



UvA-DARE (Digital Academic Repository)

Stochasticity in signal transduction pathways

Vidal Rodriguez, J.

Publication date
2009

[Link to publication](#)

Citation for published version (APA):

Vidal Rodriguez, J. (2009). *Stochasticity in signal transduction pathways*.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 6

Summarising discussion

Throughout this thesis we have made extensive use of explicit simulations in order to elucidate several relevant intracellular biochemical processes in prokaryotic cells. We started with the development of a general simulation tool to study the role of spatial localisation of reactions involving, in some reactions, a few number of molecules which make them susceptible of stochastic effects. We then focused on a model of two-component signal transduction pathway which captures the essential spatial and stochastic features and helps us to study which mechanisms affect the responsiveness to an external signal. Closely related to this topic is the study of the stochastic mechanisms of the expression of the related proteins involved in these signalling systems.

If we had to deal with the full details of biochemical reactions even for simple organisms such as prokaryotic cells, we would find ourselves with an intractable problem. We are then obliged to build simplified versions of the real system that capture the most relevant features to explain the phenomena under study. Yet we have to acknowledge that the results of a theoretical model can be sometimes hard to match to the real systems especially as the model becomes more extensive in the details and adding more variables. The range of methods to simulate biochemical processes is therefore broad. In chapter 2 we describe an approximate explicit method, GMP (Gillespie-Multiparticle), to simulate spatial and stochastic chemical processes where reactions can take place in the cytosol and also on the surface of the inner-membrane.

At the core of the method we split the two essential processes, namely the reaction and the diffusion, with the aim to reduce the complexity of the algorithm and the computational cost. The method operator-split method is inspired by the mechanisms underlying Cellular Automata for chemical reactions, which use simple rules in order to obtain an emergent behaviour. The split operator method allows us to use bulk diffusion of particles. In section 2.3 we use an alternative method based on pre-tabulated values for the exact distribution of the diffusion of few particles, namely the Binomial distribution. Unlike drawing Normal distribution numbers, which only holds for high number of particles and therefore a high concentration, drawing from the Binomial is expensive

(see Fig. 2.3). Using pre-tabulated values, for the few cases, helps avoiding the one-by-one method. Unfortunately, for parts of systems with even as much as few thousands of particles might be only possible to perform a one-by-one. As in the next-sub-volume methods, tracking the sites that change improves the performance by eliminating unnecessary checks of sites where no computations need to be executed. In these situation, the gain in performance is diminished, although qualitatively drawing uniform random numbers is more cost effective than drawing from a negative exponential distribution as used in the diffusion of particles in MesoRD.

In the comparison among several stochastic methods performed in section 3.1.2 we find that all methods converge in the dynamics of the average. Surprisingly, we find some disagreement in the noise levels between the Reaction Diffusion Master Equation-based models (GMP and MesoRD) and the Brownian Dynamic-based method Smoldyn. The culprit of such divergence is the reversible reaction, concretely the rebinding probability of a pair of molecules, for which the GFRD case is taken as a reference since it is the most physical-derived method of them all. This reaction, in the diffusion-limited case, has the most impact in the noise divergence observed.

The advantages of lattice-based methods rely on the simplicity of the method, especially that of reaction, compared to that of lattice-free methods such as GFRD and Smoldyn. One can argue against the less physical-based approach taken by lattice-based methods, and especially that of GMP, compared to GFRD. However, we find that using these methods with a reasonably chosen lattice size, as discussed in section 2.2.4, yields results that are in agreement with the reference method. As mentioned before, noise levels are the most sensible and lattice-dependent, but this behaviour is intrinsic to these methods because of the use of a reaction method at the lattice-site level that neglects the spatial phenomena.

The simplicity and generality of the lattice-based methods comes at the cost of accuracy. However, this does not render them useless and they are being used to study complex systems where the accuracy of the results might not play an important role and an approximate quantitative solution is enough to shed light on the process. In fact, measurements of the distributions presented in this thesis are difficult to measure directly. Feasible indirect measurements then are susceptible of a broad range of effects not accounted in the model. Such effects include anomalous diffusion due to crowding, and differences in the reaction mechanisms. All in all, these effects diminish the accuracy of simple models and a good approximation of the solution is often taken as valid.

Most of the analytical results, such as Langeving equations and Linear noise approximations, in fact, enable us to study homogeneous systems. Spatial analytical results are generally more difficult to obtain, and it is in this area that simulations fill this gap. Nonetheless, as in the signalling study presented in chapter 4, simulations aided us in developing analytical formulations of the problem. Such analytical treatment is possible because of the lack of interactions between membrane sensors, thus each sensor is an independent unit.

In chapters 4 we studied the influence that space and the number of molecules

have on the response time in two-component signal transduction pathways. We discussed how the spatial distribution of sensory histidine-kinases influence partly in the response time. When sensors are independent of each other scattered sensors reduce the phosphorylation time of response regulators. A detailed model of the signalling process accounts for the search process executed by each individual particle in the system. The order statistics framework, thus, allows us to calculate the first-passage activation (Eqn. 4.9). We showed that, in virtue of the double role in the phosphorylation and activation processes described in the simple model in Eqn. 4.8, the number of response regulators has a more significant impact in the reduction of the response time. Because of the higher weight of the response regulators, the response curve shows faster response time for ratios HK:RR smaller than 1. Even further altering and unbalancing the reaction ratios the optimal response time is always biased towards the right plane, for higher number of response regulators

We showed that when the system has an optimal ratio of HK:RR the system may work efficiently using fewer resources. Additionally, the total number of molecules need not to be large in order to obtain fast responses. With just 50 molecules, including HKs and RRs, the response time is one order of magnitude lower than the time necessary to synthesize a protein, or roughly the same as the initiation rate of translation. Addition of more particle, while keeping the optimal rate, reduce further the response time, however, the reduction achieved are linearly proportional with the inverse of the number of molecules, thus, only the first few particle achieve significant reductions.

In chapter 5 we analyse the expression mechanism responsible for the synthesis of the molecules necessary for two-component signalling systems. In bacteria, many genes are expressed from the same polycistronic mRNA. It is an efficient mechanism to ensure that functional related genes are expressed in similar ratios. This is of particular interest in two-component signalling system where we saw that the optimal point lies close to a 1:1 ratio. The stochasticity inherent in the transcription and translation processes has a small effect in deviating this ratio. Polycistronic mRNA with only one ribosome binding site before the first gene, may achieve this optimal ratio without the need of additional regulatory mechanisms. This together with the unusually unstable mRNA, resulting in an average half-lifetime of 36 s, produce typically tens of molecules. Despite the RR-HK ordering is under these circumstance favoured towards an optimal response time, the reverse ordering does not diverge much from the middle point.

However, we should realise that the number of active sensors and available response regulators may be just a fraction of those synthesized. Therefore, more knowledge about the particular signals detected and whether response regulators crosstalk with other systems should be taken into account in order to justify the most favourable ordering.

The literature on stochastic effects in cells is vast and has grown enormously in the last decade (see Fig. 6.1). Most of these recent efforts focus on the regulatory mechanisms of gene expression, as we did in section 3.1. This substantial increase also comes accompanied by the recent developments in microscopy technology and single cells experiments enabling detail levels that have been long

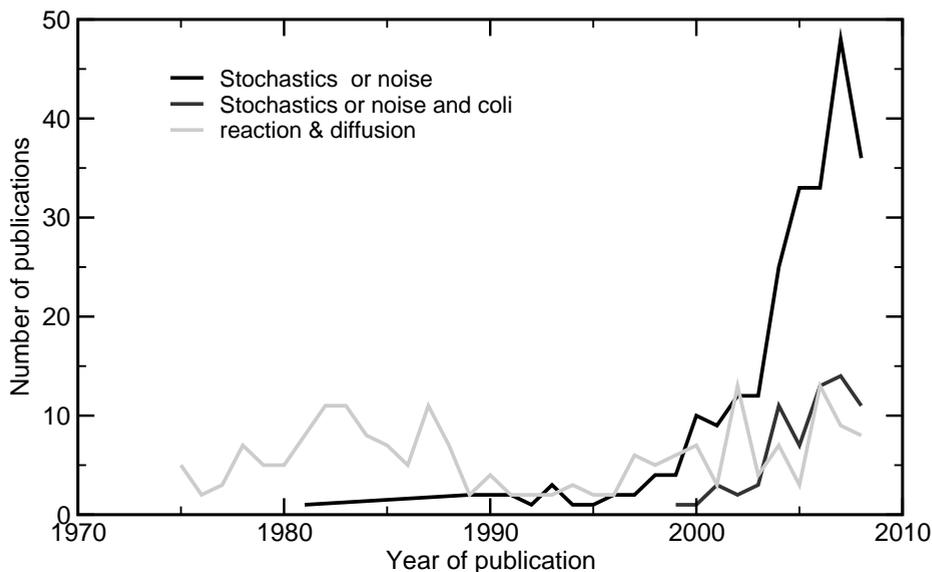


Figure 6.1: *Literature statistics for publications with key words in the title and in cell biology relevant publications. Source: <http://www.isiwebofknowledge.com/> as of October 2008.*

difficult to access. Although the beginning of reaction-diffusion methods date before this era of large interest in stochastic phenomena, there is a renewed interest in them, as we have done in this thesis.

Recapitulating, the first objective has been accomplished with the development of the Gillespie Multiparticle (GMP) method. We have shown that operator-split methods function well for simulating approximately cellular biochemical processes. As we have seen, comparable methods also have limitations in the accuracy of the results due to discretisation effects of the space into a lattice, whose size still remains difficult to estimate a priori.

Making use of the methodology developed for the first objective, we made extensive use of simulations to understand the phenomena of two-component signalling systems in bacteria to aid developing an analytical model and to characterise the signalling response time curve and its optimality. This curve suggested that bacterial organism could benefit if they had evolved a particular gene order that favours an optimal ratio of the number of molecules (histidine-kinases and response regulators) with minimal, or without, additional regulation mechanisms other than itself.

6.1 Future work

Computational methods for stochastic simulations

With ever more ambitious projects targeting whole-cell phenomena there is a need for methods that can cope with a broad variety occurring at multiple spatial, temporal and concentration scales. In this direction mesoscopic models may need to improve their performance in aspects related to diffusion-limited reactions. Increasing the lattice coarseness would offer faster computational, however, knowledge about the diffusion-limited reactions should be included in the reaction mechanism. Additionally, localisation of reactions suggest the use of irregular grids, which in addition could be made dynamic as to adjust to the system's varying conditions.

Computer parallelism is still lacking behind the massive truly parallelism found in physical systems. Inexpensive parallel computer system may enable a future generation of parallel algorithms that go beyond the limited spatial domain partitioning and include parallelism for reaction and diffusion at the particle level.

With the recent technological developments for *in vivo* measurements for single molecule phenomena and the increased temporal and spatial resolution it might be feasible to incorporate anomalous diffusion models into the simulation tools. Sub-diffusion has been observed in freely diffusive proteins in time scales similar to the size of the molecule. Large macromolecules, such a mRNA, have also been reported to have a sub-diffusive movement. Conversely, it has also been observed molecules with a super-diffusive behaviour, or active transport, using the cytoskeletal structure as a means of transport, much like RNA polymerase and Ribosome molecules perform walks on the DNA and mRNA, respectively.

Two-component signalling systems

In this thesis we have modelled the processes involved in two-component signalling systems with a generic approach in mind. We have highlighted a number of limitations in our modelling which mainly arise due to the rich number of properties and particular behaviour of the complex molecules involved in signalling and gene expression. To further our understanding the general models presented in chapter 5 should be adapted to specific cases, or more concrete classes of systems, and consider issues such as type of signal detected, signal strength, folding structure of proteins as well as secondary structure of mRNA. Also the dynamics of gene expression of operons and polycistronic mRNA which include stochastic phenomena such as translational coupling, processivity and degradation need to be understood in detail in order to shed light into the expression levels of the final proteins.

Apart from system specific problems, there are also questions about changes in the system in evolutionary scales. It also remains unanswered how two-component systems have evolved into their particular gene organisation in operons and their relative order within the operon. Have some systems originated

6. SUMMARISING DISCUSSION

with an operon architecture and due to evolutionary pressures have been separated into different operons? We have pointed out some reasons for when an operon organisation becomes inefficient and that individual regulation is required to obtain the adequate expression ratios between sensors and response regulators. Phylogenetic and bioinformatics analysis may hold the answer to these fundamental questions. However, some attempts to reconstruct phylogenetic trees have shown that this problem is hard when a large number of factors need to be taken into account.