

# Multi-Scale Spatio-Temporal Modeling: Lifelines of Microorganisms in Bioreactors and Tracking Molecules in Cells

Alexei Lapin, Michael Klann, and Matthias Reuss

**Abstract** Agent-based models are rigorous tools for simulating the interactions of individual entities, such as organisms or molecules within cells and assessing their effects on the dynamic behavior of the system as a whole. In context with bioprocess and biosystems engineering there are several interesting and important applications. This contribution aims at introducing this strategy with the aid of two examples characterized by striking distinctions in the scale of the individual entities and the mode of their interactions. In the first example a structured-segregated model is applied to travel along the lifelines of single cells in the environment of a three-dimensional turbulent field of a stirred bioreactor. The modeling approach is based on an Euler-Lagrange formulation of the system. The strategy permits one to account for the heterogeneity present in real reactors in both the fluid and cellular phases, respectively. The individual response of the cells to local variations in the extracellular concentrations is pictured by a dynamically structured model of the key reactions of the central metabolism. The approach permits analysis of the lifelines of individual cells in space and time.

The second application of the individual modeling approach deals with dynamic modeling of signal transduction pathways in individual cells. Usually signal transduction networks are portrayed as being wired together in a spatially defined manner. Living circuitry, however, is placed in highly malleable internal architecture. Creating a homogenous bag of molecules, a well-mixed system, the dynamic behavior of which is modeled with a set of ordinary differential equations is normally not valid. The dynamics of the MAP kinase and a steroid hormone pathway serve as examples to illustrate how single molecule tracking can be linked with the stochasticity of biochemical reactions, where diffusion and reaction occur in a probabilistic manner. The problem of hindered diffusion caused by macromolecular crowding is also taken into account.

---

M. Reuss (✉), A. Lapin, and M. Klann

Institute of Biochemical Engineering and Center Systems Biology, University of Stuttgart, Allmandring 31, 70569, Stuttgart, Germany  
e-mail: reuss@ibvt.uni-stuttgart.de

**Keywords** Agent-based modeling, CFD, Stirred tank bioreactor, Glycolysis of *E. coli* populations, PTS, Signal transduction, Stochastic differential equations, Fokker-Planck, Macromolecular crowding, MAPK cascade, Steroid hormone pathway

## Contents

|     |                                                                                                                                    |    |
|-----|------------------------------------------------------------------------------------------------------------------------------------|----|
| 1   | Introduction .....                                                                                                                 | 24 |
| 2   | Modeling the Dynamics of <i>E. coli</i> Populations in the Three-Dimensional<br>Turbulent Field of a Stirred Tank Bioreactor ..... | 25 |
| 3   | Stochastic Simulations of Four-Dimensional-Spatial Temporal<br>Dynamics of Signal Transduction Processes .....                     | 33 |
| 3.1 | The Multi-Scale Modeling Approach .....                                                                                            | 35 |
| 3.2 | Applications .....                                                                                                                 | 38 |
| 4   | Conclusion .....                                                                                                                   | 42 |
|     | References .....                                                                                                                   | 42 |

## 1 Introduction

Biochemical engineers are concerned with biosystems and bioprocesses. Traditionally a key focus of their activities has been the mathematical modeling in both territories. Most current models for bioreactors and biosystems are expressed as systems of nonlinear differential equations. Despite of the many benefits of such models, as well as their simplicity, they give us a rather simplified picture of the reality because of the lack of any structured and segregated details. In both applications the approach is based on the assumption of well-mixed system and biomass (unstructured or structured) as well as concentrations reflecting the intracellular states are considered as a continuum. At the same time, a major challenge is to apply more detailed approaches. Life is segregated into structural and functionally discrete entities – individual cells in heterogeneous populations of unicellular organisms [1] and the large number of molecules within the cells, which are interconnected in networks additionally characterized by molecular mobility. As a matter of fact, there is an increasing demand for modeling that captures more of the relevant complexity than the aforementioned systems of ordinary differential equations can achieve.

This contribution is aimed to illustrate the framework of individual-based modeling, which can capture the contingent nature of local interactions between individual molecules or cells. The modeling strategy will be applied to two examples with noticeable differences in scale. The first example addresses the issue of modeling the dynamics of *Escherichia coli* populations in the three-dimensional turbulent field of a stirred tank bioreactor based on a structured-segregated approach. The approach permits one to account for the heterogeneity present in real reactors in both the abiotic and biotic phases and thus deals with problems of nonideal mixing.

The second example focuses on at the discrete event based stochastic simulation applied to the four-dimensional-spatial temporal dynamics of signal transduction processes in individual cells. The modeling strategy comprises the random walk of individual molecules (diffusion) and the stochastic characteristics of the interaction of the partners of interest.

It must be emphasized that the application of this individual-based modeling framework is in no way restricted to the examples chosen for this contribution. There are many more applications with particular relevance to medical applications of systems biology, such as endocytosis, random walk of molecules in cancer tumors, lipoprotein kinetics in blood plasma, network of interacting neurons, and other interesting biological problems.

## **2 Modeling the Dynamics of *E. coli* Populations in the Three-Dimensional Turbulent Field of a Stirred Tank Bioreactor**

The physiological state of unicellular organisms in the highly dynamic environment of a bioreactor is the result of a complex interplay between the extracellular environment and the cellular machinery. The functionality of a biosystem for the purpose of bioproduction processes is therefore determined by the cooperative action of extracellular stimuli, the intracellular makeup and the dynamic response of the biological phase. The design of bioreactors in which living cells function as factories as well as the prediction of suitable operating conditions is further complicated because of the dynamic variations of the extracellular environment. Consequently, a quantitative description of these phenomena should rest upon the two interlinked aspects of structured bioprocess modeling: The first aspect concerns the complex interaction of the functional units within each cell, including the mathematical formulation of intracellular reaction rates and key regulatory responses of these networks to environmental changes. The second aspect involves the structure of the abiotic phases of the bioreactor for analysis of the impact of spatial and temporal variation in the intensities of mixing and mass transfer, leading to concentration gradients of various substrates and products. A modeling framework suited for capturing local variations in both intra- and extra-cellular concentrations should therefore rest on a link between metabolic network modeling or computational cell dynamics (CCD [2]) and computational fluid dynamics (CFD).

Previous attempts addressing this issue have been almost exclusively based on the Euler-Euler approach in which gas, liquid, and biophase are considered as a continuum [3–7]. It is only recently that more details have been included in modeling the dispersed phases. Most important are the class of models based on the Lagrangian-Euler approach, in which the liquid phase is treated as a continuum (Euler) and the dispersed phase is tracked with the aid of the Lagrangian representation. Important examples in the field of chemical engineering include the study of

gas–liquid flow in bubble column reactors [8, 9] and gas–liquid–solid three phase flow [10, 11].

In the case of bioreactors, however, modeling of the biophase most often follows the traditional approach, in which the microorganisms are lumped into a nonsegregated–non structured continuum. However, microorganisms are cellular in nature, and the continuum description is not rigorously correct. In fact, the continuum approach leads to a loss of realism if the individual history of the cells becomes the focus of attention, e.g., when considering cumulative starvation effects in cells during fed-batch fermentations or stability of plasmid containing microorganisms for production of recombinant proteins. In some instances, these problems can be tackled by combining the Euler approach for the fluid phases with population balance equations (PBEs) [12]. An inherent limitation of the PBE approach, however, is that the incorporation of a detailed intracellular reaction network leads to a computationally intractable model already for ideally mixed systems because a high-dimensional distribution function must be computed [13]. The complementary approach of combining CFD-simulations with population balances based on metabolically unstructured models of single cells faces similar practical limitations [12]. To overcome this problem, it is sometimes possible to resort to hybrid approaches combining multizonal models with CFD calculations, as reported by Bauer and Eigenberger [14, 15] for the case of gas–liquid bubble columns. Recently, Bezzo et al. [12] applied this strategy to xanthan gum production in stirred tanks. The authors combined an Eulerian description of the fluid flow with a multizonal model in which the reactor was divided into a limited number of spatial regions. These were assumed to be well-mixed and homogeneous and to be capable of material exchange with adjacent zones. Within each of these zones, a mass-structured population balance was formulated that was then combined with an unstructured kinetic model of substrate consumption and xanthan production.

Alternatively, one can account for the intracellular structure and dynamics of the cells while neglecting the spatial concentration gradients in the liquid phase. This corresponds to a Lagrange-type formulation of the equations for the corpuscular phase in an ideally mixed system [13, 16–18]. This description has the obvious limitation that bioprocesses are performed in real reactors where spatial variations in concentration usually cannot be neglected.

Detailed mathematical models capturing the variation in both the extracellular environment and the metabolism of the segregated biophase promise to aid significantly in describing the behavior of cell populations in bioreactors. This requires a combination of both approaches outlined above. For the first time this interaction between the intracellular state of the individual cells of the population and the turbulent flow field in the bioreactor has been tackled by Lapin et al. [19]. The chosen Euler-Lagrangian representation of the cell-ensemble approach permitted analysis of the lifeline of individual cells in space and time, as illustrated for the synchronization of autonomous glycolytic oscillations in yeast cells at the population level. With stirring conditions providing an environmental condition close to ideal mixing it was possible to predict the experimentally observed synchronization

of the individually autonomous oscillations at the population level. Simulation with reduced speed of agitation resulted in significant gradients of the extracellular attractor responsible for the synchronization (acetaldehyde). This leads to a dramatic loss of synchrony and eventually to almost complete desynchronization.

The work presented here picks up and summarizes the results of a second example, dealing with a problem of greater practical relevance, which, however, is also more complicated. The example, comprehensively worked out and described by Lapin et al. [20], concerns the population of the bacterium *E. coli*, which contains a phosphotransferase system (PTS) for the uptake of sugar.

The impact of extracellular gradients on biomass yield, byproduct formation and stress response of *E. coli* has been investigated in a couple of papers by the research group of Enfors [3, 21–25]. The results from measurement of glucose in a large scale bioreactor [3] during fed-batch cultures indicated that profound gradients exist which in turn give the cells an oscillatory pattern. In the case of simpler uptake systems, such as hexose transporters in yeast, a reasonable modeling approach can be based upon coupling unstructured kinetics for the biophase with various models for mixing within the abiotic phase [3, 6, 26]. In the case of *E. coli*, however, the uptake system leads to a situation in which the local uptake rate of glucose not only depends upon the locally different concentration of glucose in the tank but also upon the intracellular state which in turn may depend on the individual history of the cell.

The agent based models presented in [19, 20], which incorporate intracellular reactions, is based on an Euler-Lagrange simulation. Here, modeling of the extracellular environment is still based on the continuous Euler approach, whereas the behavior of the biophase is characterized by a discrete cell-ensemble approach (Lagrange). This allows each single-computational-“cell,” which still represents a large collective of real cells, to be endowed with its individual intracellular structure and state. In the modeling and simulation approach, mass transfer and reactions are assumed not to affect the turbulent flow field.

The three-dimensional turbulent flow for single- and two-phase systems has been simulated with the commercial software package PHOENICS employing a modified Chen-Kim  $k$ - $\varepsilon$  turbulence model [5, 6, 27]. The population of microorganisms is distributed over the reactor, which is subdivided into finite volumes.

The random movement caused by the turbulent dispersion is superimposed on the convective flow represented by the velocity field  $\vec{V}$ . In the Lagrangian approach, the position of a notional particle is governed by the stochastic differential equation [28]:

$$\vec{x}(t + \Delta t) = \vec{x}(t) + (\vec{V} + \nabla D_T)dt + (2D_T\Delta t)^{1/2}\vec{\xi}, \quad (1)$$

where  $\vec{x}(t + \Delta t)$  is the random position after a time step of  $\Delta t$ ,  $\xi_i$  signifies a Gaussian random number with a mean value zero and covariance  $\langle \xi_i \xi_j \rangle = \delta_{ij}$ , and  $D_T = f(k, \varepsilon)$  stands for the local eddy diffusivity calculated from the CFD simulation as a function of the turbulent kinetic energy  $k$  and the turbulent energy

dissipation rate  $\varepsilon$ . Slip can be neglected because the fluid velocity can be shown to surpass the slip velocity of microbial cells by several orders of magnitude. Momentum transfer between the particles and the fluid phase also does not require explicit consideration because the suspension can be treated as a quasi-single phase for particles smaller than the mesh spacing, as shown previously [9].

Thus, it can be assumed that the position of the notional particle predicted from (1) represents the behavior of the microbial cell along its trajectory in the turbulent flow field.

Prediction of the intracellular state of a single cell along the trajectory is performed by incorporating the system of intracellular balance equations into the model

$$\frac{d\vec{c}_{in,m}}{dt} = \mathbf{A}_m \vec{r}_m(\vec{c}_{in,m}(t), \vec{c}_{ex}(\vec{x}, t)). \quad (2)$$

Here,  $\vec{c}_{in,m}$  denotes the vector of intracellular metabolite concentrations in the individual cell  $m$  and  $\vec{c}_{ex}$  is the concentration vector of extracellular compounds at the position of this cell ( $\vec{x}$ ).  $\mathbf{A}_m$  stands for the stoichiometric matrix of the metabolic network in cell  $m$ , and the term  $\vec{r}_m$  represents its vector of intracellular reactions rates, which, in general, are nonlinear functions of  $\vec{c}_{in,m}$  and  $\vec{c}_{ex}$ .

Inclusion of  $\vec{c}_{ex}$  in the intracellular balance equations is required for the description of transport processes across the cell membrane, i.e., substrate uptake and product excretion. These occur in the cellular reaction rates  $\vec{r}_m$  and are also considered as a source term,  $\vec{S}$  (sink/source), in the Euler simulation of the extracellular state. Assuming that  $T$  of a total of  $R$  intracellular reaction rates represent transport rates across the cell membrane, then the system of extracellular state can be written as

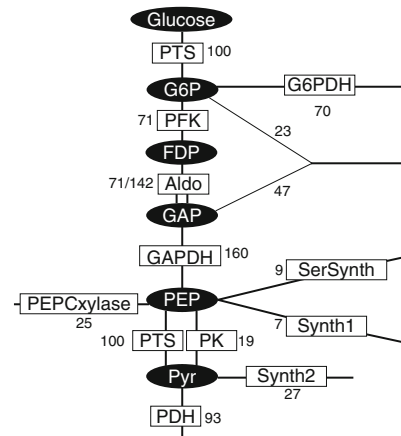
$$\frac{\partial \vec{c}_{ex}(\vec{x})}{\partial t} + (\vec{v} \nabla) \vec{c}_{ex}(\vec{x}) = \text{div}(D_T \text{grad } \vec{c}_{ex}(\vec{x})) + \vec{S}(\vec{x}) \quad (3)$$

with

$$\vec{S}(\vec{x}) = \frac{\varphi V_R}{N_C} \sum_{m=1}^{N_C} \mathbf{A}_{T,m} \vec{r}_{T,m}(\vec{c}_{in}(t), \vec{c}_{ex}(\vec{x}, t)) \delta(\vec{x} - \vec{x}_m) \quad (4)$$

where  $\vec{S}(\vec{x})$  equals the vector of net transport rates of all metabolites across the cellular membrane of all cells present at position  $\vec{x}$ . This is obtained by multiplying every addend of the sum over all  $N_C$  cells by Dirac's delta function, which signifies whether cell  $m$  (with position  $\vec{x}_m$ ) is present at  $\vec{x}$  or not. The subvector  $\vec{r}_{T,m}$  of  $\vec{r}_m$  contains only those  $T$  reaction rates accounting for transport across the cell membrane.  $\mathbf{A}_{T,m}$  represents the associated submatrix of  $\mathbf{A}_m$  consisting of columns that correspond to the transport reactions included in  $\vec{r}_{T,m}$ . The term preceding the sum reflects the influence of the total cell volume on the intensity of the exchange:  $\varphi$

**Fig. 1** Reduced metabolic network model for the sugar uptake system, glycolysis and pentose phosphate shunt. Reduction of the original model [29] is based on the hierarchy of flux control coefficients. The numbers alongside the enzymes depict the metabolic fluxes related to glucose uptake rate 100



denotes the ratio of the whole cellular volume  $N_C V_m$  and the reactor volume  $V_R$ , where  $V_m$  corresponds to the volume of a (simulated) single cell. Equation (3) accounts for convection and turbulent diffusion in the liquid phase and describes the coupling between the extracellular environment and the intracellular metabolism of the single cells.

Next we set-up the metabolically structured model, i.e., the system of intracellular balance equations introduced with (2) is substantiated. The selected biological example involves the sugar uptake system of the bacterium *E. coli*, a group translocation (PTS), where transport is associated with the phosphorylation of glucose to glucose-6-phosphate (G6P). The phosphate is donated by phosphoenolpyruvate (PEP), which is converted to pyruvate (PYR) (Fig. 1). Thus actual sugar uptake rate not only depends upon the local extracellular glucose concentration but also upon the concentrations of intracellular metabolites, which in turn are governed by the dynamics of the carbon and energy metabolism, essentially glycolysis and pentose phosphate shunt. The initial point of the model development is the dynamic model of Chassagnole et al. [29] that deals with the metabolic network of the central metabolism of *E. coli* wild-type strain W3110. It comprises modules for the PTS and the Emden–Meyerhoff–Parnas pathway providing PEP as well as PYR. The model also considers the pentosephosphate pathway which is linked to the glycolysis via fructose 6-phosphate (F6P) and glyceraldehyde 3-phosphate (GAP). The original model has been developed based on the measurement of metabolites in a continuous culture that has been perturbed by a glucose pulse. The model comprises 25 state variables and 30 kinetic rate expressions have been assigned. This model, of course, is inappropriate for the problem in hand as the computational effort (q.v. (1)) cannot be handled. It is therefore inevitable to perform a systematic model reduction to decrease the number of state variables.

A well proven concept for model reduction in metabolic engineering is based on the time hierarchy of the metabolism. The kernel of this method is a modal analysis, which considers the eigenvalues and eigenvectors of the Jacobian associated with the dynamic model [30]. The application of this time scale separation resulted in

assumptions of quasi steady state conditions for 11 eigenvectors possessing the highest values of  $\text{Re}(\lambda_i)$ . The result of this reduction, which shows reasonable agreement between the dynamic response of the original and reduced model, yielded, however, a differential-algebraic system. Because the algebraic equations did not allow an explicit analytical solution it is necessary to resort to an advanced and efficient solver for differential-algebraic systems. Thus, in context with the modeling concept developed in this chapter the decreased number of state variables would be compensated by the increased effort for the numerical solution.

As a promising alternative to the modal analysis we employed a sensitivity analysis based on the flux control coefficients. These coefficients relate the fractional change of the steady state fluxes to the infinitesimal changes in the total enzyme concentrations [30]. From the hierarchy of these flux control coefficients-predicted from the original model – the reactions with the highest values in relation to the flux control coefficient of the glucose uptake were selected. The resulting network is depicted in Fig. 1. Because of low flux control coefficients, the reactions for the phosphoglucose isomerase, the triose phosphate isomerase, the phosphoglycerate kinase, the phosphoglycomutase and the enolase could be neglected. By the same reasoning the glucose-6-phosphate dehydrogenase was selected as the rate determining enzyme for the pentose phosphate shunt. It is important to emphasize that the entire set of kinetic parameters identified from the measured intracellular metabolites [29] are the same as in the original model. The remaining set of balance equations, the kinetic rate expressions and the kinetic parameters are listed in original paper [20].

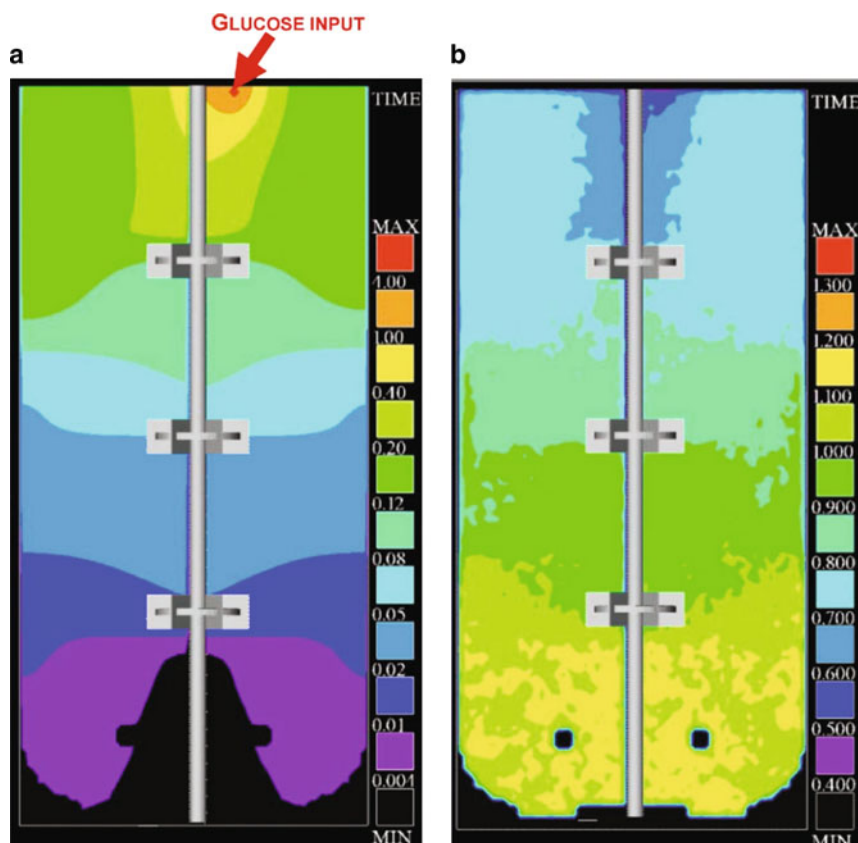
In context with the modeling task of linking the spatial variations of extracellular glucose with the dynamics of the individual cells it is important to emphasize that the system of balance equations for the intracellular state has been reduced to the feasible number of 5.

The simulations have been performed for a stirred bioreactor with 900 L operating volume, equipped with three six-bladed Rushton impellers (tank diameter 0.83 m, height 1.76 m, impeller diameter 0.33 m). The speed of agitation is 400 rpm. The number of control volumes is given by  $65 \times 65 \times 128$ .

The results from the simulation with 150,000 *E.coli* cells are displayed in Fig. 2. The combined model predicts distinct gradients in the extracellular glucose concentration (Fig. 2a). Glucose concentrations are highest at the top of the tank where the feed of concentrated glucose solution (concentration  $600 \text{ kg m}^{-3}$ ) is introduced. In the center of the bottom of the tank the glucose concentrations are close to zero, indicating strong limitations of sugar supply in this particular part of the tank. At first glance one should worry about serious starvation effects in case *E. coli* cells frequently enter this region. The pronounced gradients in the tank reflect the poor axial pumping and thus mixing capabilities of Rushton turbines.

In Fig. 2b the distribution of the ratio of the concentrations of the two intracellular key metabolites PEP and PYR are shown. The simple kinetic expression for the uptake of the sugar [20] indicate that this ratio makes a pivotal impact on the corresponding reaction rate. As evident from Fig. 2b the behavior of the ratio of PEP to PYR is reversed to the gradient of glucose – high values are observed at the

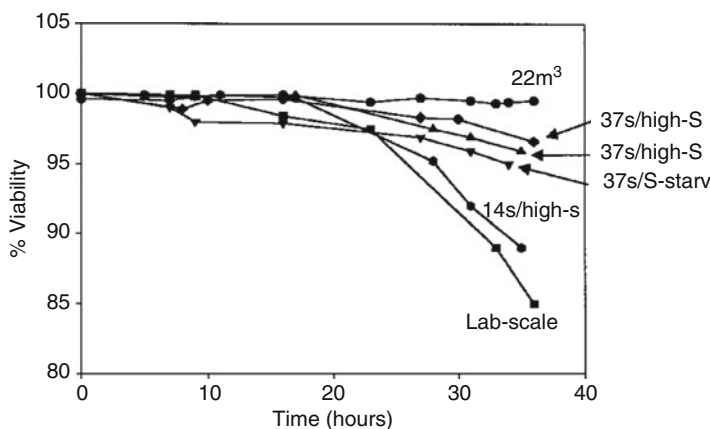




**Fig. 2** Concentration fields during a fed-batch process in a 900 L bioreactor equipped with three Rushton turbines. (a) Extracellular glucose, (b) distribution of the ratio of the intracellular concentration of phosphoenolpyruvate (PEP) and pyruvate (PYR)

top, low values at the bottom of the tank. The predicted results are in agreement with expectations. Due to the high glucose concentrations at the top of the tank, the glucose uptake rate is also high. Because the transport is associated with the phosphorylation, PEP immediately decreases and is converted to PEP. Once the cells enter the regions with lower glucose concentrations the delayed flux through glycolysis leads to a refill of the pool of PEP and readjustment of the concentration of PEP.

For further interpretation of these simulation results we refer to the experimental observations of Hewitt et al. [31–33]. These authors have been concerned about studying the influence of scale of cultivation and, as such, different intensity of mixing, on the viability of *E. coli* populations during fed-batch fermentations with constant feeding rates. For these purposes, multiparameter flow cytometry has been used. With the introduction of specific fluorescent dyes, valuable quantitative information on cell physiology and particularly viability could be obtained. The analysis revealed that a temporally varying environment with respect to



**Fig. 3** *E. coli* viability (measured with flow cytometry) in a lab-scale, a 22 m<sup>3</sup> scale and in a scale-down reactor with 14–37 s passages through a glucose starvation or glucose excess zone (reprinted from [32] during fed-batch operation with constant feeding rate)

glucose concentration has a profound effect on the viability of the cells. The comparison between a 22 m<sup>3</sup> and a 5 L scale demonstrates distinct differences in the cell viability (Fig. 3). Obviously the small-scale, well-mixed fermentation gave the lowest cell viability. The relatively poor mixed conditions in the large scale fermenter were found to lead to high cell viability. The reasoning regarding the positive influence of fluctuations in the microenvironment on cell viability were further supported with corresponding observations in an experimental set-up for scaled-down simulations. In these scaled-down experiments a small well-mixed stirred tank reactor (STR) was coupled to a plug flow reactor (PFR) via recycling flow. Highly concentrated glucose solution was fed into the STR and into the inlet of the plug flow. This set-up allows the successful simulation of poor mixing in large scale reactors. The same authors started to discuss possible reasons for these results in terms of environmental stress associated with the ever-increasing glucose limitation in the well-mixed case under conditions of constant feeding. It was furthermore argued that with the large scale and, depending on the residence time in the PFR also with the scale-down simulations, cells periodically encounter regions of relatively higher glucose concentrations.

It appears obvious to interpret our own simulation results further in the context of the aforementioned experimental observations. It appears beneficial to use the additional information about the distribution of intracellular state variables (PEP, PYR in Fig. 2b) to sustain the hypotheses regarding the impact of interlinked regulatory stress response pathways which was put forward by Hewitt et al. [33]. It is known that under conditions of glucose depletion, rapid increase of the intracellular “second messengers” cAMP and ppGpp is observed. cAMP is involved in the regulation processes related to the phenomenon of catabolite repression. The signal is built up by the enzyme adenylate cyclase which in turn is activated by one of the phosphorylated proteins in the PTS at conditions of sugar depletion and high ratios

of PEP/PYR. At natural conditions the signal is responsible for regulation phenomena leading to the expression of alternative sugar sources. The dynamic response of the PTS under conditions of substrate depletion is also linked to the chemotaxis allowing the bacteria to swim towards more favorable conditions e.g., higher glucose concentrations. In summary this dynamic response leading to the build up of the alarmone cAMP may be characterized by an offensive response of *E.coli* because constructive activities are mobilized to overcome the limitations in the carbon and energy limitations. In contrast to this offensive response the second alarmone ppGpp leads to regulation phenomena summarized with the term “stringent response,” which in a more defensive way reduces the energy and carbon demand by downregulating anabolic activities of the cell [34]. This phenomenon eventually leads to loss of viability via stress response (starvation). One of the key players in this stress response is the sigma factor ( $\sigma^S$ ). Interestingly, the two alarmones cAMP and ppGpp compete as transcription factors in the expression of the sigma factor [35].

Based on these molecular details regarding the link between metabolism and its regulation and the results of the simulations in Fig. 6 it is tempting to speculate on the following scenario within the large tank. Those cells traveling through the region of high sugar concentrations at the top of the tank respond with a corresponding high sugar uptake rate and a drop in the phosphorylation potential (PEP/PYR). At the same time as the cells are moving towards the bottom of the tank where they are exposed to extremely low sugar concentration, the ratio of PEP/PYR is very high. The dynamic response of the PTS system under these conditions should result in a fast increase of cAMP. Even under conditions of lasting sugar limitations the peak of cAMP would counteract the stringent response signal of ppGpp and thus would prevent the stress response. In the case of the small well-mixed reactor the sugar concentration would always remain at a very low level, the ppGpp could prevail and eventually pronounced stress response is initiated which also impacts cell viability by inducing the programmed death of the bacteria [36].

### 3 Stochastic Simulations of Four-Dimensional-Spatial Temporal Dynamics of Signal Transduction Processes

Traveling along lifelines of individual cells and populations in bioreactors is only one but nevertheless an important example of the application of the agent based or individual modeling in bioprocess and biosystems engineering. The issue of detailed quantitative modeling of spatiotemporal effects at the scale of individual cells or multicellular systems comprises a manifold of important problems. Important examples comprise chemotaxis and quorum sensing in bacteria, signal transduction pathways, endocytosis, phagocytosis as well as movement of drug molecules in complex tissues such as solid tumors, to name a few.

Many cellular signaling events occur in small subcellular volumes and involve low-abundance molecule species. This context introduces a major difference and additional complication compared to the bioreactor modeling illustrated above.

Reactions involving a low number of molecule species occur in a probabilistic manner. Thus, in addition to the random walk simulating the diffusive motion of the individual molecules the stochasticity of the reaction has to be taken into account. As such, the assembly approach presented in the aforementioned example is extended to a coupled reaction–diffusion process in which the individual agents change their characteristic properties through interactions.

Spatial aspects of cellular signaling have already been the subject of various experimental and theoretical investigations [37, 38]. Modeling the interaction of diffusion and reaction is most often based on the continuum approach, thus investigating the corresponding system of partial differential equations. The pioneering work of Kholodenko [38, 39] addressed for the first time the issue of spatially heterogeneous and time varying cellular signal transduction cascades. They developed computational models of the mitogenic signaling network to analyze the complex structure of the spatial distribution of the activated compound ERK. At a distance larger than several microns from the plasma membrane the phosphorylation signal is attenuated practically to basal levels, provided that the phosphatase activity in the cytosol is sufficiently high. These and similar investigations by Howe [40] are, however, restricted to a macroscopic, deterministic continuum approach, which neglects the random walk of individual molecules and the stochastic characteristics of the interaction of the partners of interest. These stochastic properties can only be modeled using Monte-Carlos simulations.

The most important approach for stochastic simulation of coupled chemical reactions trace back to the famous work of Gillespie [41], solving master equations which describe the evolution of the so-called grand probability for the number of molecules of the different species. This equation, however, is only valid for spatially homogenous mixtures, in other words, ideally well-mixed systems. There are a couple of attempts to couple diffusion problems to this master equation [42, 43]. A critical assessment of these approaches, in which the system is divided into small sub-volumes, demonstrates two important drawbacks. First, the sub-volume is assumed to be well-mixed and the diffusion processes are restricted to the boundaries of the grids. Second, the molecules are represented as point particles. Therefore, it is not possible to reproduce crowded conditions because volume exclusion from both reactive and nonreactive crowded molecules cannot be represented explicitly.

On the other hand microscopic models including all molecules of a cell as well as all interactions – leading to a complete molecular dynamics (MD) simulation – are computationally expensive and therefore unable to cover the dimensions of a complete cell [44]. Only mesoscopic models are until now able to span all necessary ranges to model signaling processes with sufficient detail. The underlying Smoluchowski model for diffusion according to Brownian motion and reactions in a reaction–diffusion system are implemented by different groups. Particles [45, 46] or so called agents [47] perform a random walk representing the diffusion process. Reactions will take place if the distance between reacting species falls below a predefined reaction-distance. While MCell [46] restricts reactions to fixed positions, all these models have the drawback that there is no physical property defining the fixed model-dependent reaction-distance.

The method presented in this chapter aims at studying the interconnected effects of signal transduction networks by a new multi-scale approach. The probability of collision and reaction between the interesting species is modeled by the solution of the Fokker–Planck equation providing the probability that two molecules will collide and react in the next interval  $\Delta t$ . The trajectories for the random walk of protein molecules are modeled by stochastic differential equations (Lagrangian approach). This allows incorporating the effects of macromolecular architecture of the cells and thus to investigate the hindered diffusion due to crowding caused by the cytoskeleton. The criterion for change on the model scale is a threshold value for the distance between the reaction partners. The new modeling approach will be exemplified for the RAS-MAPK pathway and a steroid hormone pathway.

### 3.1 The Multi-Scale Modeling Approach

#### 3.1.1 Random Walk of the Molecules

The random walk simulations are based on the numerical solution of the stochastic differential equation for single molecule tracking:

$$\vec{x}(t + dt) = \vec{x}(t) + (2Ddt)^{1/2} \vec{\xi} \quad (5)$$

with the three-dimensional position vector  $\vec{x}$ , the diffusion coefficient  $D$  and the Gaussian random number  $\xi_i$  with mean zero and covariance  $\langle \xi_i \xi_j \rangle = \delta_{ij}$

#### 3.1.2 Fokker–Planck Equation

Our interest is to study the coupling of this random walk of molecules with the probability of collisions with potential reaction partners and eventually the degree of the diffusion-limited reaction itself. For this purpose the random movements of both molecules (e.g., phosphorylated protein and phosphatase or steroid and steroid-receptor) are simulated according to the aforementioned strategy. If paths of interacting proteins come within a distance  $\varepsilon$  of each other, a change in the scale of modeling is performed by switching to the analytical solution of the Fokker–Planck equation for the probability density function (p.d.f.)  $P$  of diffusive movement given by:

$$\frac{\partial P}{\partial t} = (D_A + D_B) \Delta P \quad (6)$$

We only need to keep track of the magnitude of the difference of position [48] and hence one molecule can be considered to be in a fixed position whereas the second one diffuses through the three-dimensional space. The boundary conditions for (6) are given by the following considerations: The particles can not overlap: if

the particles reach the distance  $r = R_A + R_B$  either the reaction will occur (leading to a flux of the particles into each other) or the particles will be reflected [49]. This leads to the partially reflecting boundary condition

$$r \rightarrow \infty: \quad P = 0 \quad r = R_A + R_B: \quad (D_A + D_B) \frac{\partial P}{\partial r} = k' P \quad (\text{reaction})$$

with  $R_A, R_B$  radii of the two molecules,  $r$  radial coordinate and  $k'$  surface reaction constant. The relation with the macroscopic reaction constant for well-mixed systems can be derived easily:

$$k' = \frac{k_{\text{macro}}}{4\pi(R_A + R_B)D - k_{\text{macro}}}$$

(Dimensionless parameters and variables simplify the calculations:  $k_{\text{macro}} = k_{\text{macro}}(R_A + R_B)^2/D$ ;  $r = r/(R_A + R_B)$ ;  $t = tD/(R_A + R_B)^2$ ). Inert obstacles can be modeled with  $k' = 0$  leading to pure reflection.

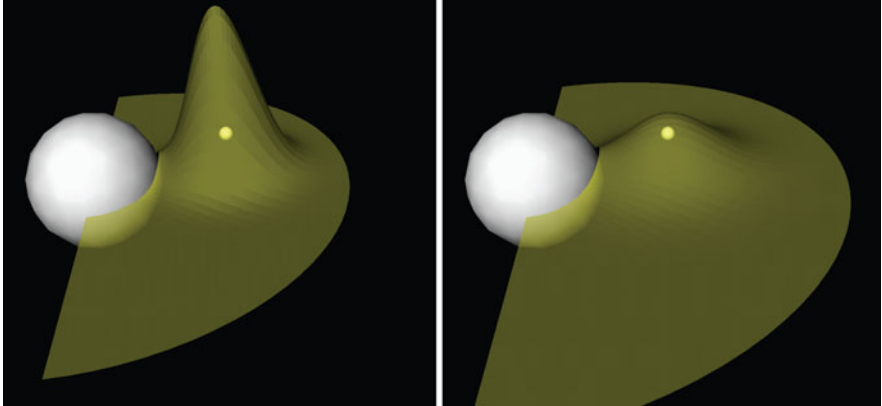
For the initial condition that the particles are separated by  $\vec{r}_0$  at  $t = 0$ , the initial p.d.f. is  $P = \delta(\vec{r} - \vec{r}_0)$ , and the analytical solution of (6) reads:

$$P = \int_0^\infty e^{-\lambda^2 t} \sum_{n=0}^\infty (n + 1/2) P_n(\cos \theta) R_{n\lambda}(r) R_{n\lambda}(r_0) d\lambda$$

$$R_{n\lambda} = \frac{L_{yn\lambda} j_n(\lambda r) - L_{jn\lambda} y_n(\lambda r)}{\sqrt{L_{jn\lambda}^2 + L_{yn\lambda}^2}} \quad (7)$$

$$L_{yn\lambda} = \lambda y'_n(\lambda) - k y_n(\lambda)$$

$$L_{jn\lambda} = \lambda j'_n(\lambda) - k j_n(\lambda)$$



**Fig. 4** Time development of the probability density function for the possible position of a particle after time  $\Delta t$  and  $2\Delta t$  (analytical solution of the Fokker–Planck equation). The initial position is marked with the small sphere, the large sphere represents the distance  $R_A + R_B$ . The function broadens due to diffusion and the overall probability is reduced by the reaction probability

where  $P_n$  is the Legendre polynomial and  $j_n$  and  $y_n$  are spherical Bessel functions. Equation (7) describes the time evolution of the p.d.f. of the relative position of the molecules. Figure 4 shows a graphical presentation of this solution for two snapshots in time.

Eventually the reaction probability has to be estimated. The calculation is based on the simple fact that the sum of the probabilities that a molecule is still present within the p.d.f.  $P$  and the probability that a molecule disappeared through reaction must be 1.

Position probability + reaction probability = 1, or

$$R_P = 1 - \int_{r \geq R_A + R_B} P(\vec{r}, t) d\vec{r}$$

$$R_P = \frac{k'}{k' + 1 + x_0} \times \left[ \operatorname{erfc}\left(\frac{x_0}{\sqrt{\tau 4}}\right) - \exp(x_0 + \tau) \operatorname{erfc}\left(\frac{x_0 + 2\tau}{\sqrt{\tau 4}}\right) \right] \quad (8)$$

again with dimensionless variables:  $x_0 = (r_0 - 1)(1 + k')$ ,  $\tau = (1 + k')^2 \Delta t$

A uniform random number  $\zeta (0 \leq \zeta \leq 1)$  is then used as an indicator of a successful reaction, such that:

$$(1) \text{ if } \zeta \leq R_P \longrightarrow \text{reaction} \quad (2) \text{ if } \zeta \geq R_P \longrightarrow \text{no reaction}$$

In the case of (2), the molecule continues moving on according to  $P(\vec{r}, t)$ . As soon as it reaches the critical distance  $\varepsilon$  between the two molecules, it is further tracking according to (5).

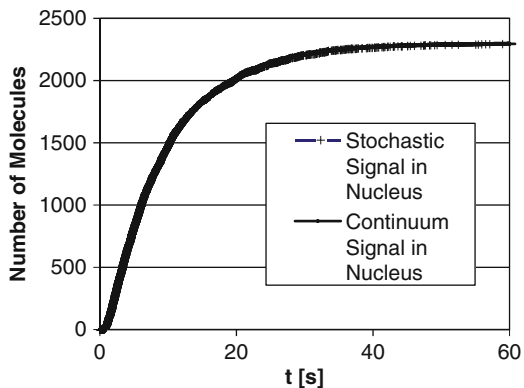
### 3.1.3 Effect of Macromolecular Crowding

Molecular crowding and the cytoskeleton have to be taken into account to get a more realistic consideration of the cellular architecture [38]. The cytoskeleton is simulated by inserting randomly distributed cylinders into the cell. Thus the free volume is reduced by 30%. To model the interaction with the cytoskeleton cylinders we rejected steps that would end inside a cylinder. The effective diffusion coefficient is about proportional to the free volume in this case [50]. The effective concentration of reaction partners is increased because the free volume is reduced, and thus molecular crowding increases the reaction speed.

### 3.1.4 Advantages of the Multi-Scale Modeling Approach

By switching to the solution of the Fokker–Planck equation we ensure that the effect of reactions as well as obstacles is properly considered in the diffusion path of the diffusing molecules. Fixed reaction distances like in Smoldyn [45] simply cut

**Fig. 5** Comparison of stochastic and continuum signal



off the Gaussian probability density function for the diffusion steps, ignoring the unknown positions of the particles between  $t$  and  $t + \Delta t$ . The probability density function according to the Fokker–Planck equation is written to a lookup table so there is only a minor increase in computation time. The comparison under spatial homogeneous continuum conditions shows excellent agreement (see Fig. 5).

## 3.2 Applications

### 3.2.1 Impact of Spatial Separation of Kinases and Phosphatases on the Output Signal of the MAPK Cascade

The MAPK cascades contain three interconnected cycles of MAPK, a MAPK kinase (MAPKK) and a MAPKK kinase (MAPKKK). In the most well characterized MAPK/Erk cascade, the system consists of ERK, MEK and Raf. Upon stimulation and Ras activation, the cytosolic Raf is recruited to the cell membrane, where it binds to and phosphorylates MEK. The phosphorylated MEK drifts into the cell interior where it phosphorylates ERK. ERK then travels through the cytoplasm into the nucleus, where it triggers the expression of certain genes. During its random walk ERK can be attacked by various phosphatases and after successful reaction would lose its activation state [39].

The example is, first of all, used to examine the reliability of the approach and the various numerical methods required for the simulations. For this purpose we use a simulation of the temporal-spatial distribution of 4,000 ERK molecules starting at time zero from the cell membrane. According to the fluctuation theory the variance is proportional to  $1/\sqrt{N}$ , where  $N$  is the number of particles. With  $N = 4,000$  one should expect an agreement between continuum and discrete simulation within 1.6%. For the continuum approach we solve the partial differential equation for



diffusion and first order reaction (the number of phosphatase molecules is not changed in the process).

$$\begin{aligned}
 \frac{\partial c(r, t)}{\partial t} &= D \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial c(r, t)}{\partial r} - Kc(r, t) \\
 B.C. : \quad c(r, t)|_{r=R_0} &= 0 \quad (R_0 \text{ is the radius of the nucleus}) \\
 D \frac{\partial c(r, t)}{\partial r} \Big|_{r=R} &= \phi_0 \delta(t) \quad (\text{initial flux from the cell membrane}) \\
 I.C. : \quad c(r, t=0) &= 0
 \end{aligned} \tag{9}$$

Analytical solution:

$$\begin{aligned}
 c(r, t) &= \sum_{n=1}^{\infty} \frac{2\phi_0}{rD} \frac{e^{-(\lambda_n^2/(1-R_0)^2 + Da)t} \sin(\lambda_n(r - R_0)/(1 - R_0))}{(1 - R_0)(\lambda_n \sin(\lambda_n) - R_0 \cos(\lambda_n))} \\
 \text{with Damkohler Number } Da &= \frac{KR_{\text{cell}}^2}{D} \\
 \text{and eigenvalues } \lambda_n \cos \lambda_n &= (1 - R_0) \sin \lambda_n
 \end{aligned}$$

From this solution the flux into the nucleus and the concentration inside the nucleus can be derived:

$$\phi_{\text{nucleus}}(t) = D \frac{\partial c(r, t)}{\partial r} \Big|_{r=R_0} \quad c_{\text{nucleus}}(t) = 4\pi R_0^2 \int_{t'=0}^t \phi_{\text{nucleus}}(t') dt'$$

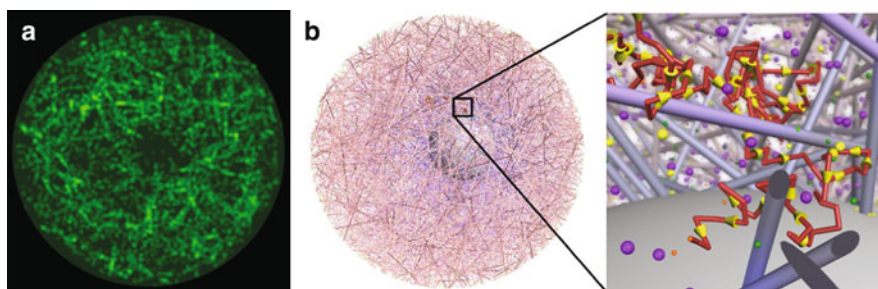
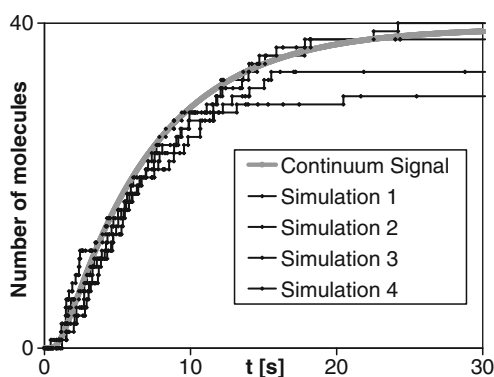
The result of the comparison (Fig. 5) is more than satisfactory as one cannot see any difference in the dynamic response between the continuum and stochastic simulation of the problem.

As expected, the differences are much more pronounced if the number of molecules is decreased.

Figure 6 shows simulation result for a scenario with only 100 ERK molecules. There are differences between the discrete and continuum simulations and, additionally, the dynamic response in case of the stochastic simulation differs from run to run.

For adequate representation of the simulation results, a visualization framework has been implemented [51]. It allows an interactive exploration of the data from the simulation. Two visualization styles have been developed (see Fig. 7): a microscopic and a more schematic representation. A virtual microscope creates images, which look like the results from a confocal fluorescence microscope. With this representation the simulation can easily be compared with microscope images from experiments. The schematic visualization is more abstract and visualizes all

**Fig. 6** Comparison of stochastic and continuum signal for 100 ERK molecules (the remaining ERK molecules are dephosphorylated and do not reach the nucleus)



**Fig. 7** Visualization of the virtual cell. (a) fluorescence microscope. (b) detailed image showing the cytoskeleton, nucleus, crowding molecules (*violet, green and yellow*) and signaling molecules (*orange*), including the path of a signaling molecule (*red, arrow cones in yellow show the direction of the path*)

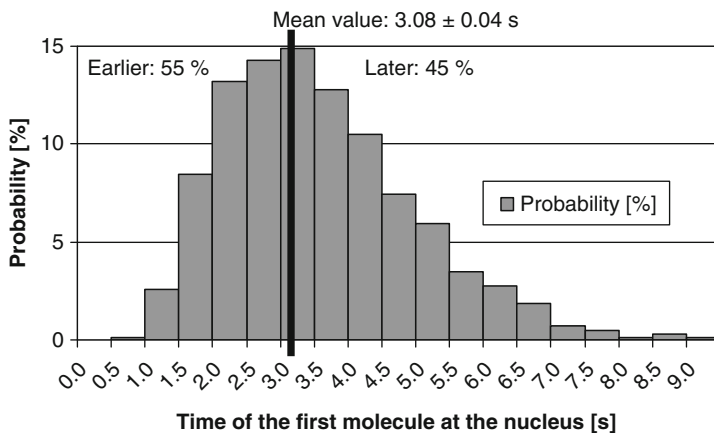
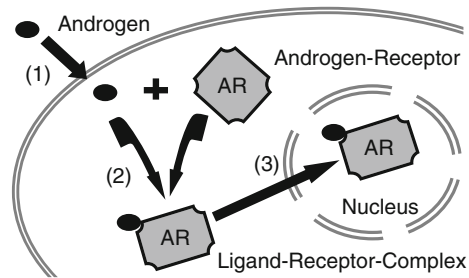
simulated components. Crowding and signaling molecules are represented by spheres with the respective radius, and the filaments of the cytoskeleton are represented by cylinders. Molecule paths can be highlighted to follow a molecule of interest.

### 3.2.2 A Steroid Hormone Pathway: A Case Study of a Bimolecular Reaction

The steroid pathway with its ligand activated steroid hormone receptor (androgen, androgen-receptor) differs from the membrane anchored receptor–ligand interactions and mobilization of phosphorylated proteins illustrated for the MAP kinase in that the steroid hormone is able to penetrate the cell membrane and then binds to the receptor (see Fig. 8). The receptor–ligand complex then travels to and is imported into the nucleus [52].

From the point of view of kinetics, the important difference is that the problem is characterized by a real bimolecular reaction  $A+B=C$ . It is the nonlinearity of the

**Fig. 8** Steroid hormone pathway for androgen: androgen-hormone diffuses into cell (1), binds to androgen-receptor (2) and the complex is imported into the nucleus (3)



**Fig. 9** Probability distribution for the first passage time of the first androgen–receptor–complex reaching the nucleus

reaction which should lead to much more pronounced differences between discrete and continuum simulations.

It is not possible to derive an analytical solution for the continuum in case of a bimolecular reaction so we solved (10) with the bimolecular reaction term  $Kc_1(r,t)c_2(r,t)$  numerically. In the stochastic simulation the first passage time of the androgen–receptor–complex was recorded in 1,008 trials with 500 androgen molecules initially located at the cell membrane and 500 receptors randomly distributed in the cell. The statistical analysis reveals that 55% of the complexes reach the nucleus earlier than the sample mean of 3.08 s with a standard error of 0.04 s (see Fig. 9). In the continuum approach the first particle arrives 4% later (3.20 s); already 57% of the trials arrived within that time. This shows, that local density fluctuations have an effect on the average reaction of bimolecular reactions (and this example has not even been optimized for a maximum nonlinearity effect; it was designed to reflect natural conditions).

## 4 Conclusion

The basic idea central to the agent based or individual modeling approach presented in this chapter are entities as objects (cells or molecules) traveling along paths which are computed from stochastic differential equations. In the case of the bioreactor model the random walk of individual cells is calculated from CFD-simulations, in which turbulent dispersion is superimposed to the three-dimensional convective movement in the turbulent flow field. The example presented deals with the impact of sugar transport into bacterial cells containing a PTS. The method allows the population behavior to be described at the outcome of the interaction between the intracellular state of its individual cells and the turbulent flow field in the bioreactor. The chosen Euler-Lagrange representation of the cell-ensemble approach permits analysis of the lifelines of individual cells in space and time. The approach presented integrates CFD with a segregated description of a cell population in a stirred tank thereby accounting for a detailed intracellular structure of the single cells. The biological example tackled with this approach is of great practical relevance. The simulation results point to serious differences in the dynamics of the intracellular states at different scale of operation with significant impact on the viability of the cells.

The second application of this approach is the random walk of molecules in individual cells. The two examples chosen comprise the MAP kinase and a hormonal stimulation. In contrast to the first example, in which the communication between the objects is moderated by the extracellular environment, thus neglecting a direct interaction, the signal transduction examples involve a direct molecular interaction via biochemical reactions. To overcome the problem of a step size dependent influence upon the collision frequency of molecules the reaction probability is estimated from the theoretical solution of the Fokker–Planck equation. This switch in the model approach is a special kind of multi-scale modeling characterized by the transition from the physical into the probabilistic space.

**Acknowledgements** The authors acknowledge support of the Deutsche Forschungsgemeinschaft (DFG) within the collaborative research center “Sonderforschungsbereich 412” and the Ministry of Science, Research and Arts Baden-Württemberg within the Center Systems Biology University Stuttgart.

## References

1. Frederickson AG, McGee RD, Tsuchiya HM (1970) *Adv Appl Microbiol* 23:419
2. Kitano H (2006) *Nat Rev Mol Cell Biol* 7:163
3. Larsson G et al (1996) *Bioproc Eng* 14:281
4. Noorman H et al (1993) *Proceedings of the 3rd international conference on bioreactor and bioprocess fluid dynamics*, Cambridge 241:258
5. Reuss M, Schmalzriedt S, Jenne M (2000) *Bioreaction engineering: modeling and control*. Springer, Berlin 207:246

6. Schmalzriedt S et al (2003) *Adv Biochem Eng* 18:68
7. Trägårdh C (1988) The second international conference on bioreactor fluid dynamics, King, Applied Sciences Publishing, Cambridge 117:134
8. Lapin A et al (2002) *Chem Eng Sci* 57:1419
9. Sokolochin A et al (1997) *Chem Eng Sci* 52:61
10. Bourloutski E, Sommerfeld M (2002) Proceedings of the 10th workshop on two-phase flow predictions, Merseburg 113:223
11. Van Sint Annaland M, Deen NG, Kuipers JAM (2005) *Chem Eng Sci* 60:6188
12. Bezzo F, Macchietto S, Pantelides CC (2003) *AIChE J* 49:2133
13. Henson MA, Müller D, Reuss M (2002) *Biochem J* 368:433
14. Bauer M, Eigenberger G (1999) *Chem Eng Sci* 54:5109
15. Bauer M, Eigenberger G (2001) *Chem Eng Sci* 56:1067
16. Ataii MM, Shuler ML (1985) *Biotechnol Bioeng* 27:1026
17. Domach MM, Shuler ML (1984) *Biotechnol Bioeng* 26:877
18. Kim BG, Shuler ML (1990) *Biotechnol Bioeng* 36:581
19. Lapin A, Müller D, Reuss M (2004) *Indus Eng Chem Res* 43:4647
20. Lapin A, Schmid J, Reuss M (2006) *Chem Eng Sci* 61:4783–4797
21. Bylund F, Collet E, Enfors S-O, Larsson G (1998) *Bioprocess Eng* 18:171
22. Bylund F et al (1999) *Bioprocess Eng* 20:377–389
23. Schweder T et al (1999) *Biotechnol Bioeng* 65:151
24. Teich A et al (1999) *Biotechnol Prog* 15:123
25. Xu B et al (1999) *Appl Microbiol Biotechnol* 51:564
26. Bajpai RK, Reuss M (1982) *Can J Chem Eng* 60:384
27. Jenne M, Reuss M (1999) *Chem Eng Sci* 54:3921
28. Fox RO (2003) *Computational models for turbulent reacting flows*. University Press, Cambridge
29. Chassagnole C et al (2002) *Biotechnol Bioeng* 79:53
30. Heinrich R, Schuster S (1996) *The regulation of cellular systems*. Chapman & Hall, New York
31. Hewitt CJ, Nebe-Von Caron G (2001) *Cytometry* 44:179
32. Hewitt CJ et al (1999) *Biotechnol Bioeng* 63:705
33. Hewitt CJ et al (2000) *Biotechnol Bioeng* 70:381
34. Lengeler JW, Drews G, Schlegel HG (1999) *Biology of the prokaryotes*. Thieme, Stuttgart
35. Lewis K (2000) *Microbiol Mol Biol Rev* 64:503
36. Loewen PC et al (1998) *Can J Microbiol* 44:707
37. Takahashi K, Arjunan SNV, Tomita M (2005) *FEBS Lett* 579:1783
38. Kholodenko BN (2006) *Nat Rev* 7:165
39. Kholodenko BN (2003) *J Exp Biol* 206:2073
40. Howe CL (2005) *Theor Biol Med Model* 2:43
41. Gillespie DT (1976) *J Comp Phys* 22:165
42. Stundzia AB, Lumsden CJ (1996) *J Comput Phys* 127:196
43. Ander M, Beltrao P, Di Ventura B, Ferkinghoff-Borg J, Foglierini M, Kaplan A, Lemerle C, Tomás-Oliveira I, Serrano L (2004) *Sys Biol* 1:129
44. Tolle DP, Le Novère N (2006) *Curr Bioinform* 1:3
45. Andrews SS, Bray D (2004) *Phys Biol* 1:137
46. Stiles JR, Bartol TM (2000) In: de Schutter E (ed) *Computational neuroscience: realistic modeling for experimentalists*. CRC, Boca Raton, FL, p 87
47. Pogson M, Holcombe M, Smallwood R, Qwarnstrom E (2006) *BioSystems* 85:37
48. Batada NN, Shepp LA, Siegmund DO (2004) *Proc Natl Acad Sci USA* 101:6445
49. Rice SA (1985) In: Bamford CH, Tripper CFH, Compton RG (eds) *Diffusion-limited reactions*. Amsterdam, Elsevier
50. Trinh S, Arce P, Locke BR (2000) *Transp Porous Media* 38:214
51. Falk M, Klann M, Reuss M, Ertl T (2009) Proceedings of IEEE pacific visualization symposium 2009, Beijing 169:176
52. Tyagi RK, Lavrovsky Y, Ahn SC, Song CS, Chatterjee B, Roy AK (2000) *Mol Endocrinol* 14:1162

Biosystems Engineering II

Linking Cellular Networks and Bioprocesses

Wittmann, C.; Krull, R. (Eds.)

2010, XIV, 166 p. 48 illus., 16 illus. in color., Hardcover

ISBN: 978-3-642-13865-2