Minireview

Space as the final frontier in stochastic simulations of biological systems

Caroline Lemerle, Barbara Di Ventura, Luis Serrano*

European Molecular Biology Lab, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Accepted 4 February 2005

Available online 11 February 2005

Edited by Robert Russell and Giulio Superti-Furga

Abstract Recent technological and theoretical advances are only now allowing the simulation of detailed kinetic models of biological systems that reflect the stochastic movement and reactivity of individual molecules within cellular compartments. The behavior of many systems could not be properly understood without this level of resolution, opening up new perspectives of using computer simulations to accelerate biological research. We review the modeling methodology applied to stochastic spatial models, also to the attention of non-expert potential users. Modeling choices, current limitations and perspectives of improvement of current general-purpose modeling/simulation platforms for biological systems are discussed.

Keywords: Spatial modeling; Stochastic simulation; Reaction-diffusion; Simulation algorithm; Modeling software

1. Introduction

Computational cell biology could well go for a medal of most exciting and active area of cross-disciplinary research these days. Not only are new modeling/simulation platforms getting invaluable input from biologists, computer scientists, theoretical physicists, applied mathematicians, and engineers, mainstream biology is learning to integrate these new tools to the core of their scientific investigation: generating and validating new hypotheses [1,2]. Benefiting from a concurrent revolution in experimental tools able to probe living cells down to atomic levels, collaborative projects have emerged setting the objective to one day run molecular-scale models of whole cells [3,4], a goal by now commonly expressed in systems biology. The modular architecture of cell function justifies placing the challenge today at building detailed models of individual cell biological processes, realistic enough to help capture the logic of how they are designed to behave. This calls for modeling simulation software that can account not only for the stochastic nature of molecular interactions, but also for the spatial heterogeneity of cellular matter. Only recently has the technological forefront started to produce the wealth of quantitative data needed to run simulations. Also, computational methods have improved only recently to allow useful simulations of such complex systems, with conventional computers fast enough to run them. Finally, only recently has biological knowledge begun to include descriptions of when, where and with whom molecules interact to carry out their function in vivo [5], crucial information to start experimenting with spatio-temporal models [6]. With space and stochasticity, we are reaching the final frontier of the in silico exploration of living cells, with the mission to understand, control or redesign them.

Present modeling/simulation platforms able to deal with stochastic spatially resolved models are rapidly expanding to offer a wealth of features that the non-expert user can find overwhelming. We attempt in this review to present the state-of-the-art of such general-purpose software, describing the modeling methodology, highlighting the approximations inherent to modeling objects of such complex nature, and presenting perspectives of refinement. References are made to the following actively developed and freely available programs: Virtual Cell [7], E-Cell [8], MCell [9], Smoldyn [10], MesoRD [11], and STOCHSIM [12] as well as SmartCell [13], developed in our lab as a tool to analyze stochastic spatially relevant biological processes (see Table 1).

2. Modeling natural or synthetic systems

A paradox with the massive amount of experimental data presently accumulating is that, despite its being invaluable for constructing spatio-temporal models of biochemical processes, for any given natural system, the available data remain largely insufficient to run useful simulations. For instance, while data regarding expression patterns, protein–protein interactions and localization of proteins within cells are available for many systems, quantitative data regarding concentrations, reaction rates, diffusion or transport rates are in general missing.

In principle, a proper chemical physics theory giving a force law that predicts how any two molecules interact would "suffice" to simulate stochastically the life of a whole cell with molecules resolved to their atomic detail, provided one could supply as initial conditions a snapshot of the system at atomic-resolution and enough computing power. Higher-level parameters that describe the behavior of molecules and their assemblies into cellular structures could all be derived from such knowledge, alleviating the present need to extract them from macroscopic experiments. With equally tremendous progress in the imaging field, the simulation output produced could be compared to "biochemical imaging" data, tracking each molecule simultaneously in vivo over time, for refinement of the theory. In reality, this is not yet the case, leaving

*Corrresponding author. Fax: +49 6221 387 306.
E-mail address: serrano@embl-heidelberg.de (L. Serrano).

0014-5793/$30.00
doi:10.1016/j.febslet.2005.02.009

Table 1

	SmartCell	Virtual Cell	mesoRD	Stochsim	MCell	Smoldyn
Model editor	Yes	Yes	No	Yes	No	No
Geometry editor	Yes	Yes	CSG, not graphical	Incomplete	Associated program DReAMM	No
Database of model geometries availability	Yes	Yes	No	Yes	No	No
Continuous space supported	No	Yes	No	No	Yes	Yes
Space discretization supported	Yes (voxelization)	No	Yes (voxelization)	Yes (2D regular lattice)	No	No
Use of analytical formulae for defying geometry	No	Yes	No	No	No	No
Image processing tool	No	Yes	No	No	DReAMM	No
Geometry dimension availability	0–3D	1–3D	3D	2D	3D	1–3D
Lower-dimension spatial elements representation	Yes	Not 1D	No	No	Not 1D	2D (only delimiting walls)
Representation of molecular entities	Populations	Concentrations	Populations	Individuals	Individuals	
Individuals						
Features of molecular entities	Size, charge, MW	Linking to databases	Size	Conformational state	None	None
Diffusion supported	Yes	Yes	Yes	No	Yes	Yes
Enabling conditions for biomolecular reactions	In same or neighboring voxels	Term in PDEs	In same voxel	Nearest-neighbours	Collision detected	Binding radius
Mathematical framework	Stoch	Det	Stoch	Stoch	Stoch	Stoch
Sensitivity analysis availability	No	Yes	No	No	No	No
Checkpointing	Under dev.	Yes	No	Yes	Yes	Yes

modelers with two alternatives: either considering a tremendously simplified model of a cell or very large interaction networks to prove general concepts, or analyzing simpler systems for which enough data is available. In the second case, the user can focus on parts of the whole cell network, simulating the corresponding processes in isolation or with loose connectivity to the rest of the cell. Alternatively, modelers can design simple systems from scratch, for which parameters are under control. Knowledge gained from these different approaches can be used to analyze progressively larger systems.

3. What to include in a model

Building a useful stand-alone spatio-temporal model of an isolated process requires selecting for each species a level of representation depending on its functional role in the context of this specific process, and a set of relevant reactions. This choice is a compromise between current knowledge of the process, computer speed and minimal representation able to capture the assigned function, keeping in mind that the presence in the reaction network of molecular species involved in other cellular processes (such as signaling molecules, "housekeeping" proteins, etc.) are potential sources of cross-talk unaccounted for in the isolated system [14,15].

Molecular species participating in the chosen reaction network are modeled explicitly as reactive species, while the species inert in respect to the network still account for the cellular environment embedding it. Those molecules responsible for cell structure, be it those forming 2D assemblies (such as biological membranes), or those assembled in 1D structures (DNA, cytoskeleton elements, etc.), compose the model geometry (respectively 2D or 1D cellular structures), while the bulk of remaining molecular species, whose contribution goes to the crowding of cellular space, is typically accounted for only in the local behavior of reactive species (possibly as obstacles in the 3D compartments) [16,17].

4. Model specification using a modeling/simulation software

While countless numbers of custom-made programs designed for local users used to be the rule, general-purpose software tools, relying on basic computer proficiency of users, are being developed with obvious efforts to adapt to a broader community [18], offering extensive documentation, workspaces that help structure the modeling process, rich graphical user interfaces, links to databases or to the community (Virtual Cell), and standards in model specification languages (open standards like SBML [19] and CellML [20] based on the XML markup language).

Modeling issues raised for the different steps of model building are discussed below.

4.1. Model geometry

With the addition of space, not only can cellular structures and heterogeneous distributions begin to be modeled, so can the random movement of molecules in cellular space. The main schemes used to build model geometries in current software are presented here (see Fig. 1).

Compartmental models. Well suited for systems where a time-persistent homogeneous distribution of molecular species within cellular compartments can be assumed, compartmental model geometries consist of a set of interconnected non-spatially resolved compartments: cellular structures (volumes

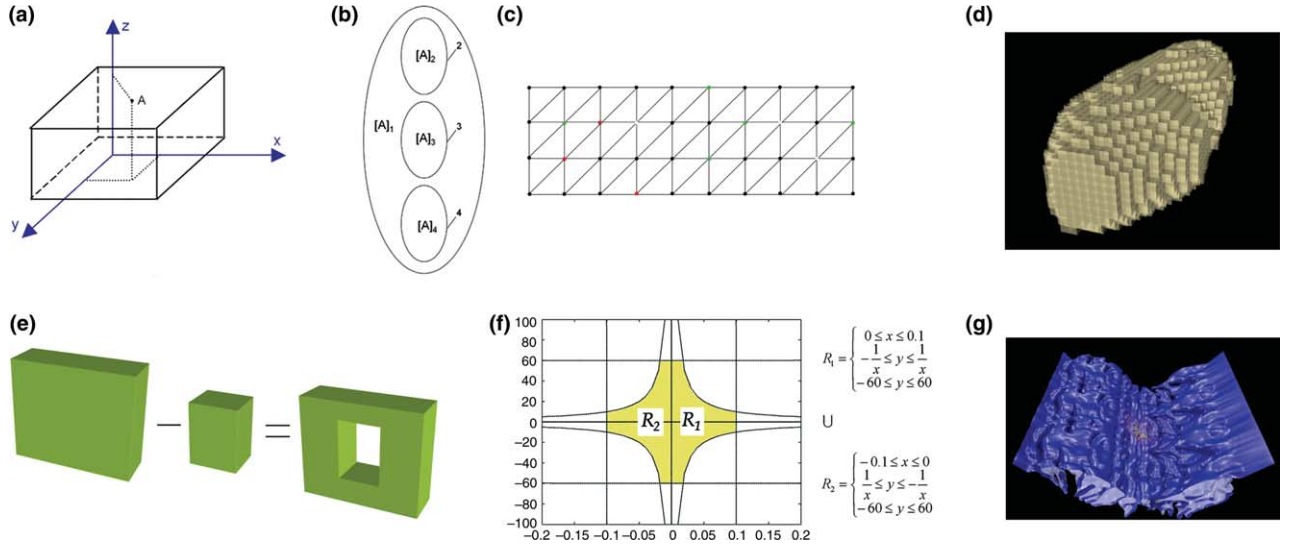

Fig. 1. Different methods for handling space in modeling and simulation of biochemical networks. (a) Continuous space bounded by walls: the geometry consists of a simple box, placed in a system of coordinates, which delimits the simulation space (Smoldyn). (b) Compartmental geometry: space is subdivided into containers, corresponding to cellular compartments, with no geometrical properties. Inclusion relationships between compartments have to be specified (Virtual Cell, E-Cell). (c) 2D lattice: surfaces are represented as two-dimensional grids, which can be made of squares, triangles or hexagons, where molecular entities are placed at grid points (STOCHSIM). (d) Voxelization: space is discretized into voxels, whose faces and edges are used to represent, respectively, 2D and 1D cellular structures (SmartCell). (e) CSG (constructive solid geometry): complex geometries are built by applying predefined operations to elementary objects called geometric primitives (MesoRD). (f) Analytical formulae: the geometry is described in terms of analytical formulae (Virtual Cell). (g) Surface reconstruction: the geometry is defined by surfaces (reconstructed from images) which can constitute the boundaries of 3D compartments or can be open structures (MCell) Schemes e, f, g are independent from space discretization. Molecular entities inside the geometry can be followed either using real-valued coordinates (MCell, Virtual Cell) or assigning them to voxels (MesoRD).

without shape) with specific molecular compositions and reactivities, and surfaces to support fluxes across them (Virtual Cell). This scheme is the least computationally expensive, but as a drawback from not handling diffusion it does not allow for the formation of gradients or heterogeneous distributions inside compartments.

1-3D spatially resolved models. Higher levels of spatial resolution (at a higher computing cost) are necessary if the ability to represent spatial heterogeneity of reactive species is an essential feature of the model [21], for instance, to achieve a concentration gradient within a cellular compartment.

(a) Realistic models. Experimentally derived images can be used to provide a realistic description of the position and shape of cellular structures. Schematically, construction of a model geometry from such sampled data is done following either a volume- or surface-oriented procedure, ending with an automatic or manual annotation of the extracted features. In the volume-oriented procedure, the geometric objects in the image are converted into a set of voxels that best approximates the continuous object (SmartCell). In the surface-oriented procedure, contours can be extracted from images, yielding sets of edges used to reconstruct filaments and surfaces (as mosaics of connected polygons for instance) defining the boundaries of 3D compartments (DReAMM for MCell).

(b) Idealized models. In the absence of input data, various methods exist to build a geometry from scratch. Volumetric datasets can be created where each 3D compartment is represented by a set of voxels from the volume buffer. The faces of the voxels at the interface between two volumetric compartments are 2D cellular structures, while 1D structures can be defined as a series of connected edges (SmartCell). In another volume-oriented approach, constructive solid geometry (CSG, commonly used in computer-aided design), the model geometry is constructed from a limited set of elementary volumetric objects (the geometric primitives) to which various geometric transformations can be applied (rotation, substraction, etc.). Alternatively, an intuitive and concise method is to specify the geometry analytically, mapping each compartment to the set of points in continuous space whose coordinates are solutions of a system of analytical expressions.

4.2. Reactive species

Depending on the spatial scheme used in the model geometry, three levels of resolution are available to represent reactive species: microscopic (individual molecular entities occupying unique positions, generally in continuous space), mesoscopic (populations localized in discrete space, composed of undistinguishable molecular entities), and macroscopic (real-valued densities defined in continuous space). The appropriate choice depends on various considerations. For instance, densities fail to make sense when representing small numbers of molecules, while models with discrete representations get slower as numbers are larger. With reactive species that can undergo multiple minor modifications affecting their reactivity (through conformational changes or small covalent modifications), microscopic representations are most suitable as the state of each molecule can be tracked (using flags to specify the state methylation at given sites, for instance) (STOCHSIM). However, all three schemes handle poorly cases where many species can form higher-order complexes in potentially huge numbers of combinations (the so-called ''combinatorial explosion''), leading to models with an unmanageably high number of species and reactions to work with (unless, as suggested in a recent paper, [22], these can be created dynamically, providing a law to

calculate the rates of newly created reactions). In very extended models, simplified naming schemes to identify species can be complemented with user-entered annotations or even links to external database entries (Virtual Cell), which can be used to retrieve information about the species (reactions, but potentially also parameter values). Additionally, important features assigned to species are those used in the mathematical model: diffusion and permeability rates, charge (for instance when membrane polarity is considered), and information related to the reactions they take part in.

4.3. Reactions

In principle, reactions in a spatio-temporal model are specified similarly to non-spatial models, only that they can now map to specific cellular structures, or even to specific locations within these structures.

Unlike continuous descriptions, which consider space as an independent variable, stochastic discrete-particle schemes can dynamically reduce the effective number of mono- (respectively bi-)molecular reactions to consider by conditioning the occurrence of such reactions to the presence (respectively encounter) of molecules, such that the time-dependent distribution of molecules determines which reactions have non-zero probability to occur (especially useful when small numbers of molecules are scattered over space). In discretized space, the criterion of physical proximity for bimolecular reactions reduces to considering identical positions (or neighboring ones if the reaction involves molecules from different compartments reacting at the interface, like a ligand binding to a membrane receptor). Reaction propensities (probability to occur per unit time) are then calculated as a function of (macroscopic) reaction rates and, when relevant, the population sizes of reacting species (SmartCell). In continuous space, collisions are first detected (by comparing the distance between two molecules with the sum of their respective radii in a hard-sphere model), the productiveness of which is then decided probabilistically (MCell). Alternatively, these two steps can be combined into one by providing a binding radius for each type of reacting pairs of molecules that incorporates the likelihood that a given collision between the two is actually productive (Smoldyn).

4.4. Movement

Tractable models of cell biological processes treat mainly three distinct types of movement: bounded diffusion, 1D transport, and membrane crossing.

Diffusive movement in bounded space. In discrete-particle schemes, the diffusive movement of single molecules is represented by finite displacements, representing many non-productive collisions. In continuous space, displacements have arbitrary orientations and lengths that follow the statistical law of Einstein–Smoluchowski (Smoldyn, MCell), and in discretized space, jumps are executed to neighboring spatial elements (mesoRD, SmartCell), corrected by boundary conditions whenever they apply. The calculation of diffusive jump rates becomes problematic in complex geometric schemes, for instance those coupling a mesoscopic representation of particles to compartment boundaries defined in continuous space [23].

1D directed movement. The directed movement of molecules across cellular space, whether of cargo molecules bound to molecular motors on microtubule filaments during transport, or RNA polymerase interacting with DNA during transcription, presents a representational difficulty. Microtubules, for example, are highly dynamic structures of such considerable length that the generally used dimensionless point representation of species is highly inadequate, on the other hand dynamic 1D cellular structures are still widely unsupported (see Table 1). As a compromise, SmartCell can model the dynamics of microtubules by allowing a fluctuating set of reactive species to align in a static 1D structure, representing the microtubule. A recent model of actin networks pushing bacterial cells across eukaryotic host cells addresses this issue using dedicated software [24].

Membrane crossing. Unless one is modeling intracellular prokaryotic processes in simple model geometries (Smoldyn), the presence of multiple compartments can be modeled that imply the exchange of molecular species between them. Different levels of mechanistic detail are available to model the crossing of biological membranes, energy-dependent (active transport) or not (passive transport), depending on whether participating molecular species can or cannot be represented in the membranes. In the simplest model, the translocation of a particular species from a given compartment to a target compartment is represented by a flux term with a specific permeability constant, possibly asymmetric (to represent active transport). Alternatively, membrane transport can be represented as a set of complexation reactions occurring at the interface, between a membrane-located carrier (for active transport) or channel protein (for either active or passive transport), and molecular species on either side of the membrane. Depending on the rates assigned to each binding and unbinding reaction, the ratio of steady-state concentrations at either side of the membrane can reach any value (a value different from one would be a signature of active transport).

4.5. Math and algorithms

Beyond the description of a geometry, molecular species and all allowable movement and reaction types, a model becomes computable only once both a kinetic theory (e.g., Fick's laws for diffusion in the deterministic approach and the chemical master equation, CME, for the stochastic formulation) and an associated algorithmic implementation are decided upon.

In their use of differential equations, macroscopic models can in principle accept any arbitrary analytical expression to represent the effects of reactions and mobility of molecular species on the time-evolution of state variables. Most software tools generally offer standard terms to choose from, like flux terms for membrane crossing, Fick's laws for diffusion, the law of mass action (LMA) as well as other various empirical expressions (E-Cell, Virtual Cell).

For models representing discrete molecular entities, a stochastic simulation algorithm (SSA) will generally be implemented accounting for events of movement and elementary chemical reactions (usually up to bimolecular) [25,26].

Different algorithms vary in their applicability and the user is expected to choose an appropriate one depending on multiple criteria: system being modeled, expected stochastic effects, modeling purpose, also available computer power [27].

A central yet non-trivial issue is whether a stochastic discrete approach or a continuous one (or a combination of the two) should be used for a given system [26,28–30]. Arguments supporting stochastic treatments in general are based on its real-

ism (molecular events are probabilistic by nature), and on the breakdown of the concept of concentration when few molecules are present, as occurs in numerous biological examples. A third argument often made a posteriori (because analytical approaches to the study of stochastic systems are far more limited than running simulations [31]), is that discrete random events can lead to phenotypic differences which continuous schemes cannot capture [32–34]. While deterministic simulations can at best provide the average behavior observed in stochastic runs (these often coincide, when the same macroscopic rates are used), stochastic continuous schemes can yield higher order statistics, while single realizations can only be studied using discrete stochastic algorithms. Stochastic mesoscopic models have an important limitation however: they commonly apply spatial variants of Gillespie's exact algorithm [27] which are valid provided that each voxel is "well-mixed", i.e., molecules assigned to that voxel can be considered to undergo few productive collisions relative to non-productive ones (a realistic assumption for cellular environments). This imposes that the timescale for diffusion events is smaller than that for reaction events, setting a lower limit to voxel size. This could result either in a too large number of voxels for simulation or, in the extreme case if the voxels become too small, this yields simulation results whose accuracy is difficult to evaluate [28]).

Continuous systems make sense when dealing with a very large number of particles for all reactive species, or when considering average population behavior.

4.6. Parameters

When modeling a system ab initio, all parameter values need to be retrieved (if not, constrained or guessed) from the various data repositories available: personal data, literature and databases, either manually or computer-aided (Virtual Cell). Since even a very good model can give aberrant results if unrealistic values are used, their quality, assessed constraining the value by considering experimental techniques and conditions used, should when possible be compared to the results of a parameter sensitivity analysis (Virtual Cell), giving an idea of the reliability of the observed behavior. Additionally, if experimental data or optimization functions are available, parameter estimation techniques can be applied.

The availability of parameters can condition modeling choices, for instance in modeling a Michaelis–Menten type of reaction, a continuous kinetic model will have two degrees of freedom (K_m and V_{max} in the lumped term representing the reaction), while a discrete scheme might decompose the reaction into three, each requiring its own rate.

Other problems can arise from the modelization of the reaction. For instance for two proteins making a complex, the outcome of the simulation will change dramatically if we consider degradation only for the isolated components or also for the complex [13]. Thus with similar parameters the outcome can be different depending on how the actual reactions are modeled.

4.7. Output and analysis

The end of the first step in the modeling cycle is the analysis of the simulation outputs, commonly provided in a wealth of different visualization formats (timeplots, snapshots, movies, etc.). This step offers a wide variety of choices, depending on the purpose sought by the user. Typically the simulation results will be analysed for a given behavior (e.g. the result of experimental observations), for instance, to validate the model [1]. If other models of the same system are available, results could be cross-checked between them. For directly comparing experimental data to simulation data automatic tools are sometimes available. If the construction of a model is a first step to building bigger models, a reduction task might be undertaken to rid the model of all unnecessary components. Other modeling purposes range from carrying out perturbation analysis, to studying system robustness or checking new hypotheses [2].

5. Conclusion

Stochastic spatio-temporal models are opening the door to simulating complex cell biological processes, bringing challenges to all disciplines involved in the design and use of software packages dedicated to that purpose.

State-of-the-art experimental tools are helping to lift the shortage of quantitative data available for modeling [35], but in order to assist modelers in the tedious task of collecting the data, curated databases and tools such as automatic data extraction from literature are being developed.

Addressing the structure of molecular species with mechanical functions such as the cytoskeleton and coatomers (molecules that curve and stabilize membranes), or macromolecules of highly dynamic and complex structures like DNA is a great representational challenge [24]. So is properly accounting for in vivo conditions (fluctuating, inhomogeneous medium with various obstacles) [16,17,36], where reaction dynamics might differ considerably from those measured in a test tube.

Dynamic geometries where compartments expand or contract are features under development in many software tools, and could be applied to modeling processes such as mitosis or cytokinesis, but changes in topology (for instance membrane budding or membrane fusion) remain a future theoretical challenge to be addressed [37].

While real cellular events happen concurrently, corresponding events simulated using conventional computers are processed (semi-)sequentially. Major ongoing efforts involve devising algorithms that reduce the number of computations by applying approximations to those reactions occurring at a faster time-scale [38–41], while adaptive algorithms are the next objective [42]. Alternatively, computer hardware has been redesigned dedicating single processors to single reactions, whereby decisions to execute events are taken at each clock cycle, resulting in considerable speed-up [43]. Benchmarking test suits to compare the performance of different simulation algorithms are now available [42].

For the user, the challenge is not only to decide how much of a system to model, and get familiarized with representation standards and terminology used in the software, but also to walk through the jungle of spatial representation and computation schemes available, avoiding modeling and numerical errors (and trust that programming errors are absent from the software). Finally, rationalizing the data analysis to let the modeling effort yield interesting insights is a challenge for all in the computational cell biology field [33,44].

While there is still room for improvement and a lot of caution should be taken when laying out a model, the time is ripe

for molecular biologists to explore the potential of stochastic spatial modeling.

Acknowledgement: This work was partly financed by the EU STREP Grant LSHG-CT-2004-503568 (COMBIO).

References

[1] Amonlirdviman, K., Khare, N.A., Tree, D.R., Chen, W.S., Axelrod, J.D. and Tomlin, C.J. (2005) Mathematical modeling of planar cell polarity to understand domineering nonautonomy. Science 307, 423–426.
[2] Arkin, A.P. (2001) Synthetic cell biology. Curr. Opin. Biotechnol. 12, 638–644.
[3] Tomita, M. (2001) Whole-cell simulation: a grand challenge of the 21st century. Trends Biotechnol. 19, 205–210.
[4] Available from: www.projectcybercell.ca.in.
[5] Eils, R. and Athale, C. (2003) Computational imaging in cell biology. J. Cell Biol. 161, 477–481.
[6] Van Helden, J., Naim, A., Mancuso, R., Eldridge, M., Wernisch, L., Gilbert, D. and Wodak, S.J. (2000) Representing and analysing molecular and cellular function using the computer. Biol. Chem. 381, 921–935.
[7] Schaff, J., Fink, C.C., Slepchenko, B., Carson, J.H. and Loew, L.M. (1997) A general computational framework for modeling cellular structure and function. Biophys. J. 73, 1135–1146.
[8] Tomita, M., Hashimoto, K., Takahashi, K., Shimizu, T.S., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J.C. and Hutchison, C.A. 3rd (1999) E-CELL: software environment for whole-cell simulation. Bioinformatics 15, 72–84.
[9] Stiles, J.R. (2000) in: Monte Carlo Methods for Simulating Realistic Synaptic Microphysiology Using MCell in Computational Neuroscience: Realistic Modeling for Experimentalists (De Schutter, E., Ed.), pp. 87–128, CRC Press, Boca Raton.
[10] Andrews, S.B. and Bray, D. (2004) Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. Phys. Biol. 1, 137–151.
[11] Elf, J., Doncic, A. and Ehrenberg, M. (2003) Mesoscopic reaction-diffusion in intracellular signaling. Paper presented at the SPIE's First International Symposium on fluctuations and noise.
[12] Le Novere, N. and Shimizu, T.S. (2001) STOCHSIM: modelling of stochastic biomolecular processes. Bioinformatics 17, 575–576.
[13] Ander, M., Beltrao, P., Di Ventura, B., Ferkinghoff-Borg, J., Foglierini, M., Kaplan, A., Lemerle, C., Toma's-Oliveira, I. and Serrano, L. (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.
[14] Vilar, J.M., Guet, C.C. and Leibler, S. (2003) Modeling network dynamics: the lac operon, a case study. J Cell Biol. 161, 471–476.
[15] Endy, D. and Brent, R. (2001) Modelling cellular behaviour. Nature 409, 391–395.
[16] Ellis, R.J. (2001) Macromolecular crowding: obvious but underappreciated. Trends Biochem. Sci. 26, 597–604.
[17] Luby-Phelps, K. (2000) Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. Int. Rev. Cytol. 192, 189–221.
[18] Massar, J.P., Travers, M., Elhai, J. and Shrager, J. (2005) BioLingua: a programmable knowledge environment for biologists. Bioinformatics 21, 199–207.
[19] Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., Arkin, A.P., Bornstein, B.J., Bray, D., Cornish-Bowden, A., Cuellar, A.A., Dronov, S., Gilles, E.D., Ginkel, M., Gor, V., Goryanin II, Hedley, W.J., Hodgman, T.C., Hofmeyr, J.H., Hunter, P.J., Juty, N.S., Kasberger, J.L., Kremling, A., Kummer, U., Le Novere, N., Loew, L.M., Lucio, D., Mendes, P., Minch, E., Mjolsness, E.D., Nakayama, Y., Nelson, M.R., Nielsen, P.F., Sakurada, T., Schaff, J.C., Shapiro, B.E., Shimizu, T.S., Spence, H.D., Stelling, J., Takahashi, K., Tomita, M., Wagner, J. and Wang, J. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19, 524–531.
[20] Lloyd, C.M., Halstead, M.D. and Nielsen, P.F. (2004) CellML: its future, present and past. Prog. Biophys. Mol Biol. 85, 433–450.
[21] Batada, N.N., Shepp, L.A. and Siegmund, D.O. (2004) Stochastic model of protein-protein interaction: why signaling proteins need to be colocalized. Proc. Natl. Acad. Sci. USA 101, 6445–6449.
[22] Lok, L. and Brent, R. (2005) Automatic generation of cellular reaction networks with Moleculizer 1.0. Nat. Biotechnol. 23, 131–136.
[23] Isaacson, S.A. and Peskin, C.S. (2004) Incorporating diffusion in complex geometries into stochastic chemical kinetics simulations, personal publication.
[24] Alberts J.B. and Odell, G.M. (2004) In silico reconstitution of *Listeria* propulsion exhibits nano-saltation. PLoS Biol. 2, e412.
[25] Schnell, S. and Turner, T.E. (2004) Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. Prog. Biophys. Mol. Biol. 85, 235–260.
[26] Turner, T.E., Schnell, S. and Burrage, K. (2004) Stochastic approaches for modelling in vivo reactions. Comput. Biol. Chem. 28, 165–178.
[27] Gibson, M.A. and Bruck, J. (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. J. Phys. Chem. 104, 1876–1889.
[28] Wolkenhauer, O., Ullah, M., Kolch, W. and Cho, K.H. (2004) Modeling and simulation of intracellular dynamics: choosing an appropriate framework. IEEE Trans. Nanobiosci. 3, 200–207.
[29] Srivastava, R., You, L., Summers, J. and Yin, J. (2002) Stochastic vs. deterministic modeling of intracellular viral kinetics. J. Theor. Biol. 218, 309–321.
[30] Gonze, D., Halloy, J. and Goldbeter, A. (2003) Deterministic and stochastic models for circadian rhythms. Pathol. Biol. (Paris) 51, 227–230.
[31] Orrell, D., Ramsey, S., de Atauri, P. and Bolouri, H. (2005) A method for estimating stochastic noise in large genetic regulatory networks. Bioinformatics 21, 208–217.
[32] Elowitz, M.B., Levine, A.J., Siggia, E.D. and Swain, P.S. (2002) Stochastic gene expression in a single cell. Science 297, 1183–1186.
[33] Paulsson, J. (2004) Summing up the noise in gene networks. Nature 427, 415–418.
[34] Rao, C.V., Wolf, D.M. and Arkin, A.P. (2002) Control, exploitation and tolerance of intracellular noise. Nature 420, 231–237.
[35] Sundararaj, S., Guo, A., Habibi-Nazhad, B., Rouani, M., Stothard, P., Ellison, M. and Wishart, D.S. (2004) The CyberCell Database (CCDB): a comprehensive, self-updating, relational database to coordinate and facilitate in silico modeling of Escherichia coli. Nucleic Acids Res. 32 (Database issue), D293–D295.
[36] Turner, T.E., Schnell, S. and Burrage, K. (2004) Stochastic approaches for modelling in vivo reactions. Comput. Biol. Chem. 28, 165–178.
[37] Giavitto, J.L. and Michel, O. (2003) Modeling the topological organization of cellular processes. Biosystems 70, 149–163.
[38] Rao, C.V. and Arkin, A.P. (2003) Stochastic chemical kinetics and the quasi-steady-state assumption: application to the Gillespie algorithm. J. Chem. Phys. 118, 4999–5010.
[39] Takahashi, K., Kaizu, K., Hu, B. and Tomita, M. (2004) A multi-algorithm, multi-timescale method for cell simulation. Bioinformatics 20, 538–546.
[40] Chavez, F. and Kapral, R. (2002) Oscillatory and chaotic dynamics in compartmentalized geometries. Phys. Rev. E: Stat. Nonlin. Soft Matter Phys. 65, 056203.
[41] Rousseau, G. and Kapral, R. (2000) Asynchronous algorithm for integration of reaction-diffusion equations for inhomogeneous excitable media. Chaos 10, 812–825.
[42] Vasudeva, K. and Bhalla, U.S. (2004) Adaptive stochastic-deterministic chemical kinetic simulations. Bioinformatics 20, 78–84.
[43] Salwinski, L. and Eisenberg, D. (2004) In silico simulation of biological network dynamics. Nat. Biotechnol. 22, 1017–1019.
[44] Pettinen, A., Aho, T., Smolander, O.P., Manninen, T., Saarinen, A., Taattola, K.L., Yli-Harja, O. and Linne, M.L. (2005) Simulation tools for biochemical networks: evaluation of performance and usability. Bioinformatics 21, 357–363.