A hybrid CME-ODE method for efficient simulation of the galactose switch in yeast

David M. Bianchii, Joseph R. Petersonj, Tyler M. Earnestt, Michael J. Hallocc, Zaida Luthe-Schultenl

Abstract: It is well known that stochasticity in gene expression is an important source of noise that can have profound effects on the fate of a living cell. The most dramatic cases occur with genetic switches that can drive a population into two or more distinct phenotypes in a non-deterministic manner. In the galactose genetic switch in yeast, the unbinding of a transcription repressor is induced by high concentrations of sugar particles activating gene expression of sugar transporters. This response, which is amplified through feedback loops within the cell, results in high propensity for all reactions involving interactions with the metabolite. The reactions for gene expression, feedback loops, and transport are typically described in terms of Chemical Master Equations (CME). Sampling the CME using the Stochastic Simulation Algorithm (SSA) results in large computational costs as each event is evaluated explicitly. To improve the computational efficiency of cell simulations involving such genetic switches and other high particle number systems, we have implemented a hybrid stochastic-deterministic (CME–ODE) method into our GPU–based Lattice Microbes (LM) software suite and its Python Interface pyLM which are both publicly available [1–4]. LM and pyLM provide a convenient way to simulate complex cellular systems and interface with high-performance RDME/CME/ODE solvers. As a test of our implementation, we apply the hybrid CME/ODE method to the galactose switch in Saccharomyces cerevisiae and investigate the appropriate communication time between the ODE and CME solvers at various concentrations. Our robust hybrid implementation is 10–50X faster, while yielding protein distribution patterns that are similar to the exact SSA CME.

1 Introduction

Many processes within living cells, especially gene expression, have a high degree of randomness and a high degree of non-determinism that leads to stochastic effects, such as heterogeneity in a population of cells. The Chemical Master Equation (CME) is a useful formalism for describing the dynamics of stochastic events in biological systems. It describes a chemical process as a continuous time Markov chain on a state space comprising particle numbers of each chemical species; thus, capturing the discreteness of particles and the random nature of individual chemical reactions. The widely used Stochastic Simulation Algorithm (SSA) of Gillespie [5, 6] provides an effective method for obtaining unbiased realizations of these Markov processes. However, the SSA is limited by the fact that reaction events are accounted for explicitly. Systems with high particle counts—those containing metabolites in millimolar concentrations—or those with large reaction rate constants, have a high propensity (probability per unit time) for a reaction event to occur; thus, they evolve on fast time scales and incur large computational cost (e.g. time to solution).

To overcome the computational expense of solving the SSA for high concentration systems researchers have devised hybrid approaches that decrease time-to-solution for stochastic simulations while faithfully capturing the results of stochasticity in important chemical species (e.g. transcription factors). For a brief review of methods that improve computational efficiency by reformulating the original SSA scheme in a more economical fashion, see Jahnke and Kreitz [7]. Notably, Cao et al. describe a method by which the chemical system of interest is separated into a set of reactions with fast rates to be simulated deterministically and a set of slow reactions to be simulated stochastically [8]. Alfonsi et al. present a hybrid model in which a CME Markov jump process describing the dynamics of the species with low particle number is coupled to ODEs (Ordinary Differential Equations) representing the highly abundant species [9]. Jahnke and Kriem validated this technique through a rigorous error analysis of the CME-ODE partitioning which was compared to a CME treatment for a small reaction system [7], Pájaro and Alonso examined the applicability of approximate methods to modeling genetic circuits [10], and Smith et al. showed the applicability of hybrid methods to metabolic networks [11]. Alternative frameworks exist, for example coupling CME with the chemical Langevin equation [12]. Algorithms to handle stochastic reaction-diffusion processes typically partition the system into small spatial subvolumes and use the SSA to describe the reaction events within them [1, 2, 13]. The hybrid CME-ODE method discussed in this work will also accelerate spatially resolved simulations of processes described by reaction diffusion master equations (RDME) over such subvolumes, where reactions within each compartment are treated by the CME.

A challenging and typical scenario arises when species participating in slow reactions are also changed by firings of one or more fast reactions, making the dynamics of the slow reactions dependent on the fast reactions. While it is tempting to assume a partial thermodynamic limit for the fast reactions involving a large number of species and simply rescale the rate constants so the entire system can be treated stochastically, this assumption cannot be made for genetic switches involving nutrients. During the early phase of sugar/inducer/metabolite uptake, the system evolves quickly and errors in the approximation can compound leading to incorrect switching dynamics and even incorrect switched steady-states. The study of nutrient-induced metabolic switches is of particular importance in understanding survival fitness. As a cell adapts to a change
in the composition its environment, genes that will help in the processing and efficient usage of the new metabolite are upregulated, and genes that are no longer needed are downregulated, in an attempt to optimize its fitness.

2 Test System

We simulate the well-studied galactose switch in *Saccharomyces cerevisiae* [14–16]. This system, composed of 37 species and 75 reactions, is summarized in Fig. 1 and provided in Appendix A. The switch has four feedback loops that respond to the presence of the sugar galactose. The transcription factor G80, in dimer form (G80d), binds to the promoters of the genes for several of the proteins (G1, G2, G3, and G80) involved in the galactose switch process, and represses their expression. When galactose binds to G3 it creates an activated complex (G3i) that can bind to G80Cd (G80d in the cytoplasm). The G3i sequesters G80 in the cytoplasm causing the genes to be in an unpressed, active state. The proteins G2, which transports galactose into the cell, and G1, which metabolizes galactose within the cell, also play key roles in the cell’s response to changes in the galactose concentration in its environment.

The positive (G1, G2, G3) and negative (G80) feedback loops of the system work in the following manner. When exposed to a galactose rich environment the cell begins to take up sugar which subsequently sequesters the transcription factor G80 in the cytoplasm, allowing the expression of the genes for G1, G2, G3, and G80. The associated mRNAs diffuse out of the nucleus into the cytoplasm where they can be translated into their protein products. As G2 transporter proteins accumulate, they subsequently motivate a flood of more galactose into the cell. As a counter-balance, more G1 proteins are also produced to metabolize the galactose. Increased G3 counts results in the production of more G3i which can sequester the newly formed G80 produced via the lone negative feedback loop of the system. A transcriptional activator G4, is constitutively expressed and dimerizes before binding to the each of the genes (G1, G2, G3, G80) in the system to promote transcription. The G4 dimer (G4d) is inactive when G80 is bound at the promoter. To measure the activity of G4d in transcription the expression of a reporter protein under the control of the G1 promoter is quantified. The reporter is a yellow fluorescent protein (YFP) added to each yeast cell whose expression level can be measured experimentally by fluorescence microscopy. The amount of reporter present is used to quantify the overall switching behavior of the system.

Galactose can exist in up to millimolar concentrations in a yeast cell, therefore, to test our method we examine scenarios ranging from 0.055 to 2.0 mM galactose concentrations (approximately 1–50 million particles) that mimic the experiments performed by Oude-naarden et al. [14] and Ramsey et al. [15]. We model a yeast cell with a volume of 35.7 fL and set the initial intracellular galactose concentration to zero. These initial conditions allow us to study the dynamics of the genetic switch, while necessitating the use of a deterministic–stochastic hybrid method to track the millions of sugar particles that rapidly rush into the cell to faithfully capture stochastic gene expression. Simulations were run with a constant concentration of extracellular galactose over the course of a simulation.

The system reactions were treated either by the CME or the ODE (Fig. 1). Transcription, translation, and transcription factor–promoter reactions are solved by the CME simulator. Species involved in these reactions were present in (relatively) low copy number throughout the course of a simulation and exhibit high variability, motivating stochastic treatment. Reactions with high propensities, due to highly abundant reactants or high reaction rate constants, such as the transport, binding and metabolism of galactose, are evolved deterministically using an ODE solver. The choice of this partitioning is crucial in the effectiveness of the hybrid algorithm in capturing the stochastic behavior seen in the pure CME. In fact, when we began this investigation we had defined the reaction of G3 with the extracellular galactose as an ODE reaction and had not considered the relatively low amount of G3 that is biologically present. This assignment led to increased errors in effectively capturing CME distributions especially when the exact SSA-ODE method (described in Section 5) where CME–ODE communication occurs at every CME reaction event was considered. This is a problem because hybrid simulations will converge to this exact result as smaller timesteps are used. Since the propensities associated with this reaction increased in the order of those for the transport reactions it should be defined as a CME reaction, without a significant deleterious effect of run–times. A potential flaw that could arise from the above partitioning is that species involved in fast reactions are described by the ODE solver even when they are present in very small numbers (at early simulation times). At early times these species may not have smooth trajectories and can be poorly approximated by differential equations. However, we did not observe these errors playing a noticeable part in simulations of the galactose switch system, as the trajectories of these species seemed to match those obtained from the pure CME, although the hybrid method at times reached steady-state values slightly sooner than the pure CME (Fig. 3c).

We will now describe the implementation of the hybrid CME-ODE algorithm, its fidelity to pure SSA CME simulations (which we accept as a ground truth) and the relative increase in computational efficiency. We study accuracy relative to the pure CME and computational speedup as a function of sugar concentrations seen in experimental studies of the galactose switch system and provide insight into the differences that are observed. We also examine the effect of the choice of interval at which information is exchanged between the CME and ODE solvers (the communication time τ), comparing results and run–times to the the exact SSA-ODE algorithm where communication occurs between the stochastic and deterministic simulation regimes at every CME reaction event. This hybrid method utilizes the Lattice Microbes (LM) software suite [1], and is implemented using its Python interface pyLM [4]. This study showcases the hybrid implementation as a new feature of LM that allows users to easily define and simulate complex biochemical systems and to achieve computational performance that was previously unavailable in LM for systems with chemical species present in millimolar concentrations.

3 Hybrid CME-ODE Algorithm Description

The hybrid algorithm is implemented such that the states treated by the CME and ODE are distinct and evolve independently over certain time intervals. At intervals set by the user (e.g. the communication timestep τ, shown by thick vertical lines) state data is transferred between treatments. Fig. 2 illustrates the manner in which the hybrid algorithm connects the two descriptions by communicating species counts, which can be poorly approximated by differential equations. However, we did not observe these errors playing a noticeable part in simulations of the galactose switch system, as the trajectories of these species seemed to match those obtained from the pure CME, although the hybrid method at times reached steady-state values slightly sooner than the pure CME (Fig. 3c).

The communication timestep is the key parameter of the hybrid algorithm. At the conclusion of each communication timestep, the state in the deterministic regime is updated with the species counts obtained from the end of the previous CME timestep. The ODE
solver is simulated forward by \( \tau \) to evolve the high particle number species in the deterministic regime. The LSODA solver is used to allow for adaptive timesteps. Then, the rates of reactions in the stochastic regime that involve species in the deterministic regimes are updated accordingly. In effect the end of the previous timestep and the beginning of the current timestep happen simultaneously since no simulation time elapses between these points in the algorithm. The process then repeats itself until the total user-defined simulation time is completed.

4 Model Modifications

The model of Ramsey et al. [15] adopted for this work contained several Michaelis-Menten and Hill reactions. To make the model amenable to simulation with the SSA, which is only valid for elementary reactions [5, 6], these reactions were decomposed into first and second order reactions. Transport of galactose was originally modeled as a reversible Michaelis-Menten reaction, which we decomposed into six bimolecular reactions describing the underlying Michaelis-Menten reaction explicitly (see Appendix Section 9.3.4). Gene repression in the Ramsey et al. model was modeled using Hill functions [15]. These reactions were simplified into four binding or unbinding reactions for each gene. These reactions consist of: 1) binding of G4 dimer to the gene, 2) unbinding of G4 dimer from the gene, 3) binding of the G80 dimer to the G4 dimer when it is bound to the gene, and 4) unbinding of the G80 dimer:G4 dimer from the gene (see Appendix Section 9.3.2). These modifications allow the model (hereafter referred to as the “reduced model”) to be directly implemented in a spatially-resolved stochastic reaction-diffusion framework (RDME) where the assumptions underlying Hill and Michaelis-Menten reactions do not always hold [17].

Rate constants for the newly simplified reactions were fit to recapitulate the cooperative behavior of the more complex rate equations. ODEs of the original and reduced model were simulated for 750 minutes and used to fit the rate constants. Briefly, during the fitting procedure the sum of squared differences between concentrations of each species in the original and reduced solutions was minimized. The reduced model generally reproduced the results of the original model less than 5% error.

5 Agreement with Pure SSA CME

The communication times between the stochastic and deterministic descriptions as well as the time steps for each solver must be evaluated to understand under which conditions the hybrid method is appropriate. In this section we provide insight into the effect the choice of \( \tau \) has on the behavior of different chemical species at varying concentrations of extracellular galactose. The hybrid CME-ODE simulation algorithm effectively captures the stochastic dynamics of the genetic switch process at galactose concentrations ranging from micromolar to millimolar. At a small enough communication timestep, both protein distributions and average traces of key species approach pure SSA CME simulation results, while allowing for sufficient performance enhancement to make the method worthwhile. It is important to note however, that even in the limit in which the communication timestep approaches zero the results from our hybrid implementation should not converge to the pure CME results, but rather to an exact hybrid SSA-ODE model, where every time a reaction fires in the CME the ODE solver is called with updated species counts due to the CME reaction.

Significant error arises in the protein distributions of the unbound G2 transporter (hereafter referred to as G2) when large communication timesteps are used (Fig. 3a). As an example, consider a simulation with an extracellular galactose concentration of 0.055 mM where a 5 minute timestep was chosen. This results in an overestimation of the mean and variability in the G2 count. As G2 is affected by reactions in both the CME (gene expression) and ODE (transport), and plays a key role in bringing galactose into the cell, error in the average or the noise in this species could give rise to differences in the switching dynamics or steady-state copy number distributions. This underscores the importance of choosing an appropriate communication timestep between the simulation regimes. As \( \tau \) is decreased from 5 minutes to 1 second the protein distribution begins to closely match the pure SSA CME result, with runtime only increasing from approximately 25 minutes to 45 minutes in the worst case.

The optimal communication timestep is dependent on the concentration of the external galactose. At an extracellular galactose concentration of 0.055 mM a 1 minute timestep seems to closely approach the CME distribution results (Fig. 3a). However, as demonstrated by the G2 average at 2.0 mM galactose (Fig. 3b) the 1 minute
timestep is on the order of a few 1000 proteins away from the pure CME result as opposed to 100s of proteins when using a 1s timestep at an increase of simulation time from 45 minutes to approximately 1 hour. We have observed the trend that a smaller communication timestep is often needed to capture stochastic behavior at higher inducer concentrations. The exact SSA-ODE trace (small dashes in Fig. 3b), which was calculated using 250 replicates, can be used to determine whether moving to a smaller communication timestep will provide an increase in accuracy since smaller timesteps will converge to this result.

The dynamics of a chemical specie can also be observed by witnessing the times taken to reach steady-state values. At an extracellular galactose concentration of 2 mM the CME-ODE hybrid with \( \tau = 1s \) takes similar times to reach steady-state concentration as the pure CME (Fig. 3c). For G2 the mean and median times to reach 80 percent of the average steady-state value were less 10 minutes faster for the CME–ODE than the CME. While the hybrid method did seem to evolve slightly faster than the CME for some of the species with reactions handled by the ODE solver the difference is relatively small (generally on the order of 5 - 10 minutes faster and in the worst case approximately 20 minutes faster).

As a quantitative measure of disagreement between the hybrid method and the pure CME, the Kolmogorov–Smirnov Statistic (KS Test) for protein distributions from each method was computed. The KS Test calculates the maximum difference in cumulative probability between empirical cumulative distribution functions of two samples (i.e. Pure CME and Hybrid CME-ODE protein distributions). Therefore, identical distributions give a KS Statistic of 0.00 and completely non-intersecting distributions would give a value of 1.00. The formula for KS Statistic is:

\[
KS = \sup_x |F(x)_{CME} - F(x)_{Hybrid}| \tag{1}
\]

where \( \sup_x \) is the supremum of the set of distances, \( F(x)_{CME} \) is the pure CME protein empirical cumulative distribution function and \( F(x)_{Hybrid} \) is the Hybrid CME-ODE protein empirical cumulative distribution function.

For species with reactions in the deterministic regime (i.e G2) a decrease in communication timestep coincides with a decrease in KS Statistic (increase in agreement between protein distributions) although the observed decrease is larger at the higher external galactose concentration of 2 mM as opposed to 0.055 mM (Fig. 3d). The p–values associated with the KS Test also decrease from \( 5 \times 10^{-7} \) to \( 2 \times 10^{-3} \) as timestep is decreased from 1 minute to 1s, showing greater agreement at smaller timesteps. However, for species that participate only in reactions in the stochastic regime such as the reporter protein, KS Statistic is relatively constant with respect to communication timestep.

Table 1 The hybrid algorithm using a 10s and 1s communication interval can give 10-50X speedup respectively versus a pure CME SSA implementation. The time given is the wall-time required to simulate 1000 replicates (250 for Exact SSA-ODE) of the system using 1 node per replicate. Simulations were performed on a Cray XE machine (NCSA Blue Waters) containing AMD 6276 "Interlagos" processors.

<table>
<thead>
<tr>
<th>Model</th>
<th>[Galactose] (mM)</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>CME (hrs)</td>
<td>2.1</td>
<td>47.4</td>
</tr>
<tr>
<td>Exact SSA-ODE</td>
<td>4.7(^a)</td>
<td>47.9</td>
</tr>
<tr>
<td>(0.45)(^b)</td>
<td>(0.99)(^b)</td>
<td></td>
</tr>
<tr>
<td>Hybrid ((\tau=10s))</td>
<td>0.4 (5.2)</td>
<td>1.1 (43.1)</td>
</tr>
<tr>
<td>Hybrid ((\tau=1s))</td>
<td>0.8 (2.6)</td>
<td>1.8 (26.3)</td>
</tr>
</tbody>
</table>

\(^a\) Times are presented in the number of hours required to simulate 750 minutes of cell growth.
\(^b\) Values in parenthesis indicate the speedup relative to pure CME.

6 Computational Performance

Having established the fidelity of the hybrid method we now demonstrate the dramatic increase in simulation efficiency it provides. Wall-time required to simulate 750 minutes of the galactose switch using the hybrid method, exact SSA–ODE and the pure SSA CME are shown in Table 1 along with the relative speedup.

While the pure SSA CME Gillespie Direct Method takes approximately 2 days to simulate a cell introduced to 2 mM external galactose, the hybrid method using \( \tau=1s \) runs in less than 2 hours. Even at the lowest concentration, 0.055 mM external galactose, simulations are executed in 0.75 minutes instead of the 6-80 minutes. This speedup provides researchers with much higher throughput and can assist informing the design of more complex investigations.

The differences in the simulation times achieved when using a 1 minute communication timestep versus a 1 second timestep at 2 mM external galactose demonstrates that at times the user must make a choice between accuracy (see Fig. 3a,d) and simulation speedup. While at the lower galactose concentration a 1 minute communication interval appears sufficient, the G2 traces and KS Statistic at 2 mM galactose (Fig. 3b,d) show that a 1 second timestep is more appropriate to capture the dynamics of the transporter specie. For those considering even coarser timesteps, the run-time for a simulation using a 5 minute communication timestep is approximately 50 minutes compared to approximately 110 minutes for a 1 second timestep at 2 mM external galactose. By choosing a larger timestep the user may lose accuracy in describing the stochastic behaviors in the cell while not gaining a worthwhile decrease in simulation time.

While the exact SSA-ODE method is an effective tool to determine the accuracy of hybrid methods and the appropriate communication timestep to use, it is much less computationally efficient than comparable fixed timestep runs taking approximately 2 days to run at 2 mM external galactose compared to less than 2 hours for \( \tau = 1s \). At the lower concentration of 0.055 mM galactose the exact method is still 5–10 times slower than when using 10-1 s timesteps. The increase in simulation time for the exact method is due to the fact that makes orders of magnitude more computationally costly calls to Python to run the ODE Solver than what is made with a fixed timestep. The exact method calls the ODE Solver at every CME reaction and we have observed the time between CME reactions to be on the order of micro to nano seconds at 2 mM external galactose (much smaller than a typical \( \tau = 1s \)).

7 Conclusion

The hybrid CME-ODE algorithm implementation described in this study, now compatible with Lattice Microbes/pyLM, provides an effective method for the simulation of a genetic switch system containing 37 species, 75 reactions, 4 feedback loops and millions of metabolite particles within a eukaryotic cell. The 10–50 fold computational performance increase relative to a pure SSA CME simulation for sugar concentrations ranging from micromolar to millimolar makes this simulation method an intriguing option for researchers in the field of computational biology. However, the user must determine a suitable communication timestep between the stochastic and deterministic regimes to ensure that data is passed with enough frequency to maintain the protein distributions and stochastic effects that are observed when using a pure SSA CME implementation. The results gained from these efficient hybrid CME-ODE simulations can be used to inform simulation setup conditions (communication timestep etc.) for hybrid simulations of much more computationally expensive, spatially resolved whole cell Reaction–Diffusion Master Equation (RDME) studies. Hybrid simulations can utilize LM features developed for RDME simulations, such as multiple–GPU computation [2] and optimized propensity calculation [18], without any further work on the user’s part. LM provides the fastest method for RDME simulations to date and with the addition of this hybrid stochastic–deterministic method, simulations of systems with chemical species present in millimolar concentrations (such as those found in nutrient based genetic switches) are now computationally accessible.
Fig. 3: The choice of communication timestep is crucial in recovering the stochastic dynamics of the system. (a) The distributions of the unbound G2 transporter (G2) at 0.055 mM extracellular galactose when the galactose switch system reaches a steady state at 700 minutes of simulation time. (b) The average species count of G2 as a function of time, with 2 mM extracellular galactose as an initial condition. (c) A Kernel Density Estimate with a histogram below of the times for G2 at 2 mM extracellular galactose to reach 80 percent of its average steady-state value. CME–ODE results are given for $\tau = 1$ s. (d) The Kolmogorov-Smirnov Statistic (showing divergence from pure CME distributions) of the protein distributions for G2 and the reporter protein at 0.05 mM and 2 mM extracellular galactose at 700 minutes simulation time.

Acknowledgment

D.M.B. thanks Piyush Labhsetwar for helping in simplifying the galactose switch system model. J.R.P. thanks support from the National Science Foundation (NSF) Graduate Fellowship Program under Grant DGE-1144245. This research was supported in part by NSF Grants: PoLS (PHY-1505008, PHY-1026550), MCB-1244570 and CPLC (PHY-1430124) as well as NIH grant: NIH-9P41GM10460123 and by NASA Grant no. NNA13AA91A. This research is part of the Blue Waters sustained-petascale computing project at NCSA and the University of Illinois Urbana-Champaign and is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois.

8 References

9 Appendices

9.1 Galactose Switch Model

The Python PyLM model for the galactose switch system as well as the code for using the hybrid algorithm through the “hookSimulation” feature of Lattice Microbes is available at http://www.scs.illinois.edu/schulten/software/ODE_CME.tar.gz.

The reaction model is included below for completeness. All rates are stochastic rates (i.e. the volume of the cell has been factored into reaction rate) and are presented in units of min\(^{-1}\).

9.2 Species

9.2.1 Genes:

\( DG1 \) - Gene encoding Gal1 with nothing bound

\( DG1\_Gal1\_G80d \) - Gene encoding Gal1 bound to the G4 dimer and G80 dimer

\( DG2 \) - Gene encoding Gal2 with nothing bound

\( DG2\_Gal2\_G80d \) - Gene encoding Gal2 bound to G4 dimer and G80 dimer

\( DG3 \) - Gene encoding Gal3 with nothing bound

\( DG3\_Gal3\_G80d \) - Gene encoding Gal3 bound to the G4 dimer and G80 dimer

\( DG80 \) - Gene encoding Gal80 with nothing bound

\( DG80\_Gal80\_G80d \) - Gene encoding Gal80 bound to G4 dimer and G80 dimer

\( DGrep \) - Gene encoding the reporter protein (YFP) with nothing bound

\( DGrep\_Gal4\_G80d \) - Gene encoding reporter bound to the Gal4 dimer and Gal80 dimer

9.2.2 mRNAs:

\( R1 \) - mRNA for Gal1

\( R2 \) - mRNA for Gal2

\( R3 \) - mRNA for Gal3

\( R4 \) - mRNA for Gal4

\( R80 \) - mRNA for Gal80

\( reporter\_rna \) - mRNA for the reporter gene

9.2.3 Proteins and Metabolites:

\( G1 \) - Gal1; galactokinase that metabolizes galactose

\( G2 \) - Gal2; galactose transporter

\( G3 \) - Gal3; galactose sensing transcription factor

\( G4 \) - Gal4; a monomer of the Gal4 transcriptional repressor

\( G4d \) - Gal4 dimer; the transcriptional repressor dimer in the nucleus

\( G80 \) - Gal80; nuclear; the monomer of the transcriptional repressor

\( G80C \) - Gal80; cytoplasmic; the monomer of the transcriptional repressor in the cytoplasm

\( G80d \) - Gal80 dimer; nuclear; a dimer of the transcriptional repressor

\( G80Cd \) - Gal80 dimer; cytoplasmic; a dimer of the transcriptional repressor in the cytoplasm

\( G80G3i \) - Gal80 dimer bound to Gal3i; the transcriptional repressor sequestered in the cytoplasm

\( GAI \) - Intracellular galactose

\( GAE \) - Extracellular galactose

\( G2GAI \) - Galactose bound to the Gal2 transporter on the intracellular side

\( G2GAE \) - Galactose bound to the Gal2 transporter on the extracellular side

\( G1GAI \) - Galactose bound to the Gal2 transporter on the extracellular side

\( reporter \) - A yellow fluorescence reporter protein (YFP)

9.3 Reaction model

9.3.1 Transcription:

\[
DG1\_G4d \quad k_d = 0.7379 \quad DG1\_G4 \quad k_r = 0.02236 \quad R1 + DG1\_G4d \quad \rightarrow R1 \quad DG1\_G4d

DG2\_G4d \quad k_d = 2.54154 \quad DG2\_G4d \quad k_r = 0.077035 \quad R2 + DG2\_G4d \quad \rightarrow R2 \quad DG2\_G4d

DG3\_G4 \quad k_d = 0.426572 \quad DG3\_G4d \quad k_r = 0.02666 \quad R3 + DG3\_G4d \quad \rightarrow R3 \quad DG3\_G4d
\]

\( DGrep\_G4d \quad k_d = 0.089902 \quad DGrep\_G4d \quad k_r = 0.0247552 \quad DGrep\_G4d \quad k_s = 1.1437 \quad reporter\_rna \quad DGrep\_G4d \quad reporter\_rna \quad k_s = 0.03466 \quad DGrep\_G4d \quad reporter\_rna \quad k_s = 0.028881 \quad R80 + DG80\_G4d \quad R80 \]

9.3.2 DNA Regulation:

\[
DG1 + G4d \quad k_s = 0.384615 \quad DG1\_Gad

DG1\_Gad \quad k_d = 0.085317 \quad DG1\_Gad + G80d

DG2 + G4d \quad k_s = 0.1 \quad DG2\_Gad

DG2\_Gad \quad k_d = 10.01 \quad DG2 + G4d

DG3 + G4d \quad k_s = 0.1 \quad DG3\_Gad

DG3\_Gad \quad k_d = 0.143989 \quad DG3\_Gad + G80d
\]

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.
9.3.3 Translation:
\[ R1 \quad k_1 = 1.92541 \rightarrow G1 + R1 \]
\[ G1 \quad k_2 = 0.003851 \rightarrow \emptyset \]
\[ R2 \quad k_3 = 13.4779 \rightarrow G2 + R2 \]
\[ G2 \quad k_4 = 0.003851 \rightarrow \emptyset \]
\[ R3 \quad k_5 = 55.4518 \rightarrow G3 + R3 \]
\[ G3 \quad k_6 = 0.01155 \rightarrow \emptyset \]
\[ R4 \quad k_7 = 10.7091 \rightarrow G4 + R4 \]
\[ G4 \quad k_8 = 0.0069315 \rightarrow \emptyset \]
\[ reporter_rna \quad k_9 = 0.7762 \rightarrow reporter + reporter_rna \]
\[ reporter \quad k_{10} = 0.01155 \rightarrow \emptyset \]
\[ R80 \quad k_{11} = 3.67368 \rightarrow G80 + R80 \]
\[ G80 \quad k_{12} = 0.0069315 \rightarrow \emptyset \]

9.3.4 Transport and Enzymatic:
\[ G2GA1 \quad k_1 = 4350 \rightarrow G2GAE \]
\[ G2GAE \quad k_2 = 4350 \rightarrow G2GAI \]
\[ G2GAE \quad k_3 = 2392.5 \rightarrow G2 \]
\[ G2 \quad k_4 = 0.00031353 + GAE \]
\[ G2GA \quad k_5 = 0.003851 \rightarrow \emptyset \]
\[ G2 + GAI \quad k_6 = 0.00031353 \rightarrow G2GAI \]
\[ G2GAI \quad k_7 = 2392.5 \rightarrow G2 + GAI \]
\[ G2GAI \quad k_8 = 0.003851 \rightarrow GAI \]
\[ G1 + GAI \quad k_9 = 0.00462426 \rightarrow G1GAI \]
\[ G1GAI \quad k_{10} = 1842.5 \rightarrow G1 + GAI \]
\[ G1GAI \quad k_{11} = 3350 \rightarrow G1 \]
\[ G1GAI \quad k_{12} = 0.003851 \rightarrow GAI \]

9.3.5 Transcription Factor:
\[ G3 + GAI \quad k_1 = 7.45 \rightarrow G3i \]
\[ G3i \quad k_2 = 890 \rightarrow G3 + GAI \]
\[ G3i \quad k_3 = 0.01155 \rightarrow GAI \]
\[ G80G3i \quad k_4 = 0.0159616 \rightarrow G3i + G80Cd \]
\[ G3i + G80Cd \quad k_5 = 0.25716 \rightarrow G80G3i \]
\[ G80G3i \quad k_6 = 0.005775 \rightarrow G4 + G4 \]
\[ G4 + G4 \quad k_7 = 0.0009315 \rightarrow G4d \]
\[ G4d \quad k_8 = 0.0069315 \rightarrow G80d \]
\[ G80 + G80d \quad k_{10} = 0.0009315 \rightarrow G80 + G80 \]
\[ G80d \quad k_{12} = 0.0069315 \rightarrow G80C + G80Cd \]
\[ G80C + G80Cd \quad k_{14} = 0.001 \rightarrow G80C + G80Cd \]
\[ G80d \quad k_{16} = 0.530504 \rightarrow G80d \]
\[ G80d \quad k_{17} = 0.530504 \rightarrow G80d \]
\[ G80d \quad k_{18} = 50 \rightarrow G80C \]
\[ G80C \quad k_{19} = 50 \rightarrow G80C \]

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. To cite the paper please use the doi provided on the Digital Library page.