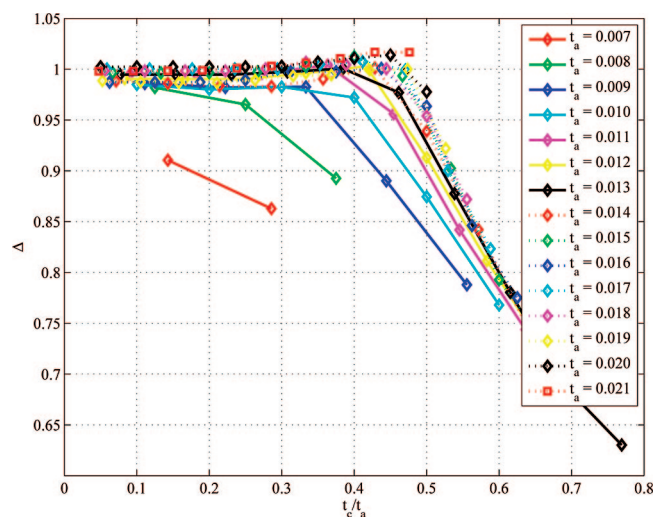
4.3. Stability Boundary. To analyze the location of the stability boundary as a function of energy charge and ATP demand, we performed a series of stochastic experiments by varying mean time between ATP regeneration events (phosphorylation of ADP to ATP), t_c, and mean interarrival times for glucose, t_a. The simulation times were fixed at 2×10^3 min. We conducted one simulation per parameter combination. Figure 5 shows Δ as function of t_c and t_a. In this figure, one can distinguish the following three regimes:

Overloaded system due to slow regeneration of ATP. For $t_c/t_a \gtrsim 0.4$ and $t_a \geq 0.011$, Δ decreases below 1. In this regime, slow ATP regeneration starves the ATP-dependent enzymes Glk and Pfk of available free energy, resulting in a buildup of Gluc and depletion of F6P. As a result, total outflux, and therefore Δ, both decrease. When $t_a \geq 0.011$, Δ is determined entirely by the ratio t_c/t_a, as can be seen in Figure 5, since all curves in which $t_a \geq 0.011$ are essentially identical.

Stable system. The system is stable for $t_c/t_a \lesssim 0.4$ and $t_a \geq 0.011$. Importantly, since phosphorylation of a single glucose molecule requires two ATP molecules, the deterministic stability

boundary is $t_c/t_a = 0.5$. But, due to the system's stochastic behavior, ATP is more often unavailable than deterministic theory would allow, and hence the region of stability is smaller. Once again, this points to the importance of studying the effects of stochasticity on metabolic dynamics.

Increasing the glucose influx beyond enzyme capacity. Most graphs of scaled outflux, Δ, for different values of t_a follow one identical path, starting at $\Delta \approx 1$ and dropping below unity when $t_c/t_a > 0.4$. But, for $t_a \leq 0.010$ the drop in scaled outflux Δ occurs earlier. In this region it becomes harder to stabilize

Figure 5. "Blow-up" parameter Δ as a function of t_c/t_a for different values of t_a.

Single-Molecule Enzyme Networks

J. Phys. Chem. B, Vol. 113, No. 16, 2009 **5543**

the system by shortening the conversion time t_c. In fact, when $t_a = 0.007$, glucose is arriving so rapidly that the system cannot be stabilized by any reasonable increase in ATP regeneration rate. In this case, one can see from Table 2 that Pfk has the slowest processing rate and hence is the *bottleneck* of the network. As a consequence the concentration of F6P will increase without bound. Note that deterministically, the resulting stability boundary is at $t_a = 0.004$. Again, stochasticity reduces the capacity of the enzymes to process glucose.

5. Conclusion

We have shown that simulation techniques routinely applied in production systems and supply chain modeling can be useful to simulate complex biochemical pathways where one can follow the time evolution of single molecules through a network. We model a single-molecule enzymatic reaction via two stochastic processes: a search process whose time to completion is a stochastic variable modeled by an exponential distribution that depends on the number of substrate molecules, and a reconfiguration process whose time to completion is a stochastic process that we model via a Γ-distribution. Simulation tools from discrete event simulations monitor all reactions in the network as concurrent processes. This technique allows simulations of arbitrarily complex time dependent processes, far exceeding the applicability of Markov approximation-based analytic solutions for stationary processes. The associated process algebra allows verification of the simulation tool and invests the simulations with a high degree of mathematical rigor.

As an illustration of the technique we have generated a stochastic simulation for the glucose phosphorylation steps of the EMP pathway. Our goal is not to study the EMP pathway per se, but rather to demonstrate the technique. Nevertheless, our simulation of even the simplified system yields biological insight. For instance, the simulations predict that the distribution of intermediate metabolites, and therefore the metabolic state of the cell, has a nearly stepwise dependence on cell energy charge, which is typically defined as

$$\frac{[ATP] + (1/2)[ADP]}{[ATP] + [ADP] + [AMP]}$$

although here we ignore AMP. In particular, we find a "capacity boundary" such that on one side of the boundary, the system can successfully phosphorylate all incoming glucose, whereas on the other side it "blows up" as one of the ATP-dependent enzymes becomes a bottleneck. In addition, we quantify the effects of stochastic behavior on fluxes, times to process a molecule (cycle time) or the stability of the system, and show that stochasticity can significantly alter conclusions based on deterministic theory.

A similar conclusion is reached by Kühl and Jobmann[18] who studied the influence of nonexponential (i.e., non-Markovian) distributions on the shape of dose—response curves for ligand—receptor interactions. In future work we plan to extend their approach to analyze in detail second and higher moments of reaction rate distributions for biochemical networks, including local and nonlocal feedbacks.

Tools from the paradigm of production systems can easily extend to more complicated biochemical networks than the one we study here. For instance, [3] discuss the influence of dynamic disorder on Michaelis—Menten kinetics. Dynamic disorder assumes there are multiple channels for the same

enzymatic reaction that work at slightly different speeds. For a production system, dynamic disorder corresponds to parallel machines all of which have their own slightly different cycle time distribution. It is natural for production system modeling to choose different cycle time distributions for those different parts of the whole process. Typical simulation tools (e.g., χ) support any kind of such distributions and can therefore immediately be applied to the study of dynamic disorder in biochemical networks.

It is worth noting that our simulation approach motivated by the simulation approaches used in production systems is still a discrete event simulation. That means in particular that multiscale issues leading to very long simulation times which generated τ-leaping, hybrid, and other approximate simulation models are not resolved. In the context of production systems, those issues lead to the development of aggregate models using concepts like effective process times[29] as well as hybrid simulation models based on hybrid process algebras.[24] Exploring the transfer of such aggregate models to biochemical networks is another direction for future work.

There are a number of metabolic and genetic networks in which stochasticity may play a significant or even dominant role, and to which DEMs derived from production systems can be applied. For example, we have begun simulating the Ras-Raf-MEK-ERK signal transduction pathway[30,31] and have plans to begin modeling generalized genetic regulation networks as in ref 12 and studies of starvation responses.[32] In each of these examples the system size is best described at the mesoscale, where limited numbers of molecules, and hence their fluctuations, should play a major role in the pathway's functionality.

References and Notes

(1) Rigler, R. *J. Biotechnol.* **1995**, *41* (2−3), 177.
(2) Xie, S.; Trautman, J. K. *Annu. Rev. Phys. Chem.* **1998**, *49*, 441.
(3) Kou, S. C.; Cherayil, B. J.; Min, W.; English, B. P.; Xie, X. S. *J. Phys. Chem. B* **2005**, *109*, 19068.
(4) Qian, H; Elson, E. L. *Biophys. Chem.* **2002**, *101*, 565.
(5) Meinhold, L.; Schimansky-Geier, L. *Phys. Rev. E* **2002**, *66*, 050901(R) .
(6) Falcke, M. *Biophys. J.* **2003**, *84*, 42.
(7) Hou, Z.; Xin, H. *J. Chem. Phys.* **2003**, *119*, 11508.
(8) Falcke, M. *Adv. Phys.* **2004**, *53*, 255.
(9) McKane, A. J; Newman, T. J. *Phys. Rev. E* **2004**, *70*, 041902.
(10) Kummer, U.; Krajnc, B.; Pahle, J.; Green, A. K.; Dixon, C. J.; Marhl, M. *Biophys. J.* **2005**, *89*, 1603.
(11) Li, H.; Hou, Z.; Xin, H. *Phys. Rev. E* **2005**, *71*, 061916.
(12) McKane, A. J.; Nagy, J. D.; Newman, T. J.; Stephanini, M. O. *J. Stat. Phys.* **2007**, *128*, 165.
(13) McKane, A. J.; Newman, T. J. *Phys. Rev. Lett.* **2005**, *94*, 218102.
(14) Rao, C. V.; Wolf, D. M.; Arkin, A. P. *Nature* **2002**, *420*, 231.
(15) Gillespie, D. T. *J. Comput. Phys.* **1976**, *22*, 403.
(16) Gillespie, D. T. *J. Phys. Chem.* **1977**, *81*, 2340.
(17) Li, H.; Cao, Y.; Petzold, L. R.; Gillespie, D. T. *Biotechnol. Prog.* **2008**, *24* (1), 56–61.
(18) Kühl, P. W.; Jobmann, M. *J. Recept. Signal Transduction* **2006**, *26*, 1–34.
(19) Wilkinson, D. J. *Stochastic Modelling for Systems Biology*; Chapman: London, 2006.
(20) Van Kampen, N. G. *Stochastic Processes in Physics and Chemistry*; Elsevier: Amsterdam, 1981.
(21) Bernstein, D. *Phys. Rev. E* **2005**, *71*, 041103.
(22) Law, A. M.; Kelton, W. D. *Simulation modeling and analysis*, 3rd ed.; McGraw-Hill: Boston, MA, 2000.

(23) Baeten, J. C. M.; Middelburg, C. A. *Process algebra with timing*; Springer Verlag: New York, 2002.

(24) van Beek, D. A.; Man, K. L.; Reniers, M. A.; Rooda, J. E.; Schiffelers, R. R. H. *JLAP* **2006**, *68*, 129.

(25) Hofkamp, A. T.; Rooda, J. E. Chi 1.0 reference manual, Systems Engineering Report, Eindhoven (August 2007) http://se.wtb.tue.nl/sereports.

(26) Schaechter, M. *Encyclopedia of Microbiology*; Academic Press: Boston, MA, 1994; Vol 2, pp 115−124.

(27) Ishii, N.; Suga, Y.; Hagiya, A.; Watanabe, H.; Mori, H.; Yoshio, M.; Tomita, M. *FEBS Lett.* **2007**, *581*, 413.

(28) Little, J. D. C. *Oper. Res.* **1961**, *9*, 383.

(29) Jacobs, J. H.; Etman, L. F. P.; van Campen, E. J. J.; Rooda, J. E. *IEEE Trans. Semicond. Manuf.* **2003**, *3* (16), 511–520.

(30) Hornberg, J. J.; Binder, B.; Bruggeman, F. J.; Schoeberl, B.; Heinrich, R.; Westerhoff, H. V. *Oncogene* **2005**, *24*, 5533.

(31) Roberts, P. J.; Der, C. J. *Oncogene* **2007**, *26*, 3291.

(32) Ropers, D.; de Jong, H.; Page, M.; Schneider, D.; Geiselmann, J. *BioSystems* **2006**, *84*, 124.