Systems biology

Spatial stochastic modelling of the phosphoenolpyruvate-dependent phosphotransferase (PTS) pathway in *Escherichia coli*

J. Vidal Rodríguez[1,*], Jaap A. Kaandorp[1], Maciej Dobrzyński[2] and Joke G. Blom[2]

[1]Section Computational Science, Faculty of Science, University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands and [2]Center for Mathematics and Computer Science (CWI), Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

ABSTRACT

Motivation: Many biochemical networks involve reactions localized on the cell membrane. This can give rise to spatial gradients of the concentration of cytosolic species. Moreover, the number of membrane molecules can be small and stochastic effects can become relevant. Pathways usually consist of a complex interaction network and are characterized by a large set of parameters. The inclusion of spatial and stochastic effects is a major challenge in developing quantitative and dynamic models of pathways.

Results: We have developed a particle-based spatial stochastic method (GMP) to simulate biochemical networks in space, including fluctuations from the diffusion of particles and reactions. Gradients emerging from membrane reactions can be resolved. As case studies for the GMP method we used a simple gene expression system and the phosphoenolpyruvate:glucose phosphotransferase system pathway.

Availability: The source code for the GMP method is available at http://www.science.uva.nl/research/scs/CellMath/GMP

Contact: jrodrigu@science.uva.nl

1 INTRODUCTION

There is much evidence that stochastic effects play a key role in several biological processes, such as cell differentiation or gene regulatory networks and signalling pathways. The fluctuations in these phenomena are mostly due to biochemical processes involving a relatively small number of molecules, and the relevance of these fluctuations depends ultimately upon the underlying network architecture. Some networks components, such as switches, oscillators, positive and negative feedbacks, or transcription gene networks may display particular stochastic effects that cannot be captured using deterministic methods. The analysis of such stochastic effects can yield valuable additional insights. For example, it can help explaining part of the observed fluctuations in wet experiments, namely the internal fluctuations due to the chemical processes [intrinsic noise as in Paulsson (2004)]. The main advantage of stochastic results is the probability distribution of the steady state, trajectory or limit cycle (Arkin *et al.*, 1998; Doubrovinski and Howard, 2005; Srivastava *et al.*, 2002) as opposed to just the average value obtained with deterministic methods. Although effects are known to exist from theoretical analysis, only recently

several leading groups have measured intracellular molecular fluctuations *in vivo* (Kettling *et al.*, 1998; Arkin *et al.*, 1998; Becskei *et al.*, 2005; Thattai and van Oudenaarden, 2004; Doubrovinski and Howard, 2005; Pedraza and van Oudenaarden, 2005).

Many biochemical reactions do not occur in a completely homogeneous medium. For example, metabolic and signalling pathways involve reactions between cytosolic and membrane-bound molecules, causing an inhomogeneous spatial distribution of reactants. These membrane–cytosol reactions can only be sustained by the diffusion of cytosolic species, leading to an emergence of cytosolic gradients. The reaction rate and the diffusion coefficient will determine whether such gradient is significant. It is also probable that such gradients be unlikely for small prokaryotic cells but otherwise relevant in larger cells as shown in Francke *et al.* (2003).

Partly-membrane pathways can be modelled in the continuum and deterministic state space by partial differential equations (PDE). These models treat membrane reactions as boundary conditions on infinitely thin membranes. Alternatively a more 'natural' approach, adopted in particle-based models, is to account for the position in space of each particle (or a cluster of particles). Membrane-bound particles are then restricted to the membrane surface, or an adjacent sub-volume, and treated equally as cytosolic molecules with restricted mobility into the membrane subvolume. We base our work on this particular idea to tackle partly-membrane pathways.

Several models have been proposed to incorporate stochastic and spatial effects in biochemical networks, as reviewed in Takahashi *et al.* (2005). Currently there is no single model capable of efficiently coping with the broad range of spatial, temporal and concentration scales commonly found in biochemical networks. For instance, the highly accurate microscopic methods are too computationally demanding to simulate full pathways, while macroscopic ODE models cannot cope with spatial and stochastic phenomena. For this reason mesoscopic models may represent a plausible alternative approach and a compromise between computational efficiency and spatial and stochastic accuracy. A common approach for mesoscopic models is to tackle inhomogeneities, or spatial localization of reactions, by discretizing space into smaller homogeneous subvolumes. These subvolumes are considered to be a well-stirred system for which efficient well-stirred reaction methods can be used, either stochastic or deterministic (Weimar, 2002; Chopard *et al.*, 1994; Stundzia and Lumsden, 1996; Baras and Malek Mansour, 1996; Ander *et al.*, 2004; Bormann, 2001; Hattne *et al.*, 2005).

---

*To whom correspondence should be addressed.

In this article we present a method to model biochemical networks which combines the advantages of stochastic simulation and spatial discretization: the Gillespie-Multi-Particle method (GMP). Such a model can capture, within a certain range, spatial patterns and fluctuations due to diffusion and reactions. In contrast, models assuming a well-stirred medium can only capture reactional fluctuations. The GMP model is based on the discrete-in-space, operator-split and particle-based paradigms. These methods generally use dimensionless particles located in the lattice sites; diffusion takes place between lattice sites and reaction occurs locally in each site. Particularly, GMP builds on the multiparticle lattice gas automata model developed by Chopard *et al*. (1994). GMP employs the same diffusion process for particles, namely randomly distributing them among the nearest neighbours. For reactions, however, instead of Chopard's combinatoric method, we use the stochastic simulation algorithm (SSA) of Gillespie (1977). The separation of the reaction and diffusion processes distinguishes GMP from non-split models for the Reaction Diffusion Master Equation, where both processes are interlaced, such as the Stundzia and Lumsden (1996) approach. The core functionality of the Stundzia and Lumsden (1996) approach is a modified Gillespie method in which diffusion events for single particles are treated as a monomolecular reaction (diffusion events also follow an exponential distribution). Thus both processes are interlaced in a non-deterministic manner. In the GMP method we aim at reducing the computational cost of non-split models caused by the extra calculations of individual diffusion events. The reaction and diffusion processes are executed independently of each other (operator-split method) and the diffusion process is executed at predetermined times. Thus, the species population change in each site is due to an alternating reaction and diffusion process.

To test our method we first use a simple model of gene expression as analysed by van Zon and ten Wolde (2005). It illustrates how GMP performs in the regime where spatial fluctuations are important. Next, we apply the GMP method to a partially membrane pathway, viz., the phosphoenolpyruvate:glucose phosphotransferase system in *Escherichia coli* illustrated in Table 2. This case represents a real challenge for the method due to the complexity of the reactions, the spatial confinement of reactants, number of different species, and varying parameters. We validate the average results against a PDE model of the PTS pathway described in Blom and Peletier (2000) and used in Francke *et al*. (2003).

2 ALGORITHM

We combine the multiparticle method for diffusion (Chopard *et al*., 1994) and the kinetic Monte Carlo method (Gillespie, 1977) in our GMP algorithm. To tackle space we discretize the volume of the cell into a cubical lattice with L^3 sites (see Fig. 1 for an illustration of the spatial discretization we use). Each lattice site holds a discrete number of dimensionless, uniformly distributed particles. Consequently, there is no need to account for the exact position of individual particles. As a result of the spatial discretization, the overall inhomogeneity of the system is limited by the coarseness of the lattice. The geometry of the membrane is represented by the lattice sites enclosing the cytosol (light boxes). These membrane sites (dark boxes) also hold cytosolic molecules, enabling membrane–cytosol reactions. To simulate the membrane as a wall, we use reflective boundary conditions. Owing to spherical
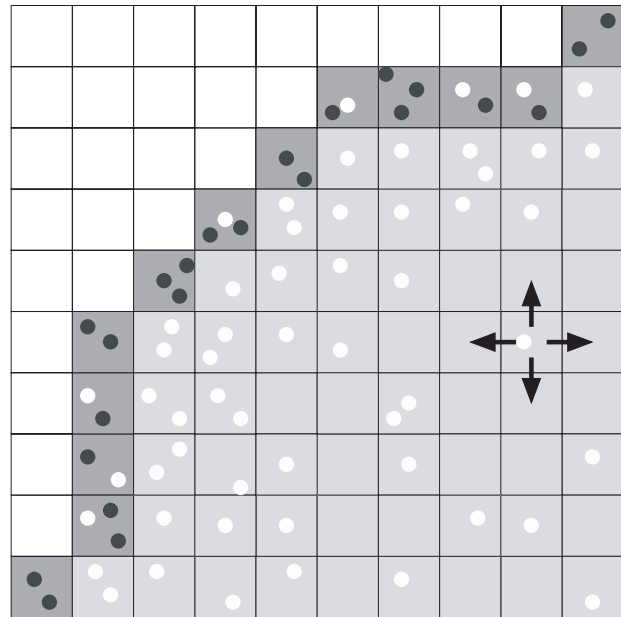

Fig. 1. 2D slice representation of the spherical cell geometry, including the membrane sites in dark grey and a possible distribution of particles. The length of a site is λ and the neighbourhood of a cytosolic site consists of four sites. Reflective boundary conditions are used for sites on the edge of the domain.

symmetry, we also use pseudo-symmetric boundary conditions on the edges of the cytosol.

An operator-split reaction–diffusion (RD) scheme alternates the execution of each process at predetermined intervals of time (see Algorithm 1). For each species S, its diffusion process will be executed at times $t_S = n_S \tau_{D_S}$, where n_S is the iteration number, and the diffusion step, τ_{D_S}, is given by

$$\tau_{\mathcal{D}_S} = \frac{1}{2d} \frac{\lambda^2}{D_S}, \qquad (1)$$

with d the dimensionality of the system.

Since each species may have a different diffusion coefficient, the next diffusion time is the minimum t_S that is larger than the current simulation time t_{sim}. Note that more than one species (even with different D_S) can occasionally have equal t_S. In this case all these species are diffused. The current simulation time t_{sim} is updated to t_S, and n_S is incremented by one for the diffused species.

Algorithm 1 Operator-split, Reaction–Diffusion algorithm.

0- Initialize $t_S = \min\{\tau_{\mathcal{D}_S}\}$ for all species S
1- $t_{\text{sim}} = 0$, $n_S = 1$ for all S
2- while $t_{\text{sim}} < t_{\text{end}}$ do
3-   while $t_{\text{sim}} < t_S$ do
4-   Reaction μ in τ_R on every lattice site
5-   Advance simulation time $t_{\text{sim}} = t_{\text{sim}} + \tau_R$
6-   Diffuse species for which $t_S = t_{\text{sim}}$
7-   Increment iteration n_S for the diffused species
8- $t_S = \min\{\tau_{\mathcal{D}_S} \cdot n_S\}$ for all S

The diffusion process is executed locally at every site. In this step particles are distributed randomly with equal probability among the

six nearest neighbours (four are shown in Fig. 1 in black arrows). Chopard *et al.* showed that this diffusion rule reproduces the macroscopic diffusion equation in the limit of $\lambda \to 0$, where λ is the lattice site's side length. Note that for large numbers of particles inside a lattice site it is not necessary to decide for a single particle where to diffuse, since the number of particles moving to a neighbouring site follows a normal distribution.

For the reaction process we use the SSA of Gillespie (1977), which solves the chemical master equation for well-stirred systems. This implies that the system is reaction limited at the scale of the lattice site. Note that we cannot have more spatial resolution than the lattice site size. Gillespie's method advances the system by executing the chemical reactions sequentially at reaction times τ_R. The multivariate distribution $P(\tau_R, \mu)$[1] for the time of next reaction τ_R that is of type μ is given by:

$$P(\tau_R, \mu) = \begin{cases} a_\mu \exp(-a_0 \tau_R), & 0 \leq \tau_R < \infty, \\ & \mu = 1, \ldots, M, \quad (2) \\ 0, & \text{otherwise} \end{cases}$$

where a_μ is called propensity function, with M the total number of reaction channels and $a_0 = \sum_{\mu=1}^{M} a_\mu$, where $a_\mu = k_\mu, n^A, n^B/(V, N_A)$, k_μ is the reaction rate, n^{A_i} is the number of particles in the site for species A, V is the site volume, and N_A Avogadro's number. We obtain τ_R and μ for a given species S as follows:

$$\tau_R = 1/a_0 \ln(1/r_1) \quad \text{and} \quad \sum_{i=0}^{\mu-1} a_i \leq a_0 r_2 < \sum_{i=\mu}^{M} a_i, \quad (3)$$

where r_1 and r_2 are uniformly distributed random numbers from the interval [0, 1]. Then reaction μ is effectuated and the simulation time t_{sim} is increased with τ_R.

Because of the operator-splitting, the reaction process executes reaction only for the duration of a diffusion interval (while $\sum \tau_R < \tau_{\mathcal{D}_S}$).

2.1 Choosing a lattice size

By adjusting the lattice size, λ, GMP's spatial resolution capabilities range from those offered by spatially resolved particle methods to well-stirred methods. However there is a limitation for the minimum lattice size λ. Despite the convergence of the diffusion process as shown by Chopard *et al.* (1994), the reaction–diffusion process does not converge for $\lambda \to 0$. In that limit no reactions can occur as the probability of finding two particles in a site is zero. Approaching this limit also implies that the maximum possible reaction rate decreases. Thus, for a pair of reacting particles and a diffusion limited process, there is a minimal lattice size $\lambda_o = 2\sigma$, where σ is the diameter of a particle. Note, however, that under this condition GMP would act almost as Brownian dynamics with reaction. For complex biochemical systems, in general we do not find purely diffusion limited reactions, which allows us to take larger subvolumes that are considered to be well-stirred. A reasonable size of the well-stirred subvolume is given by λ_{rmfp}, the reaction mean

free path[2], which is the average distance traveled between two reacting collisions (cf. also the discussion on the validity of the reaction diffusion master equation in Baras and Malek Mansour (1996)).

3 CASE STUDIES

We validate GMP against two problems that highlight two different aspects of GMP's capabilities. First, a generic gene expression model to illustrate the effects of the lattice size on the variance of the fluctuations. Second, the PTS pathway, which includes spatially confined membrane-bound reactions resulting in the inhomogeneity of the system due to cytosolic gradients.

3.1 Gene expression

The generic model for gene expression, as analysed by van Zon and ten Wolde (2005), enables us to study noise as caused by a low number of molecules and diffusion limited reactions. In van Zon and ten Wolde detailed continuous-in-space simulations are provided of the system which we use to validate our GMP method and check our reasoning to choose a proper lattice size.

The system we look at is a closed volume V of 1 μm^3 with a DNA binding site fixed in the centre surrounded by freely diffusing RNA (RNAp) polymerase. Once the DNA-RNAp complex is formed with association rate k_a, it can either dissociate back to separate DNA and polymerase (with rate k_d) or a protein P can be produced with a production rate k_{prod} with subsequent complex dissociation. The protein can further decay at the rate k_{dec}. Obviously the single protein production step in this model encompasses both transcription and translation which in fact consist of many biochemical reactions. Protein degradation is also simplified and treated as a first-order reaction. For the chemical reaction scheme and the parameters we refer to Table 1.

For choosing our lattice size we use the reasoning developed in Section 2.1. In van Zon and ten Wolde (2005) the size of the particles is 5 nm. We therefore obtain for the minimum λ a value of $\lambda_o = 10$ nm with $L = 100$. The λ_{rmfp} corresponds with $L \approx 18$.

3.2 The PTS pathway and parameter setting

To test our GMP method in a more complex biological scenario, we have used the phosphoenolpyruvate:glucose phosphotransferase system (PTS system), as illustrated in the figure in Table 2. The PTS was chosen because of its moderate network complexity, involving 17 species, fast and slow reactions, and especially because of the membrane-cytosol reactions. The PTS is a group-transfer pathway situated at the beginning of the glycolysis pathway. It is present in many bacteria and it is involved in the uptake and concomitant phosphorylation of external glucose that feeds the glycolysis. A phosphoryl group derived from PEP is carried along a series of bimolecular reactions to the membrane-bound IICB, which is responsible for taking the external glucose and transferring it into the cytoplasm. Table 2 summarizes the complete sequence of reactions as in Francke *et al.* (2003). In that paper the kinetic parameters were obtained from Rohwer *et al.* (2000) and the diffusion coefficients were based on measurements by Elowitz *et al.* (1999).

Bulk concentrations have been converted into numbers of molecules using one-eighth of a sphere's volume of 1 μm radius, discretized in a lattice of $L = 20$. Surface concentrations of

---

[1] $P(\tau, \mu)d\tau$ is the probability that, given the state X_S at time t, the next reaction μ will occur in the infinitesimal time interval $(t + \tau, t + \tau + d\tau)$.
[2] The reaction mean free path is defined as $\lambda_{rmfp} = (\langle \tau_R \rangle D_{AB})^{1/2}$, where $\langle \tau_R \rangle$ is the mean reaction time given by $1/a_\mu$, and $D_{AB} = D_A + D_B$ is the mutual diffusion coefficient.

Table 1. Reaction scheme and parameters associated with the gene expression model

Reaction		Rate	Species	Initial Concentration	Molecules	Dif.coef
DNA+RNAp $\xrightarrow{k_a}$	DNA-RNAp	3×10^9 M^{-1}s^{-1}	DNA		1	0
DNA+RNAp $\xleftarrow{k_d}$	DNA-RNAp	21.5 s^{-1}	RNAp	30 nM	18	1 μm^2 s^{-1}
DNA-RNAp $\xrightarrow{k_{prod}}$	P+DNA+RNAp	89.55 s^{-1}				
P $\xrightarrow{k_{dec}}$	∅	0.04 s^{-1}				

Table 2. Reaction list and parameters associated to the PTS model illustrated below

| Reaction		Rate[a]	Species	Initial Concentration[b]	Molecules	Species	Dif. coef.
EI + PEP $\xrightarrow{k_1}$	EI·PEP	1960	EI	5	1577	EI, EI·P, EI·P·Pyr	197.8
EI + PEP $\xleftarrow{k_{-1}}$	EI·PEP	480 000	HPr	50	15 766	EI·P·HPr	189.1
EI·PEP $\xrightarrow{k_2}$	EI·P + Pyr	108 000	IIA	40	12 613	HPr, HPr·P	378.0
EI·PEP $\xleftarrow{k_{-2}}$	EI·P + Pyr	294	IICB	3	3100*	HPr·P·IIA	262.1
HPr + EI·P $\xrightarrow{k_3}$	HPr·P·EI	14 000	PEP	2800	882 890	IIA, IIA·P	300
HPr + EI·P $\xleftarrow{k_{-3}}$	HPr·P·EI	14 000	Pyr	900	283 780	PEP, Pyr, Glc	18 000
HPr·P·EI $\xrightarrow{k_4}$	HPr·P + EI	84 000	Glc	500	70 767	all IICB	0
HPr·P·EI $\xleftarrow{k_{-4}}$	HPr·P + EI	3360	Glc·P	50	15 766		
IIA + HPr·P $\xrightarrow{k_5}$	IIA·P·HPr	21 960					
IIA + HPr·P $\xleftarrow{k_{-5}}$	IIA·P·HPr	21 960					
IIA·P·HPr $\xrightarrow{k_6}$	IIA·P + HPr	4392					
IIA·PHP·r $\xleftarrow{k_{-6}}$	IIA·P + HPr	3384					
IICB + IIA·P $\xrightarrow{k_7}$	IICB·P·IIA	880					
IICB + IIA·P $\xleftarrow{k_{-7}}$	IICB·P·IIA	880					
IICB·P·IIA $\xrightarrow{k_8}$	IICB·P + IIA	2640					
IICB·P·IIA $\xleftarrow{k_{-8}}$	IICB·P + IIA	960					
Glc + IICB·P $\xrightarrow{k_9}$	Glc·P·IICB	260					
Glc + IICB·P $\xleftarrow{k_{-9}}$	Glc·P·IICB	389					
Glc·P·IICB $\xrightarrow{k_{10}}$	Glc·P + IICB	4800					
Glc·P·IICB $\xleftarrow{k_{-10}}$	Glc·P + IICB	$5.4 \cdot 10^{-3}$					
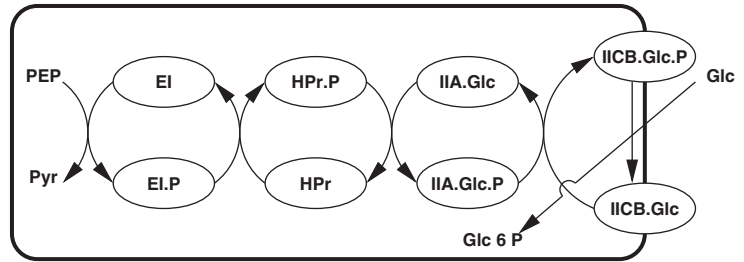
Illustration of the PTS system for *E.coli*. IICB molecules are bounded to the membrane. For the uptake of Glucose, a phosphoryl group is transported from cytosolic PEP to the membrane-bound protein IICB via the cytosolic enzymes HPr and IIA. Joined arrows indicate the formation of a compound.

[a]Bimolecular reaction rates are given in μM^{-1} min^{-1}, and monomolecular in min^{-1}.

[b]Diffusion coefficients are given in μm^2 min^{-1}.

[c]Initial concentrations for cytosolic species are given in μm, the surface concentration for IICB is given in μm, μm

These parameters were taken from the spatial PDE model discussed in Francke *et al.*, (2003), converting concentrations to number of molecules. Special consideration is put on the conversion of the membrane-bound concentration*. This parameter is adjusted to the value 2100 to obtain a better approximation of the PDE solution and is used for the results reported here.

membrane-bound molecules require special consideration. We assume that the concentration reported in Rohwer *et al.* (2000) for IICB was calculated as if the membrane-bound molecules were diluted in the cytosol and its concentration, 15 μmM, was measured. According to this, the number of IICB molecules should be 3100. Since the procedure of fitting IICB bulk concentration in the original model by Rohwer does not clearly indicate the number of membrane molecules, we have adjusted the number of IICB molecules in such a way to approximate the steady state of IIA·P·IICB and IICB·P·Glc. In addition, since we have to compare surface concentrations and numbers of molecules, we scale the surface concentrations of the PDE solution in such a way that the initial IICB concentration matches the number of particles, as is used in Fig. 2(c). We found that 2100 particles for IICB,

corresponding to an average of 4.3 particles per site, yield a good approximation for IIA·P·IICB and IICB·P·Glc but a less accurate approximation of IICB and IICB·P. Finally, external Glc molecules are confined to membrane sites in the reported concentration.

For the PTS system at steady state, the average value for λ_{rmfp} for cytosolic reactions is 0.06 ± 0.05. Since all species must share the same lattice size, $L = 20$ is a reasonable size.

4 RESULTS

4.1 Gene expression

With the parameters given in Table 1, simulations of the simple gene expression model yield an average level of protein P of 1032 ±

64, which is consistent with GFRD. However, the noise η in the levels of protein expression varies (see Table 3). We know that GMP converges to Gillespie for $L = 1$, and we observe the noise η converging to the value given by Gillespie. Using a lattice size which corresponds approximately with the reaction mean free path ($L = 18$), we obtain a noise level close to the GFRD reference value given by van Zon and ten Wolde (2005) using GFRD. By increasing the number of sites we obtain a better approximation, a noise level within 15% of the GFRD value, at the cost of more computations. However, L cannot be increased indefinitely since we then artificially reduce the probability of particle encounters and thus reduce the frequency of reactive collisions.

4.2 PTS Pathway

We have applied the GMP method to model the phosphoenolpyruvate:glucose phosphotransferase system (PTS) system as indicated in section 3.2. A first check of the validity of our method consists of looking at the System's Mass Evolution in Time (SMET) which shows the total number of particles present in the system over time regardless of their position. In Figure 2 we show the SMET of molecules, showing a good agreement with the reference solution of the PDE model by Blom and Peletier (2000). These results are a direct consequence of the initial decision to adjust the number of molecules of IICB to match the $IIA \cdot P \cdot IICB$ steady state. Nevertheless, we note in Figure 2a, a very good agreement of the EI-related molecules. In Figure 2b, despite a good initial agreement ($t < 0.005$ min), the total number of molecules starts to diverge, and settling at levels $\sim 6\%$ for HPr species, and 20% for $IIA \cdot P$. Finally, the level of the phosphorylated $IICB \cdot P$, is four times as high as the PDE solution. However, in absolute numbers, the extra ≈ 100 molecules in $IICB \cdot P$, seem to be lacking in IICB.

In Figure 3 we show the radial gradients, at steady state ($t_{sim} = 0.005$ min) for some enzymes. To obtain the gradient along the radius we average the number of particles over all sites with the centre in the range $[r - 0.5/L, r + 0.5/L]$. We then average again over 50 trials of the same systems with uniform random distributed particles. Therefore, results near the membrane will always contain far more points than those near the centre, so the average near the centre may show a larger deviation from the real mean. In Figure 3a, we observe two different behaviours of the obtained gradients: the phosphorylated enzymes HPr show only a difference in the gradient higher near the membrane (10%), whereas for $IIA \cdot P$ its gradient is upward shifted along the whole radius (30% near the membrane and 10% near the centre of the cell). These differences are linked with those observed in the SMET in Figure 2. Note, that owing to the spherical geometry, a slight divergence near the membrane has a larger contribution to the SMET difference than a divergence near the centre. Turning to the three basic enzymes shown in Figure 3b, we find for HPr a relative concentration difference between the centre and the membrane of the cell of 66% in both the PDE solution and the GMP averaged solution. However, in absolute value, GMP's solution for the radial gradient is consistently 16% lower than the PDE solution. Both the gradients for IIA and EI show an excellent agreement. Nevertheless, some of the molecules involved in their reactions, such as $IIA \cdot P$, $IICB \cdot P$ and $HPr \cdot P$, show a noticeable difference as seen in Figure 2.
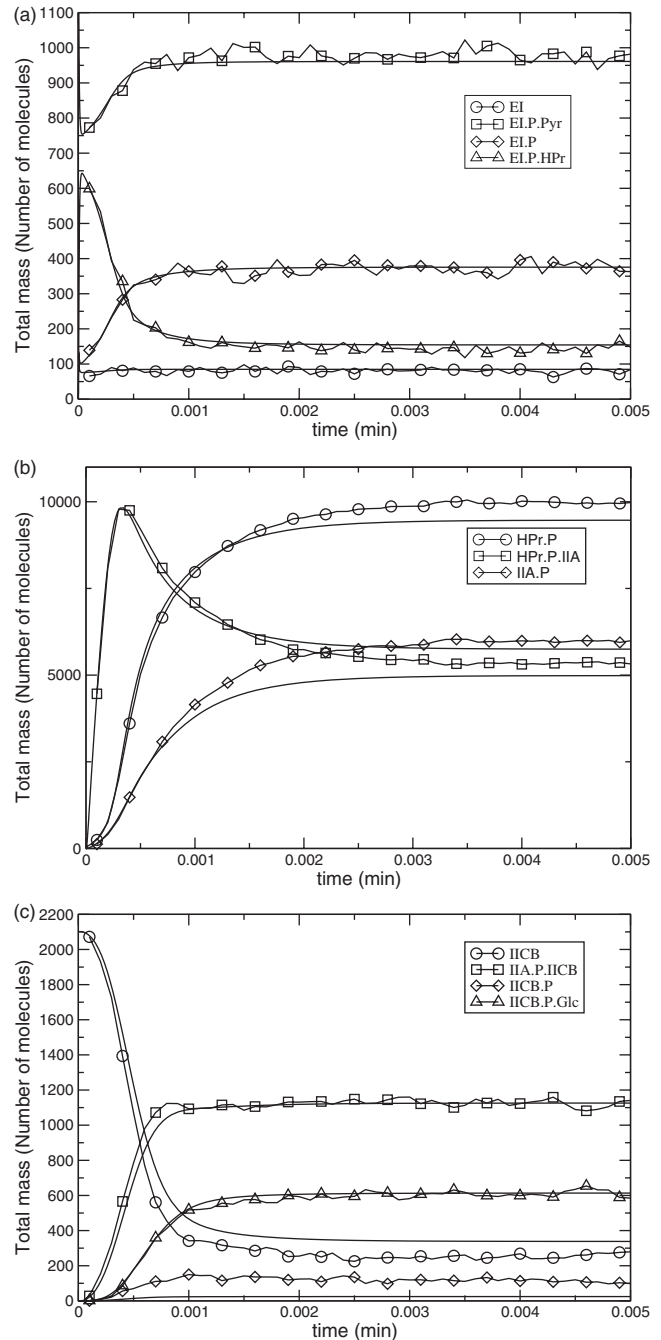

Fig. 2. System Mass Evolution in Time (SMET) graph for the PTS pathway with parameters shown in Table 2 and a cell's radius of 1 μm. Only one trial is shown here for the GMP result. EI-related enzymes show a very good agreement between the GMP result and the PDE solution. The membrane concentration of IICB enzyme in subfigure (c) was scaled to match the number of molecules at t = 1e-5.

5 DISCUSSION

We have developed the GMP reaction–diffusion method along the lines of other spatial stochastic kinetic model solvers, such as Stundzia and Lumsden (1996), Ander *et al*. (2004), Bormann
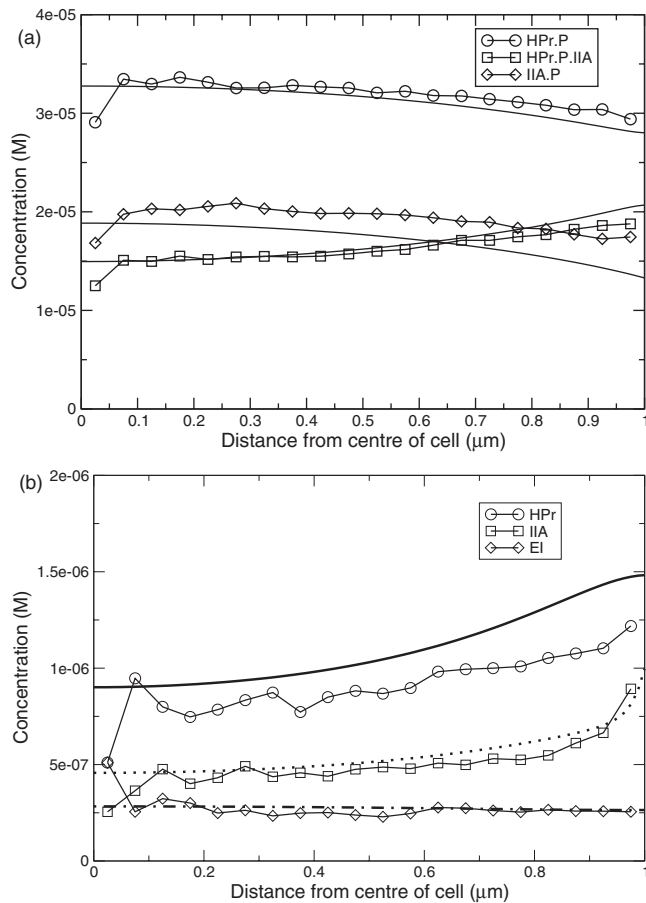
Fig. 3. Profile gradients along the radius of a 1/8th of an sphere for the PTS system. Lines without glyphs are the solution of the PDE model in Blom and Peletier (2000). Each point is the average of all sites with centre in $[r − 0.5/L, r + 0.5/L]$, and then averaged again over 50 trials. Note the good agreement of IIA, a consequence of the tuning of IICB molecules to match $IIA·P·IICB$ as seen in Figure 2c. EI has no significant gradient, while HPr differs by ∼15% from the value predicted by the PDE model. The first point near the centre shows a large deviation because it is only the average of one site and 50 trials.

(2001), Hattne *et al.*, (2005). For GMP, the distinguishing feature is the separation of the reaction and diffusion processes at the scale of a lattice site. The consequent locality of the operations results in a fast algorithm. Since at the lattice site scale no spatial information of particles is used, GMP is very suitable for parallel computing. Such computational power will be required when simulation studies of whole-cell processes become necessary as experimental biology advances. GMP can cope with both reaction limited and diffusion limited problems, as well as with confined distributions of chemicals in biological systems. For complex biochemical networks a trade-off in choosing the lattice size is necessary, since different reaction rates and diffusion coefficients result in different optimal lattice sizes. Consequently, some accuracy, compared with (hard-sphere) particle models, is lost in order to gain computational speed. Nevertheless, the results can be used to guide experimentalists since such loss can be controlled and minimized.

Table 3. Measurements of noise for the gene expression using GMP with various lattice sizes L

L	Gillespie	5	10	20	30	50	GFRD
η	0.028	0.047	0.039	0.11	0.12	0.17	0.14

Initial conditions and parameters are given in Table 1. The optimal L corresponding to the reaction mean free path is 18.

The GMP method takes into account the principal sources of fluctuations, but the class of lattice-discrete methods, to which GMP belongs, smoothes out some noise if the lattice size is too large, because of the use of Gillespie as a reaction method (cf. Table 3). Note, however, that λ_{rmfp} is a safe lattice size, which is valid for small concentrations. For larger concentrations λ_{rmfp} decreases and in that case the slope of the gradients and the error we allow determine the optimal lattice size. For more complex biochemical systems there is no single optimal lattice size. In those cases an average λ_{rmfp} can be used as an orientative value, as we did for the PTS case in Section 4.2.

For the PTS system, the main focus was on the reactions between membrane and cytosolic species, as well as on the gradients that may arise due to the spatial localization of the membrane-bound reactions and the diffusion of cytosolic molecules. Owing to the relatively high concentrations fluctuations are not significant. As shown in Figure 3, only the species related to HPr and IIA show a significant gradient, in contrast to a flat gradient by EI which has the lowest concentration. This case illustrates that spatial inhomogeneities do not depend on the concentration levels, but on the whole pathway architecture, consisting of reaction rates, diffusion coefficients, spatial restrictions and dependencies.

Having tuned the initial number of membrane-bound molecules of IICB, we obtained a good approximation of $IIA·P·IICB$ and IIA, but IICB, $IICB·P$ and $IIA·P$ differ from the PDE solution (see Figs 2 and 3). This difference is propagated down the pathway irregularly. For instance, it relatively strongly affects the gradient of HPr, though not in absolute sense ($< 2e$-$7\mu M$) (see Fig. 3b) and not of IIA. The reaction procedure between bulk (cytosol) and surface (membrane) molecules, which is approximated by a bulk to bulk reaction in a small volume, remains a topic for further research. Note, that in the GMP method we have to use a constant number of particles rather than a constant concentration for the membrane-bound molecules as used in PDE models. Although this is a more realistic representation of the underlying biological process and avoids the arbitrary use of concentrations of membrane-bound molecules, this introduces a practical difficulty, since this number will usually be unknown. In PDE models this concentration is estimated from the bulk concentration, which does not necessarily provide a correct estimation.

The GMP method can be easily extended to include diffusion over the membrane of membrane-bound particles and the effect of clusters of membrane-bound particles on a biochemical network can be studied. Other future aims are to improve the geometrical representation of membranes and the particle-based representation of cytosol-membrane reactions, and to extend the analysis of the choice of a proper lattice size.

ACKNOWLEDGEMENTS

Conflict of Interest: none declared.

REFERENCES

Ander,M. *et al.* (2004) SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *Syst. Biol.*, **1**, 129–138.

Arkin,A. *et al.* (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics*, **149**, 1633–1648.

Baras,F. and Malek Mansour,M. (1996) Reaction–diffusion master equation: a comparison with microscopic simulations. *Phys. Rev. E*, **54**, 6139–6148.

Becskei,A. *et al.* (2005) Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat. Genet.*, **37**, 937–944.

Blom,J.G. and Peletier,M.A. (2000) *Diffusive gradients in the PTS system*, Technical Report MAS-R0020, CWI, Amsterdam.

Blom,J.G. and Peletier,M.A. (2002) *The Importance of Being Cigar-shaped*, Technical report MAS-R0228, CWI, Amsterdam.

Bormann,G. (2001) Analysis and implementation of a MC method for reaction–diffusion in biophysical realistic compartimentized neuromodel. PhD Thesis, Universiteit Antwerpen, Belgium.

Chopard,B. *et al.* (1994) Multiparticle lattice gas automata for reaction diffusion systems. *Int. J. Mod. Phys. C*, **5**, 47–63.

Doubrovinski,K. and Howard,M. (2005) Stochastic model for Soj relocation dynamics in *Bacilus subtilis*. *Proc. Natl Acad. Sci. USA*, **102**, 9808–9813.

Elowitz,M.B. *et al.* (1999) Protein mobility in the cytoplasm of *Escherichia coli*. *J. Bacteriol.*, **181**, 197–203.

Francke,C. *et al.* (2003) Why the phosphotransferase system of *Escherichia coli* escapes diffusion limitation. *Biophysical*, **85**, 612–622.

Gillespie,D.T. (1977) Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, **81**, 2340–2361.

Hattne,J. *et al.* (2005) Stochastic reaction–diffusion simulation with MesoRD. *Bioinformatics*, **21**, 2923–2924.

Kettling,U. *et al.* (1998) Real-time enzyme kinetics monitored by dual-color fluorescence cross-correlation spectroscopy. *Proc. Natl. Acad. Sci. USA*, **95**, 1416–1420.

Paulsson,J. (2004) Summing up the noise in gene networks. *Nature*, **427**, 415–418.

Pedraza,J.M. and van Oudenaarden,A. (2005) Noise propagation in gene networks. *Science*, **307**, 1965–1969.

Rohwer,J. *et al.* (2000) Understanding glucose transport by the bacterial phospoenolpyruvate: glycose phosphotransferase system on the basis of kinetic measurements *in vitro*. *J. Biol. Chem.*, **275**, 34909–34921.

Srivastava,R. *et al.* (2002) Stochastic vs. deterministic modeling of intracellular viral kinetics. *J. Theor. Biol.*, **218**, 309–321.

Stundzia,A.B. and Lumsden,C.J. (1996) Stochastic simulation of coupled reaction–diffusion processes. *J. Comput. Phys.*, **127**, 196–207.

Takahashi,K. *et al.* (2005) Space in systems biology of signaling pathways—towards intracellular molecular crowding *in silico*. *FEBS Lett.*, **579**, 1783–1788.

Thattai,M. and van Oudenaarden,A. (2004) Stochastic gene expression in fluctuating Environments. *Genetics*, **167**, 523–530.

van Zon,J.S. and ten Wolde,P.R. (2005) Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics. *Phys. Rev. Let.*, **94**, 128103.

Weimar,J. (2002) Cellular automata approaches to enzymatic reaction networks. In Bandini,S. *et al.* (eds), *ACRI, Lecture Notes in Computer Science*. Springer, Berlin/Heidelberg, Germany. pp. 294–303.