

Modelling reaction kinetics inside cells

Ramon Grima † and Santiago Schnell ‡¹

† *Institute for Mathematical Sciences, Imperial College, London (E-mail: r.grima@imperial.ac.uk), ‡ *Indiana University School of Informatics and Biocomplexity Institute, 1900 E 10th St, Eigenmann Hall 906, Bloomington, IN 47406 (E-mail: schnell@indiana.edu).**

Abstract

In the past decade, advances in molecular biology such as the development of non-invasive single molecule imaging techniques have given us a window into the intricate biochemical activities that occur inside cells. In this article we review four distinct theoretical and simulation frameworks: (1) non-spatial and deterministic, (2) spatial and deterministic, (3) non-spatial and stochastic and (4) spatial and stochastic. Each framework can be suited to modelling and interpreting intracellular reaction kinetics. By estimating the fundamental length scales, one can roughly determine which models are best suited for the particular reaction pathway under study. We discuss differences in prediction between the four modelling methodologies. In particular we show that taking into account noise and space does not simply add quantitative predictive accuracy but may also lead to qualitatively different physiological predictions, unaccounted for by classical deterministic models.

¹ To whom correspondence should be addressed. **Present address:** Department of Molecular and Integrative Physiology and Center for Computational Medicine and Biology, University of Michigan Medical School, Palmer Commons, 100 Washtenaw Avenue, Ann Arbor, MI 48109-2218, USA.

Introduction

Over the past decade there has been an explosion of interest in methods for studying intracellular reactions, in particular those focusing on the determination of kinetic parameters [1]. Typically data obtained from experimental assays must be input into a mathematical model before it can be meaningfully interpreted. The bulk of these models are based on the macroscopic and deterministic models of classical physical chemistry [2]. A prominent example of this analysis is the characterization of enzymes using the Michaelis-Menten equation [3].

The use of any type of modelling methodology implicitly assumes that a number of physical, chemical or biochemical constraints are satisfied. Using a model outside of its range of application naturally implies that it is not possible to guarantee the accuracy of the model predictions. In this essay we discuss the four main modelling strategies in the context of intracellular kinetics. In particular we show how basic experimentally obtainable parameters can be used to roughly determine which model is appropriate for studying the reaction pathway under investigation. In section I, we discuss the homogeneity and continuum assumptions inherent in the use of classical biochemical models. We show how the validity of these assumptions can be checked by determining the fundamental length scales which characterize a given system. In section II, we briefly survey literature on various intracellular biochemical processes. We show that in a significant number of cases, it is not possible to uphold both the continuum and homogeneity assumptions at the same time. These cases are amenable to study via stochastic and discrete models, which are the topic of section III. In section IV, we discuss the effects of noise (stochastic vs. deterministic models) and

space (homogenous vs. heterogeneous models) on physiological predictions. We conclude with a brief discussion and summary in section V.

I. Classical models of chemical reaction kinetics

All reaction pathways can be decomposed into a series of elementary reaction steps, each of which is either unimolecular or bimolecular. By simulating each key step of a reaction, one can thus construct a model of the original pathway, independent of its complexity [4]. For these reasons we limit our discussion to modelling only these two types of elementary reactions.

Consider a chemical reaction occurring in a small cubic container with sides of length L . If the mixture is homogeneous throughout the container at all times - a condition which can be obtained by constant stirring - the state of the mixture at any point in time is fully described by the total number of particles of each molecular species in the container. Let n_I be the total number of particles of species I in the container; then the concentration of the species, denoted by $[I]$, is equal to $\frac{n_I}{L^3}$. For a reversible

bimolecular reaction, $A + B \xrightleftharpoons[k_{-1}]{k_1} C$, the forward rate of reaction (i.e. the number of molecular events leading to successful reaction per unit volume per unit time) is $k_1[A][B]$, whereas the backward rate of reaction is $k_{-1}[C]$. The rate equation (RE) for species C is then given by the familiar classical equation:

$$\frac{d[C]}{dt} = k_1[A][B] - k_{-1}[C]. \quad (1)$$

A crucial assumption implicit in REs is that the local concentration at each point in space inside the container must equal the global concentration at all times. In other words, all concentration gradients must be zero and, as a result, a non-spatial model is enough to describe the kinetics. This is called the homogeneity assumption. In industrial applications this is easily satisfied by constant (convective) stirring. Inside cells, homogeneity is only obtained if the effective diffusion coefficients are sufficiently large. The latter cannot be assumed *a priori*. The exact mathematical criterion for diffusion to suffice to maintain homogeneity inside our container is $D\tau \gg L^2$, where D is the molecular diffusion coefficient and τ is the average lifetime of a reactant molecule [5]. The distance $l_k = \sqrt{D\tau}$ is a measure of the distance travelled by a molecule during its lifetime. It is sometimes referred to as the Kuramoto length [5, 6]. If $l_k \gg L$, then any fluctuation in concentration at a particular point inside the container quickly diffuses to all parts of the container and homogeneity is maintained at all times. On the other hand, if $l_k \ll L$, the fluctuation remains localized in the region where it occurred and distant parts of the container will have different concentrations and, as a result, homogeneity cannot be assumed.

Estimation of the Kuramoto length depends on a determination of the molecular lifetime. We will illustrate how this calculation proceeds in the context of the following simple reaction pathway, $A \xrightarrow{k} C$, $C + C \xrightarrow{k'} B$, where species A is continually supplied (so that the concentration of A is constant at all times and equal to $[A_{const}]$) while species B is depleted. Suppose that the mixture of reactant molecules is initially perfectly homogeneous throughout the container and that the reaction is at steady-state such that the concentration of C is equal to $[C_{eq}]$. The average lifetime of a

molecule of type C can then be obtained by substituting $[A]=[A_{const}]$ and $[C]=[C_{eq}]+\delta[C]$ in the rate equation $\frac{d[C]}{dt}=k[A]-k'[C]^2$, where $\delta[C]$ is the fluctuation in concentration. By ignoring terms involving $\delta[C]^2$ (i.e. assuming fluctuations are very small) and solving for $\delta[C]$, one obtains the solution $\Delta[C]\propto e^{-t/\tau}$ where τ is the relaxation time given by $\tau=\frac{1}{2k'[C_{eq}]}$. This timescale is a measure of the lifetime of a molecule of species C and can, therefore, be used to estimate the Kuramoto length of species C [6]. Note that since τ is inversely proportional to the reaction rate, the Kuramoto length scale represents the relationship between diffusion and reaction rates.

In general, in order for homogeneity to be assumed, the Kuramoto length of each molecular species in the pathway must satisfy the restriction that it be much larger than the container size. If this is not the case, then one necessarily needs a mathematical description of the time evolution of the local concentrations -- in other words, a spatial model in which diffusion is explicitly described. This can be accomplished by dividing the volume of the container into many small cubic elements of linear size δL , such that in each element homogeneity prevails, i.e. $\delta L \ll l_k$. The rate of change in the concentration of species C in each element is due to (i) reaction between molecules of type A and B inside the element and (ii) due to diffusion of the same molecules of type C from neighbouring elements. Then it follows that the dynamics in each element is described by an equation of the same form as Equation (1) with an additional diffusion term:

$$\frac{\partial [C]}{\partial t} = k_1 [A][B] - k_{-1} [C] + D \nabla^2 [C] \quad (2)$$

This is called a reaction-diffusion equation (RDE). While REs and RDEs differ by whether the homogeneity assumption holds over the whole or parts of the volume in which the reaction occurs, they do share a common implicit assumption: both are deterministic approaches. The concentrations in these equations are mean concentrations and thus they can only be valid when the stochastic fluctuations about the mean are very small. In a cubic region of space of size L in which the average concentration is $[C]$, the ratio of concentration fluctuation size to the average is $r = \frac{L^{-3/2}}{\sqrt{[C]}}$. Now the average intermolecular distance is given by $l_i = [C]^{-1/3}$ from

which it follows that $r = \left(\frac{l_i}{L}\right)^{3/2}$. Hence, REs and RDEs implicitly assume that the reaction volume is much larger in size than the intermolecular separation, $l_i \ll L$ and $l_i \ll \delta L$. This assumption is known as the continuum assumption since it implies that the discreteness of molecules can be ignored.

Note that in our discussion above, we have assumed that molecules could reach any part of the container via diffusion. In this case it is clear that the size of the well mixed region of space can only depend on the magnitude of the diffusion coefficients and the molecular lifetime (i.e. the Kuramoto length scale). However, if physical structures exist inside the container which constrain molecular movement to some specific spatial regions characterized by a length scale L_s , then concentration homogeneity can only exist on a spatial scale δL if $\delta L \ll \min(L_s, L_k)$.

In summary, the classical approach is only valid if the reaction volume can be subdivided into spatial elements (of linear size δL) which are small enough to ensure well-mixed conditions inside each element (the homogeneity assumption: $\delta L \ll \min(L_s, L_k)$) but large enough to ensure that molecular discreteness can be ignored (the continuum assumption: $l_i \ll \delta L$). These are the two main assumptions implicit in the use of classical reaction kinetic equations. Diffusion-limited coalescence is a well studied case in which the above compromise is not possible (see [7] for a detailed discussion). We now turn to a discussion of whether both classical assumptions hold when modelling intracellular reaction kinetics.

II. Can we understand intracellular reactions using classical models?

Although there is limited information on the concentration and number of molecules inside cells, diverse experimental evidence (see below) suggests that many intracellular biochemical reactions involve nanomolar (nM) or smaller concentrations of reactant molecules [8-10]. In absolute number terms, 1 nM concentration roughly corresponds to 2500 molecules inside an average mammalian tissue culture cell (approximated as a sphere with a diameter equal to 20 microns), while 100 nM corresponds to 100 molecules inside an *E. coli* bacterial cell (approximated as a cylinder with one micron diameter and two microns in length). Absolute numbers of regulatory molecules have been reported to be approximately one thousand per cell for two endogenous yeast proteins [11], at most a few thousand per cell for proteins involved in the mammalian circadian clock [12], and to range from tens to hundreds per cell [13-15] for molecules in some bacteria. In contrast, only a few hundred are required to induce chemotaxis in amoeboid cells [8], to extend the growth cone in

chick dorsal root ganglion cells [16] or to induce an intracellular calcium response in HeLa cells [9].

A 1 nM concentration corresponds to an average intermolecular distance l_i approximately equal to a micron while 100 nM corresponds to l_i approximately equal to a quarter of a micron. Note that the magnitude of the intermolecular length scale is comparable to the typical size of many intracellular compartments inside cells (e.g., the size of a lysosome is 0.2 to 0.5 microns, the size of a mitochondrion is 0.5 to a micron and the size of a nucleus is 3 to 10 microns [17]).

It is important to remember that not all metabolites inside cells have concentrations in the nanomolar range. A number of glycolytic intermediates in the bacterium *L. lactis* are reported to be in the tens of millimolar range [18]. Such a concentration corresponds to an average intermolecular distance l_i approximately equal to 10 nanometres, a length scale much smaller than that of intracellular compartments.

The fact that nanomolar concentrations correspond to intermolecular distances comparable to the size of intracellular compartments has important implications for the validity of the continuum and homogeneity assumptions inside cells. The subdivision of eukaryotic cells into functionally distinct, membrane-bound compartments means that a cell is far from being a bag of well-mixed metabolites [19]. Compartments provide cells with functionally specialized aqueous spaces, separate from the cytosol. Indeed it is known that the enzymes of several metabolic pathways do not occur in the aqueous cytoplasm but are, instead, organized into structural and functional units, usually residing on membranes or on components of the cytoskeleton.

Some examples include scaffold and adaptor proteins (serving as docking sites for the assembly of various complexes at the plasma membrane) [20], enzymes involved in lipid metabolism [17] and enzymes involved in glycolysis [21]. (See Luby-Phelps [22] for a recent literature survey of this topic.)

Hence, the homogeneity assumption holds for length scales much smaller than both the typical size of an intracellular compartment (the latter length scale is equal to L_s) and than the molecular Kuramoto lengths. Since intermolecular distances are comparable to the size of intracellular compartments, the number of reactant molecules in spatial regions of this size will typically be so small that their discrete nature cannot be neglected and the continuum approximation breaks down. Hence it is unlikely that both assumptions can simultaneously hold and, consequently, the use of classical models appears inappropriate for modelling reactions with nanomolar concentrations.

It thus appears inevitable that a reasonable model of intracellular chemical kinetics will need to relax the requirement that the continuum approximation be valid for any length scale inside the cell along with the constraint that the homogeneity approximation hold for fine enough length scales. Such models are the subject of the next section.

III. Discrete and stochastic models of chemical reaction kinetics

Discrete models (i.e., models in which the temporal change in the density of particles in a region of space occurs by integer amounts), can be deterministic or stochastic. When fluctuations of the average number of particles in a given volume are small,

then a deterministic model (in terms of mean concentrations) will typically suffice. Otherwise, a stochastic model will be needed. In a cubic region of space of size δL in which the average number of particles is equal to n , the fluctuation about the average is of the order \sqrt{n} . A 1nM concentration in a 2 micron cubic box (representative of the volume occupied by a typical organelle in a mammalian cell) would imply approximately 5 ± 2 particles in the box. Since the fluctuation size is comparable to the mean, a stochastic description of discrete chemical reaction kinetics seems to be an unavoidable choice for intracellular processes where one or more of the metabolite concentrations are in the nanomolar range.

There are two types of discrete and stochastic formulations of chemical kinetics, paralleling exactly the classical RE and RDE approach discussed previously in this article. We will first discuss the case where reactants are homogeneous and then the case where they are heterogeneous in the container. In this case, the term container refers to the cell. The small elements of space into which it is subdivided are the intracellular compartments or, perhaps, even smaller spatial regions.

Stochastic kinetics in homogeneous conditions

Let us imagine a non-reversible bimolecular chemical reaction occurring in a container of size L between two different chemical species, $A + B \xrightarrow{k} C$. We assume that reactions occur in well-mixed conditions, i.e. $L \ll l_k$, but do not require that L is much larger than the typical intermolecular distance. We are, therefore, dealing with discrete (integer) numbers of molecules instead of continuous, real-valued concentrations.

Since we are assuming well-mixed conditions, the positions and velocities of individual molecules can be ignored and thus the state of the system at any time is completely specified by the number of molecules of both species (the number of molecules of type C can be ignored since they do not contribute to the reaction kinetics). Let $P(n_A, n_B, t)$ be the probability that in the container there are n_A molecules of species A and n_B molecules of species B at time t . The state vector is (n_A, n_B) . Applying the laws of probability yields the following equation describing the time evolution of the system:

$$\frac{d}{dt} P(n_A, n_B, t) = T(n_A, n_B | n_A + 1, n_B + 1) P(n_A + 1, n_B + 1, t) - T(n_A - 1, n_B - 1 | n_A, n_B) P(n_A, n_B, t), \quad (3)$$

where $T(n_A', n_B' | n_A, n_B) \Delta t$ gives the probability of a single bimolecular reaction event occurring in a time interval Δt and causing the state vector to change from (n_A, n_B) to (n_A', n_B') . The first term represents the state change $(n_A + 1, n_B + 1) \rightarrow (n_A, n_B)$ whereas the second term represents the state change $(n_A, n_B) \rightarrow (n_A - 1, n_B - 1)$. Equation (3) is called the Master Equation or the Chemical Master Equation (CME) for the bimolecular reaction. The function T is commonly referred to as the propensity function. For the reaction under consideration, T has to be proportional to both k (the familiar macroscopic rate constant) and to the number of A-B pairs in the reaction volume but inversely proportional to the container volume (since a larger volume implies rarer successful encounters). We, therefore, have $T(n_A', n_B' | n_A, n_B) = \frac{k}{L^3} n_A n_B$; this equation together with Equation (3) completes the discrete and stochastic

description of a bimolecular reaction. A similar procedure can be used to derive CMEs for any reaction process. A detailed discussion of the general derivation of propensity functions from microscopic physics can be found in [23].

CMEs can only be exactly solved in a few simple cases and, as a result, computer simulation is frequently essential. The most common simulation method is the stochastic simulation algorithm (SSA) of Gillespie [24, 25]. A discussion of this approach and its variants in the context of reactions inside cells can be found in the recent review by Turner et al [26]. There are several readily available stochastic simulation software packages for the simulation of CMEs via the Gillespie algorithm, e.g., Dizzy [27], BioNetS [28] and Dynetica [29].

Many of the current models assume well-mixed conditions throughout the whole volume of space in which the intracellular processes of interest occur. These models are built by decomposing the relevant reaction pathway into a series of unimolecular and bimolecular elementary reaction steps and then writing master equations for each step. Such a procedure has now been applied to study a significant number of intracellular processes, including genetic regulatory networks [30], enzymatic substrate cycles (a recurring control motif in biological molecular networks) [31], single enzyme kinetics [32, 33] and the circadian clocks of *Drosophila*, *Neurospora* [34] and of mammals [12].

We finish this section by noting that up till now we have assumed that the collision frequency is proportional to the product of the number of particle species; this is evident in both the form of Equation (1) and in the propensity function. If the

Maxwell velocity distribution is maintained [5, 25] then the above assumption is correct. When the dominant transport process is anomalous diffusion, the molecular velocities do not obey the Maxwell distribution; this is the case when there exists significant molecular crowding effects in the region of the cell where the reactions occur (three examples of this are protein motion inside cell membranes [35], in the nucleoplasm [36] and in the cytoplasm of HeLa cells and *E. coli* [37, 38]). The description of reaction kinetics in the presence of crowding requires more sophisticated approaches than the CME.

Stochastic kinetics in heterogeneous conditions

We now consider the case where the molecular Kuramoto lengths are much smaller than the container thus leading to significant spatial heterogeneity in the concentrations. To derive the stochastic equations governing such conditions, we adopt the same conceptual approach used previously to derive the deterministic reaction-diffusion equations with one important distinction: the continuum approximation is not made. We divide the reaction volume into m small cubes with side length δL , such that the reaction is well-mixed in each cube, i.e. $\delta L \ll l_k$. The reaction inside each cube can be treated in the same way as in the previous section. We must now, however, consider diffusion between neighbouring cubes (as was also the case for RDEs). Diffusion of species A between two adjacent cubes, i and j , can be represented as $A_i \xrightleftharpoons[k_D]{k_D} A_j$ where $k_D = D / z\delta L^2$ is the rate at which molecules diffuse between the two cubes (z is the number of nearest adjacent cubes). Hence, we see that diffusion is mathematically equivalent to a set of unimolecular processes. We can therefore describe both reaction and diffusion throughout space by writing the CME

for the set of reaction processes $\left(A_i + B_i \xrightarrow{k} C_i, A_i \xrightleftharpoons[k_D]{k_D} A_j, B_i \xrightleftharpoons[k_D]{k_D} B_j; i = 1..m \right)$. This is frequently termed a multivariate master equation [39] or a reaction-diffusion master equation (RDME). A thorough analytical treatment of these equations can be found in the recent paper by Isaacson and Peskin [40].

Efficient simulations of RDMEs in biological contexts can be achieved by extensions and variants of the Gillespie algorithm e.g. SmartCell [41] and MesoRD [42]. A crucial issue to be resolved before any spatially extended system can be simulated using RDMEs is the determination of the size δL of the small cubic elements in which homogeneity is assumed to prevail. The most straightforward way to ascertain the appropriate cube size to use for a given simulation requires the comparison of the frequency of molecular transfer (due to diffusion) between adjacent cubes, f_d , and the frequency of chemical reactions occurring inside the cubes, f_c . When $f_d \gg f_c$, homogeneity inside each cube is guaranteed; otherwise, a smaller cube size must be selected [43].

It is important to note that, in addition to the use of RDMEs (and their simulation via variations of the Gillespie algorithm), there are other mathematical and simulation frameworks available for modelling stochastic kinetics in heterogeneous conditions. Field-theoretic renormalization group methods (see [44] for a review) provide an alternative analytical approach whereas software based on Brownian dynamics (e.g., Smoldyn [45], Virtual Cell [46] and M-cell [47]) provides spatial stochastic simulation with single molecule detail (the latter feature is absent in software based on the Gillespie algorithm).

There have been only a few instances to date where spatial stochastic models have been used to model realistic intracellular kinetics; these include the bacterial chemotaxis signalling pathway [48], Min-protein oscillations in *E. coli* (see MesORD website <http://mesord.sourceforge.net/>), neurotransmission in the chick ciliary ganglion [49] and calcium waves in differentiated neuroblastoma cells [50].

In the above treatment we considered k_D to be a constant throughout space; by allowing this parameter to be a function of space one obtains a stochastic and discrete model of intracellular reactions and the complex intracellular environment in which they occur. For example, a permeable membrane could be modelled by requiring the value of k_D for each cube discretizing the membrane to be different than the value of k_D for each cube discretizing the surrounding cytosol. In this manner any number of complex cytoplasmic geometries may be easily modelled. We note that intracellular processes in which anomalous diffusion is the dominant transport process can be modelled with RDMEs by introducing a spatial distribution of impenetrable obstacles in the reaction volume (for a discussion of the relationship between anomalous diffusion and obstacle concentration see Saxton [51]).

IV. Homogeneous vs. Heterogeneous models. Stochastic vs.

Deterministic models.

As remarked in the previous section, the majority of current stochastic models of intracellular kinetics assume well-mixing in the intracellular region where the reactions occur. It is important to remember that that these models are simply an approximation of spatially-extended stochastic models and, as such, their predictions

can be expected to be physically realistic only when the Kuramoto length of each interacting molecular species is significantly larger than the cell diameter. If the rate coefficients and the approximate number of molecules are known then the above condition can be checked by explicit calculation using the formulas discussed in the beginning of this essay. Otherwise, one may experimentally tag key molecular species with green fluorescent protein and analyze the spatial distribution over a period of time.

In general, if the reaction scheme under consideration involves one or more diffusion-limited reactions, i.e., one in which the probability of reaction upon encounter is close to one (e.g., models of the kinetics of super-efficient enzymes), then the predictions of homogeneous models may be incorrect. In these cases reactant molecules move a small distance before reacting, implying that concentrations in different regions of space will evolve with time in a different manner. For example, homogeneous stochastic simulations of the diffusion-limited bimolecular reaction $A + B \xrightarrow{k} C$ with equal initial reactant concentrations via the Gillespie algorithm and analysis via the Van-Kampen expansion [52] show that the mean reactant concentration follows the classical decay law $[A] = [B] \propto 1/t$ (in all spatial dimensions) whereas spatially extended simulations [53] show that the correct long-time decay law is $[A] = [B] \propto t^{-d/4}$, where d is the dimension of space. Note that the reaction proceeds more slowly than predicted by the corresponding non-spatial models; this is because, during the course of the reaction, alternating domains of A-only and B-only particles are formed and the reaction can thus proceed only at the domain boundaries rather than throughout the whole volume of space, as assumed by non-spatial models

(Figure 1). A thorough discussion of this phenomenon can be found in Ben-Avraham and Havlin [7].

Note that discrepancies between the concentration decay laws predicted by non-spatial and spatial models increase with decreasing dimension. This has profound implications for the modelling of intracellular biochemical networks in which some of the reactions occur via an effective lower-dimensional motion. The use of combinations of three-dimensional (in the cytosol) and one or two-dimensional diffusion (e.g., gliding along a linear polymer or a membrane surface) is a mechanism first proposed by Adam and Delbruck [54]. This facilitates target location [55] and hence it is likely quite common inside cells, e.g., the one-dimensional gliding of the tetrameric *E. coli* lac repressor and RNA polymerase along the DNA chain to locate specific binding sites [56, 57].

Next we briefly discuss the predictions of stochastic and discrete models with those of the corresponding deterministic and continuous models for a given biological or biochemical system. Although it is now recognized that stochastic effects play important roles in cell biology, “they are often thought of as providing only moderate refinements to the behaviours otherwise predicted by the classical deterministic system description” ([31], p. 2310). Indeed there are many cases in which a deterministic description suffices to capture the approximate, qualitative physiological behaviour exhibited by a real biological system. Some examples include models of the budding yeast cell cycle [58], of the glycolysis pathway in yeast [59] and *L. lactis* [60] and of feedback regulation in the lactose operon [61]. However, there are also many instances in which it has been shown that stochastic effects lead to

physiological predictions which cannot be reproduced qualitatively or quantitatively by the corresponding deterministic models. Prominent examples of such cases occur in models of biochemical oscillations (e.g., circadian rhythms, calcium oscillations, and oscillations of ADP and ATP concentrations in glycolysis) where it has been shown that noise generally expands the region of parameter space in which cycles are predicted to occur by deterministic models (**Figure 2**) [12, 31, 62, 63]). In some cases noise may even induce oscillations in systems in which the corresponding deterministic model predicts that no oscillations are possible over the entire parameter space. For example, noise may induce cycles in reaction schemes involving just two chemical agents whereas deterministic models predict that at least three chemical agents are necessary [63]. In general, it has been found that fluctuations may lead to steady-states in chemical reaction-diffusion systems which are not predicted by RDEs [6].

V. Discussion

In this essay we have discussed the four main methodologies for modelling reaction kinetics in cells (**Figure 3**). The determination of which model is most relevant for the analysis of a particular intracellular reaction pathway is closely tied to the physical nature of the intracellular region in which it occurs and also with the typical chemical concentrations involved in the reaction.

In particular one needs to estimate the Kuramoto length scale, the size of the intracellular region in which reaction occurs and the average inter-molecular distances. These can be estimated from knowledge of the diffusion coefficients, the reaction constants and molecular concentrations. For example consider a reaction confined to

an intracellular compartment of micron size and in which the reacting species have a hydrodynamic radius equal to 4nm ($D \sim 0.05 \text{nm}^2 / \text{ns}$), a bimolecular rate constant equal to $k = 0.5 \times 10^6 \text{M}^{-1} \text{s}^{-1}$ and a typical concentrations of 1 μM . This implies $l_k \sim 10$ microns and $l_i \sim 0.1$ micron; since the compartment size is much smaller than l_k and significantly larger than l_i then a rate equation approach may suffice. If crowding caused the molecular diffusion coefficient to be reduced by five times, the typical concentration was 100nM and the rate constant was equal to $k = 5 \times 10^6 \text{M}^{-1} \text{s}^{-1}$ then $l_k \sim 4.5 \mu\text{m}$ and $l_i \sim 0.3 \mu\text{m}$. If this reaction forms part of a pathway which involves both the nucleus (3–10 μm) and mitochondrion (0.5–1 μm) of a mammalian cell then both homogeneity and continuum assumptions do not hold and a RDME approach will be needed.

Roughly speaking, reactions involving concentrations in the millimolar range will probably satisfy the continuum assumption whereas those involving smaller concentrations, such as nanomolar ones, will not. Reactions occurring exclusively in a small intracellular compartment will favour the homogeneity assumption; those involving multiple transport between several compartments and the cytosol or in regions where molecular crowding is significant will not satisfy this assumption.

As mentioned in the introduction, choosing the right model for the application at hand ensures the appropriate interpretation of experimental data. An example of this is the analysis of data obtained from single-molecule imaging techniques (SMI) [64-68]. The large fluctuations inherent in this data would necessarily have to be ignored if analysis proceeds via a deterministic framework. However, analysis via the stochastic

framework of master equations has shown that such fluctuations may contain important information about molecular behaviour; for example, it was recently shown that fluctuations in the rate constants of a single enzyme yield information regarding its conformational dynamics [33, 69].

We conclude by noting that in general it is not possible to *a priori* rule out the effects of noise or space on a given intracellular reaction scheme of interest. It is important to be aware of the fact that non-spatial deterministic models (based on classical REs) constitute only a first approach to modelling intracellular dynamics; their physiological predictions should be carefully checked against those of spatial stochastic models which correspond more closely to reality.

.Summary

- *Intracellular reaction kinetics can be modelled using four main different approaches (RE, RDE, CME, RDME); the right model choice for a particular reaction pathway under study depends on the concentrations and Kuramoto lengths of all molecular species, the size of intracellular space in which reaction occurs and the extent of macromolecular crowding in this region.*
- *In general, stochastic effects, typical of reactions involving nanomolar concentrations, do not simply add moderate quantitative refinements to the behaviours otherwise predicted by the classical deterministic models (RE, RDE). Noise may induce novel and qualitatively different physiological behaviour than that predicted by deterministic models.*

We are very grateful to Michelle L. Wynn for her critical comments during the revision of this essay. This work is based upon work supported by the NIH/NIGMS under Grant No. R01 GM076692. Any opinions, findings, and conclusions or recommendations expressed in this material do not necessarily reflect the views of the NIH.

References

- 1 Cornish-Bowden, A. and Cardenas, M. L. (2001) Enzyme kinetics. *Methods*. **24**, 95-96
- 2 Schnell, S. and Turner, T. E. (2004) Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Progress in Biophysics & Molecular Biology*. **85**, 235-260
- 3 Schnell, S. and Maini, P. K. (2003) A century of enzyme kinetics: Reliability of the K_M and v_{max} estimates. *Comments on Theoretical Biology*. **8**, 169-187
- 4 Crampin, E. J., Schnell, S. and McSharry, P. E. (2004) Mathematical and computational techniques to deduce complex biochemical reaction mechanisms. *Progress in Biophysics & Molecular Biology*. **86**, 77-112
- 5 Van Kampen, N. G. (2007) *Stochastic processes in physics and chemistry*. Elsevier, Amsterdam
- 6 Togashi, Y. and Kaneko, K. (2004) Molecular discreteness in reaction-diffusion systems yields steady states not seen in the continuum limit. *Physical Review E*. **70**, 020901
- 7 Ben-Avraham, D. and Havlin, S. (2000) *Diffusion and reaction in fractals and disordered systems*. Cambridge University Press, Cambridge
- 8 Ueda, M., Sako, Y., Tanaka, T., Devreotes, P. and Yanagida, T. (2001) Single-molecule analysis of chemotactic signaling in Dictyostelium cells. *Science*. **294**, 864-867
- 9 Uyemura, T., Takagi, H., Yanagida, T. and Sako, Y. (2005) Single-molecule analysis of epidermal growth factor signaling that leads to ultrasensitive calcium response. *Biophysical Journal*. **88**, 3720-3730
- 10 Rao, C. V. and Arkin, A. P. (2003) Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *Journal of Chemical Physics*. **118**, 4999-5010
- 11 Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W. and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences of the United States of America*. **100**, 6940-6945
- 12 Forger, D. B. and Peskin, C. S. (2005) Stochastic simulation of the mammalian circadian clock. *Proceedings of the National Academy of Sciences of the United States of America*. **102**, 321-324
- 13 Guptasarma, P. (1995) Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of Escherichia coli. *Bioessays*. **17**, 987-997
- 14 McAdams, H. H. and Arkin, A. (1999) It's a noisy business! Genetic regulation at the nanomolar scale. *Trends in Genetics*. **15**, 65-69
- 15 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
- 16 Tani, T., Miyamoto, Y., Fujimori, K. E., Taguchi, T., Yanagida, T., Sako, Y. and Harada, Y. (2005) Trafficking of a ligand-receptor complex on the growth cones as an essential step for the uptake of nerve growth factor at the distal end of the axon: A single-molecule analysis. *Journal of Neuroscience*. **25**, 2181-2191
- 17 Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002) *Molecular Biology of The Cell*. Garland Science, New York
- 18 Neves, A. R., Ramos, A., Nunes, M. C., Kleerebezem, M., Hugenholtz, J., de Vos, W. M., Almeida, J. and Santos, H. (1999) In vivo nuclear magnetic resonance

studies of glycolytic kinetics in *Lactococcus lactis*. *Biotechnology and Bioengineering*. **64**, 200-212

19 Clegg, J. S. (1984) Properties and metabolism of the aqueous cytoplasm and its boundaries. *American Journal of Physiology*. **246**, R133-R151

20 Pawson, T. and Scott, J. D. (1997) Signaling through scaffold, anchoring, and adaptor proteins. *Science*. **278**, 2075-2080

21 Pagliaro, L. (1993) Glycolysis revisited: A funny thing happened on the way to the Krebs cycle. *News in Physiological Sciences*. **8**, 219-223

22 Luby-Phelps, K. (2000) Cytoarchitecture and physical properties of cytoplasm: Volume, viscosity, diffusion, intracellular surface area. In *International Review of Cytology - A Survey of Cell Biology*, Vol 192. pp. 189-221

23 Gillespie, D. T. (1992) A rigorous derivation of the chemical master equation. *Physica A*. **188**, 404-425

24 Gillespie, D. T. (1977) Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*. **81**, 2340-2361

25 Gillespie, D. T. (2007) Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*. **58**, 35-55

26 Turner, T. E., Schnell, S. and Burrage, K. (2004) Stochastic approaches for modelling in vivo reactions. *Computational Biology and Chemistry*. **28**, 165-178

27 Ramsey, S., Orrell, D. and Bolouri, H. (2005) Dizzy: Stochastic simulation of large-scale genetic regulatory networks. *Journal of Bioinformatics and Computational Biology* **3**, 415-436

28 Adalsteinsson, D., McMillen, D. and Elston, T. C. (2004) Biochemical Network Stochastic Simulator (BioNetS): software for stochastic modeling of biochemical networks. *Bmc Bioinformatics*. **5**

29 You, L. C., Hoonlor, A. and Yin, J. (2003) Modeling biological systems using Dynetica - a simulator of dynamic networks. *Bioinformatics*. **19**, 435-436

30 McAdams, H. H. and Arkin, A. (1997) Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*. **94**, 814-819

31 Samoilov, M., Plyasunov, S. and Arkin, A. P. (2005) Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proceedings of the National Academy of Sciences of the United States of America*. **102**, 2310-2315

32 Stefanini, M. O., McKane, A. J. and Newman, T. J. (2005) Single enzyme pathways and substrate fluctuations. *Nonlinearity*. **18**, 1575-1595

33 Kou, S. C., Cherayil, B. J., Min, W., English, B. P. and Xie, X. S. (2005) Single-molecule Michaelis-Menten equations. *Journal of Physical Chemistry B*. **109**, 19068-19081

34 Gonze, D., Halloy, J. and Goldbeter, A. (2002) Robustness of circadian rhythms with respect to molecular noise. *Proceedings of the National Academy of Sciences of the United States of America*. **99**, 673-678

35 Saxton, M. J. and Jacobson, K. (1997) Single-particle tracking: Applications to membrane dynamics. *Annual Review of Biophysics and Biomolecular Structure*. **26**, 373-399

36 Wachsmuth, M., Waldeck, W. and Langowski, J. (2000) Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *Journal of Molecular Biology*. **298**, 677-689

- 37 Weiss, M., Elsner, M., Kartberg, F. and Nilsson, T. (2004) Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. *Biophysical Journal*. **87**, 3518-3524
- 38 Golding, I. and Cox, E. C. (2006) Physical nature of bacterial cytoplasm. *Physical Review Letters*. **96**, 098102
- 39 Gardiner, C. W. (2004) *Handbook of stochastic methods for physics, chemistry and the natural sciences*. Springer-Verlag, Berlin
- 40 Isaacson, S. A. and Peskin, C. S. (2006) Incorporating diffusion in complex geometries into stochastic chemical kinetics simulations. *Siam Journal on Scientific Computing*. **28**, 47-74
- 41 Ander, M., Beltrao, P., Di Ventura, B., Ferkinghoff-Borg, J., Foglierini, M., Kaplan, A., Lemerle, C., Tomas-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 (Stevenage)*. **1**, 129-138
- 42 Hattne, J., Fange, D. and Elf, J. (2005) Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics*. **21**, 2923-2924
- 43 Bernstein, D. (2005) Simulating mesoscopic reaction-diffusion systems using the Gillespie algorithm. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*. **71**, 041103
- 44 Tauber, U. C., Howard, M. and Vollmayr-Lee, B. P. (2005) Applications of field-theoretic renormalization group methods to reaction-diffusion problems. *Journal of Physics a-Mathematical and General*. **38**, R79-R131
- 45 Andrews, S. S. and Bray, D. (2004) Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Physical Biology*. **1**, 137-151
- 46 Loew, L. M. and Schaff, J. C. (2001) The Virtual Cell: a software environment for computational cell biology. *Trends in Biotechnology*. **19**, 401-406
- 47 Stiles, J. R., VanHelden, D., Bartol, T. M., Salpeter, E. E. and Salpeter, M. M. (1996) Miniature endplate current rise times $<100 \mu s$ from improved dual recordings can be modeled with passive acetylcholine diffusion from a synaptic vesicle. *Proceedings of the National Academy of Sciences of the United States of America*. **93**, 5747-5752
- 48 Shimizu, T. S., Aksenov, S. V. and Bray, D. (2003) A spatially extended stochastic model of the bacterial chemotaxis signalling pathway. *Journal of Molecular Biology*. **329**, 291-309
- 49 Coggan, J. S., Bartol, T. M., Esquenazi, E., Stiles, J. R., Lamont, S., Martone, M. E., Berg, D. K., Ellisman, M. H. and Sejnowski, T. J. (2005) Evidence for ectopic neurotransmission at a neuronal synapse. *Science*. **309**, 446-451
- 50 Fink, C. C., Slepchenko, B., Moraru, II, Watras, J., Schaff, J. C. and Loew, L. M. (2000) An image-based model of calcium waves in differentiated neuroblastoma cells. *Biophysical Journal*. **79**, 163-183
- 51 Saxton, M. J. (1994) Anomalous diffusion due to obstacles: A Monte Carlo study. *Biophysical Journal*. **66**, 394-401
- 52 Ben-Avraham, D. (1987) Discrete fluctuations and their influence on kinetics of reactions. *Journal of Statistical Physics*. **48**, 315-328
- 53 Toussaint, D. and Wilczek, F. (1983) Particle antiparticle annihilation in diffusive motion. *Journal of Chemical Physics*. **78**, 2642-2647

- 54 Adam, G. and Delbruck, M. (1968) Reduction of dimensionality in biological diffusion processes. In *Structural Chemistry and Molecular Biology* (Rich, A. and Davidson, N., eds.). pp. 198-215, W. H. Freeman & Company, San Francisco
- 55 Berg, O. G. and Von Hippel, P. H. (1985) Diffusion-controlled macromolecular interactions. *Annual Review of Biophysics and Biophysical Chemistry*. **14**, 131-160
- 56 Shimamoto, N. (1999) One-dimensional diffusion of proteins along DNA - Its biological and chemical significance revealed by single-molecule measurements. *Journal of Biological Chemistry*. **274**, 15293-15296
- 57 Wang, Y. M., Austin, R. H. and Cox, E. C. (2006) Single molecule measurements of repressor protein 1D diffusion on DNA. *Physical Review Letters*. **97**, 048302
- 58 Chen, K. C., Csikasz-Nagy, A., Gyorffy, B., Val, J., Novak, B. and Tyson, J. J. (2000) Kinetic analysis of a molecular model of the budding yeast cell cycle. *Molecular Biology of the Cell*. **11**, 369-391
- 59 Teusink, B., Passarge, J., Reijenga, C. A., Esgalhad, E., van der Weijden, C. C., Schepper, M., Walsh, M. C., Bakker, B. M., van Dam, K., Westerhoff, H. V. and Snoep, J. L. (2000) Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *European Journal of Biochemistry*. **267**, 5313-5329
- 60 Srividhya, J., Schnell, S., Crampin, E. J. and McSharry, P. E. (2007) Reconstructing biochemical pathways from time course data. *Proteomics*. **7**, 828-838
- 61 Yildirim, N. and Mackey, M. C. (2003) Feedback regulation in the lactose operon: A mathematical modeling study and comparison with experimental data. *Biophysical Journal*. **84**, 2841-2851
- 62 Orrell, D. and Bolouri, H. (2004) Control of internal and external noise in genetic regulatory networks. *Journal of Theoretical Biology*. **230**, 301-312
- 63 McKane, A. J., Nagy, J. D., Newman, T. J. and Stefanini, M. O. (2007) Amplified biochemical oscillations in cellular systems. *Journal of Statistical Physics*. **128**, 165-191
- 64 Sako, Y., Minoghchi, S. and Yanagida, T. (2000) Single-molecule imaging of EGFR signalling on the surface of living cells. *Nature Cell Biology*. **2**, 168-172
- 65 Schutz, G. J., Kada, G., Pastushenko, V. P. and Schindler, H. (2000) Properties of lipid microdomains in a muscle cell membrane visualized by single molecule microscopy. *Embo Journal*. **19**, 892-901
- 66 Sako, Y. (2006) Imaging single molecules in living cells for systems biology. *Molecular Systems Biology*
- 67 Xie, S. N. (2001) Single-molecule approach to enzymology. *Single Molecules*. **2**, 229-236
- 68 Watanabe, N. and Mitchison, T. J. (2002) Single-molecule speckle analysis of actin filament turnover in lamellipodia. *Science*. **295**, 1083-1086
- 69 English, B. P., Min, W., van Oijen, A. M., Lee, K. T., Luo, G., Sun, H., Cherayil, B. J., Kou, S. C. and Xie, X. S. (2006) Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nature Chemical Biology*. **2**, 87-94

Figure captions

Figure 1. Snapshot of the particle positions in two dimensional space, at a particular time, of two types of particles, A (white circles) and B (solid black circles), which are involved in a bimolecular reaction, $A+B \rightarrow C$. The simulation is spatial and stochastic. The segregation into A-rich and B-rich domains leads to a reduced reaction rate compared to the prediction of non-spatial models. (Reprinted with permission from Journal of Chemical Physics, Vol. 78, D. Toussaint and F. Wilczek, Particle-antiparticle annihilation in diffusive motion, pp. 2642-2647. Copyright 1983, American Institute of Physics.)

Figure 2. Comparison of the predictions of RE (A) and CME (B, C) models for the circadian oscillator in *Drosophila*. The parameter V (x-axis) is the gene transcription rate whereas M_p frequency (y-axis) is the frequency of the mRNA concentration cycles. The colour represents the log power of the frequency. Note that oscillations are paradoxically enhanced by noise; for low transcription rates, the stochastic models predict significant oscillations whereas the deterministic one predicts none. (Reprinted from Journal of Theoretical Biology, Vol. 230, D. Orrell and B. Hamid, Control of internal and external noise in genetic regulatory networks, pp. 301-312, 2004, with permission from Elsevier.)

Figure 3. Schematic showing the four modelling strategies and the conditions underlying their use and validity. The magnitude of the Kuramoto lengths and the intermolecular distances compared to the system size determines whether the homogeneity and continuum assumptions hold, respectively.

Figures

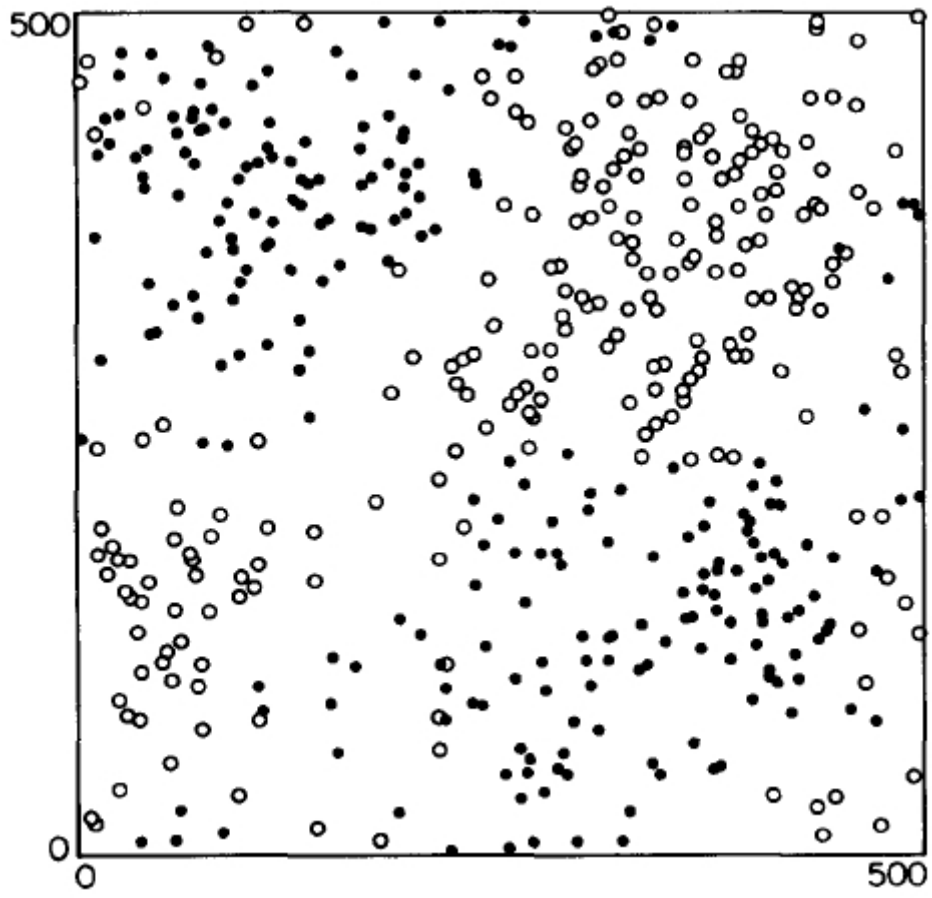


Figure 1.

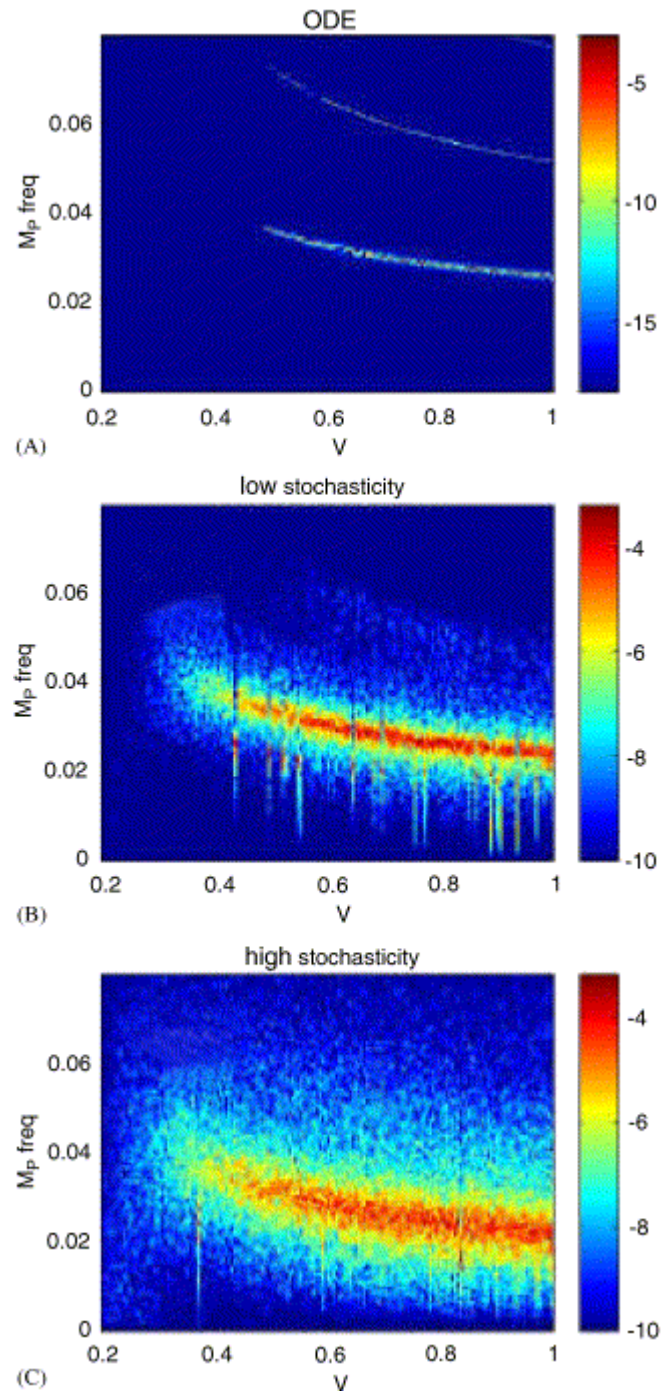


Figure 2.

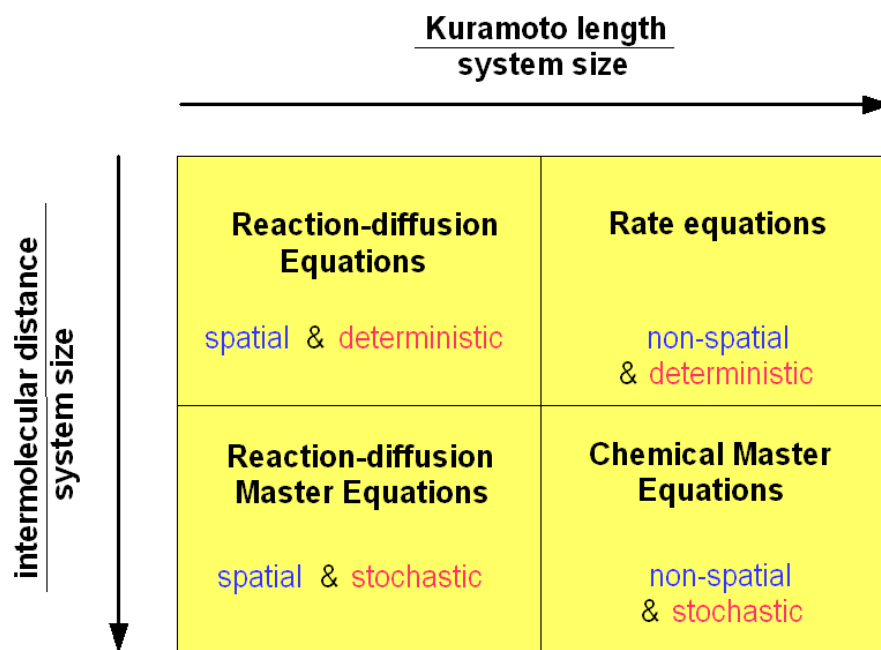


Figure 3.