Particle-based stochastic simulations

Nicolas Le Novère and Dominic Tolle
EMBL-EBI
Wellcome-Trust Genome Campus, Hinxton CB10 1SD, UK

Abstract

Most of the modelling approaches used in Computational Systems Biology are population-based, discarding reactant individuality. While those methods have been, and still are, tremendously useful, the complexity of cellular reactions calls for another type of modelling approach, based on particle-based simulations. The discrete stochastic approaches permit to track the state and position of each single molecules, paving the way to the use of multistate components and accurate simulation of spatial events. Those features are particularly important when it comes to modelling signal transduction machinery.

1 Introduction

The vast majority of mathematical models developed to understand biochemical or cellular behaviours are based on assumptions coming from classical chemical kinetics. In particular, the Mass Action Law, expressed by Waage and Guldberg in 1864 [20] still largely dominates the scene. In modern words, this law stipulates that the rate of a reaction is proportional to the product the "activity" of the reactants. In a diluted solution, those activities are equal to the concentrations of the reactants. As a consequence biochemical modelling and simulation largely relied on deterministic methods describing the evolution of populations with differential equations.

However, this law is valid only if the behaviour of a particular reactant is ignored and only the average behaviours of the populations are considered. This ideal situation is of course rarely matched in Cell Biology, where the

small number of reacting entities cause the individual behaviours to be significant. This led to the development of stochastic methods, the most widely used belonging to the family of approaches stemming form Gillespie's algorithm [8]. While those methods permit to handle the noise mechanistically, they are still intrinsically *population-based* in most of the cases. Therefore, one cannot track a particular reactant across time steps. Two major consequences are the explosion of the number of component and reactions due to the multistate molecules and the impossibility to build "true" spatial models.

2 Combinatorial explosion and multistates

In a population-based model, one has to keep track of the concentrations of all reactants, whether those reactants are different types of molecules, or the "same" molecules under different states (e.g., covalent modifications, 3D conformations etc.). For instance, a model representing the MAPK cascade on a molecular scaffold [12], comprises almost 50 different species representing the various phosphorylation states of RAF, MEK and ERK, bound or not to the molecular scaffold, while there are only 4 different proteins involved. Very large assemblies of proteins such as the post-synaptic machinery can reach a number of different states far more important than the total number of molecules in the system. Furthermore, each new state doubles the number of potential reactions, resulting in very inefficient simulations, where most species pools are empty, and most of the reactions cannot occur.

This problem has been tackled in *particle-based* simulators using multistate molecules, such as in the simulator StochSim [15]. In StochSim, the model of MAPK cascade described above has only four multistate molecules. The number of possible states is still important. However, because the software tracks every existing molecule, rather than all possible ones, there is no wasting computation on unpopulated pools.

3 Spatial modelling

An overview of the methods used to simulate spatial events in Systems Biology has been presented by Takahashi [19]. A popular method used to run stochastic simulations bearing spatial information is the extension of Gillespie's next reaction method to 3D lattices of small volumes [18]. This

approach is used for instance in the simulators MesoRD [9] or SmartCell [1]. Here again, the use of a population-based approach rather than a particle-based approach precludes the simulation of situations where the relative positions and states of the molecules are important.

Several attempts have been made to realistically represent diffusion of single molecules. One of the earliest was MCell [4]. The software has been used mainly to model synaptic transmission. Its accurate representation of space allowed for instance to investigate the influence of the relative positions of various neurotransmitter receptors [7]. The main problem of MCell is that only small molecules are mobile, the proteins being represented as reactive surfaces, and the reactions decided on the basis of ray-tracing, i.e. occuring at surfaces.

An alternative approach has been used in the software Smoldyn [2], where the diffusion of all molecules is considered using Brownian dynamics. The program allowed the development of highly accurate simulations of *E. coli* chemotaxis [14, 3]. However, Smoldyn considers molecules with different states as different species. While the problem of the number of species types is irrelevant (since single molecules are tracked rather than pools), the combinatorial explosion of reactions is still not addressed.

4 Signalling lattices

Signalling lattices are example cellular systems, where both the position and the state of individual elements are important. Those lattice typically functions as memory devices, and the history of each element is recorded in their state. Because of the complexity of even the simplest such lattice, ad-hoc programs are often used to model them. A recurrent pattern is the circular lattice of multistate elements, such as the flagella motor of the bacteria [6], or the calcium/calmodulin kinase II [21]. The software StochSim has been extended to handle bi-dimensional lattices of various geometry [11]. Each cell of the lattice can contain one multistate molecule, and the lattice functions as a stochastic cellular automaton. Very accurate simulations the lattice of chemotaxis receptors of *E. coli* [16] have been developed using StochSim [13, 17].

5 Generic particle-based simulation of neuronal signalling

The Abstracted Protein Simulator (APS) is a software which enables proteins to be simulated at a highly abstracted level, initially developed by Dan Mossops and Fred Howell at the University of Edinburgh. We extended the software to render the simulations more realistic. Each protein is built from simple geometric elements (spheres and cylinders). Binding sites can be added anywhere on the resulting assemblies. APS can be extended to define compartments of various dimensionality, with different diffusion laws. The reactions taking place upon molecule encounters are treated probabilistically.

When several molecules assemble in a cluster, they retain their individuality, and therefore their state, a crucial feature when modelling biological processes such as synaptic plasticity. Our simulations permit not only to study the effect of the relative positions of neurotransmitter receptors, as with MCell, but also the consequences of their movements [5] and their aggregation on the signal processing. We are working to develop further our models, in an attempt to provide the most detailed simulations of a post-synaptic density so far.

6 Conclusion

Particle-based simulations offer an alternative to classical population-based approaches, tackling molecular behaviours in a more realistic way. Despite the fact that they generally requiring dedicated software and are more computationally demanding, the steady increase in computing power, associated with the continuous improvement of single-molecule recordings, should make the use of there methods more widespread in the future.

References

[1] Ander, M., Beltrao, P., Di Ventura, B., Ferkinghoff-Borg, J., Foglierini, M., Kaplan, A., Lemerle, C., Tomás-Oliveira, L., Serrano., L., SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *IEE Syst Biol* **1**, 129–138 (2004).

[2] Andrews, S.S., Bray, D., Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Physical Biol* **1**, 137–151 (2004).

[3] Andrews, S.S., Serial rebinding of ligands to clustered receptors as exemplified by bacterial chemotaxis. *Physical Biol* **2**, 111–122 (2005).

[4] Bartol, T.M. Jr, Land, B.R., Salpeter, E.E., Salpeter, M.M., Monte Carlo simulation of miniature endplate current generation in the vertebrate neuromuscular junction *Biophys J* **59**, 1290–1307 (1991).

[5] Choquet, D., Triller, A., The role of receptor diffusion in the organization of the postsynaptic membrane. *Nat Rev Neurosci.* **4**, 251–265 (2003).

[6] Duke, T.A.J., Le Novère, N., Bray, D., Conformational Spread in a Ring of Proteins: A Stochastic Approach to Allostery. *J Mol biol* **308**, 541–553 (2001).

[7] Franks, K.M., Stevens, C.F., Sejnowski, T.J., Independent Sources of Quantal Variability at Single Glutamatergic Synapses. *J Neurosci* **23**, 3185–3195 (1976).

[8] Gillespie, D.T., A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Physics* **22**, 403–434 (1976).

[9] Hattne, J., Fange, D., Elf, J., Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics* **21**, 2923–2924 (2005).

[10] Hucka, M., Bolouri, H., Finney, A., Sauro, H., Doyle, J., *et al.*, The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524–531 (2003).

[11] Le Novère, N., Shimizu, T.S., StochSim: modelling of stochastic biomolecular processes. *Bioinformatics* **17**, 575–576 (2001).

[12] Levchenko, A., Bruck, J., Sternberg, P.W., Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. *Proc Natl Acad Sci USA* **97**, 5818–5823 (2000).

[13] Levin, M.D., Shimizu, T.S., Bray, D., Binding and diffusion of CheR molecules within a cluster of membrane receptors. *Biophys J* **82**, 1809–1817 (2002).

[14] Lipkow, K., Andrews, S.S., Bray, D., Simulated Diffusion of Phosphorylated CheY through the Cytoplasm of Escherichia coli. *J Bacteriol* **187**, 45–53 (2005).

[15] Morton-Firth, C.J., Bray, D., Predicting temporal fluctuations in an intracellular signalling pathway. *J Theor Biol* **192**, 117-128 (1998).

[16] Shimizu, T.S., Le Novère, N., Levin, M.D., Beavil, A.J., Sutton, B.J., Bray, D., . Molecular model of a lattice of signalling proteins involved in bacterial chemotaxis. *Nat Cell Biol* , 792–796 (2000).

[17] Shimizu, T.S., Aksenov, S.V., Bray, D., A Spatially Extended Stochastic Model of the Bacterial Chemotaxis Signalling Pathway. *J Mol Biol* **329**, 291–309 (2003).

[18] Stundzia, A.B., Lumsden, C.J., Stochastic Simulation of Coupled Reaction-Diffusion Processes. *J Comput Physics* **127**, 196–207 (2003).

[19] Takahashi, K., Arjunan, S.N.V., Tomita, M., Space in systems biology of signaling pathways — towards intracellular molecular crowding *in silico*. *FEBS Lett* **579**, 1783–1788 (2005).

[20] Waage, P., Guldberg, C.M., Studier over affiniteten I. *Forhandlinger I Videnskabs-selskabet I Christiania*, 35–45 (1864).

[21] Zhabotinski, A.M., Bistability in the Ca(2+)/calmodulin-dependent protein kinase-phosphatase system. *Biophys J.* **79**,2211–2221 (2000).

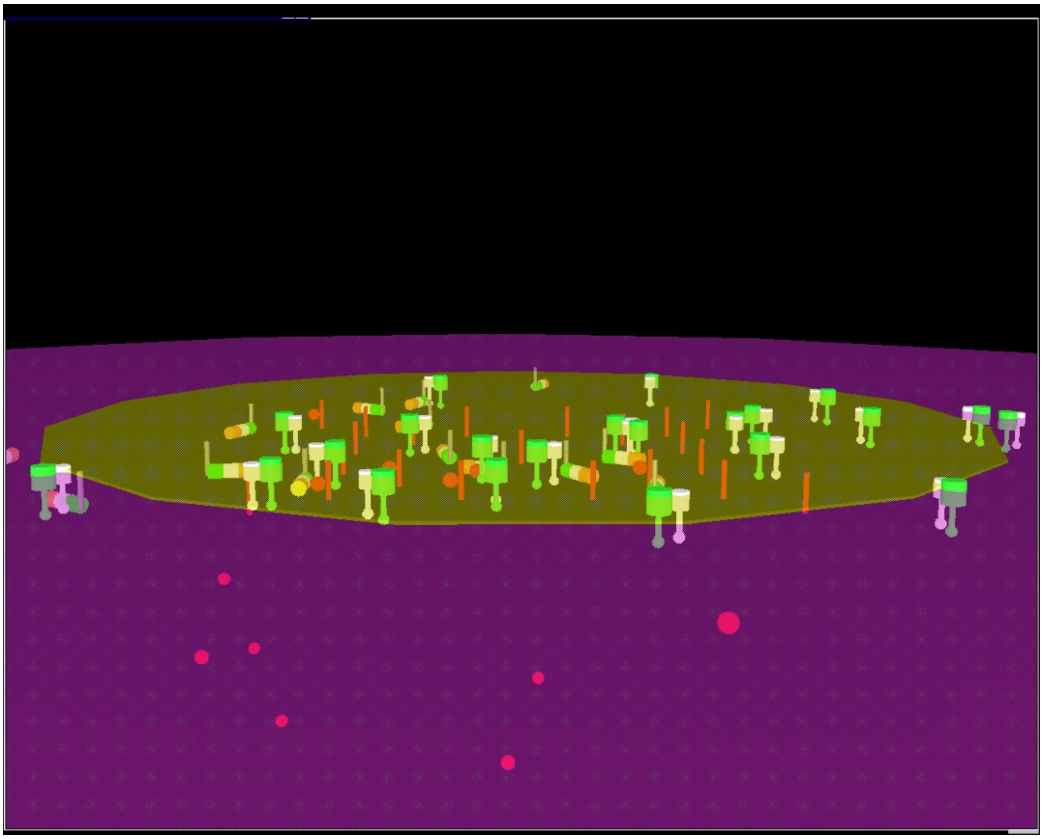

Figure 1: Dynamical model of the post-synaptic density of a glutamatergic synapse. The AMPA glutamate receptor (green), stargazin (white), PSD95 (tricolour), cytoskeletal proteins (red rods) and adaptor proteins (red spheres) are represented. The post-synaptic membrane is yellow, while extrasynaptic surface is purple.