

Reaching new levels of realism in modeling biological macromolecules in cellular environments



Michael Feig^{d,b,*}, Yuji Sugita^{a,b,c}

^a RIKEN Advanced Institute for Computational Science, 7-1-26 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

^b RIKEN Quantitative Biology Center, International Medical Device Alliance (IMDA) 6F, 1-6-5 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

^c Theoretical Molecular Science Laboratory, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama, Wako-shi 351-0198, Japan

^d Department of Biochemistry & Molecular Biology and Department of Chemistry, Michigan State University, 603 Wilson Road, BCH 218, East Lansing, MI 48824, United States

ARTICLE INFO

Article history:

Received 4 July 2013

Received in revised form 14 August 2013

Accepted 19 August 2013

Available online 28 August 2013

Keywords:

Crowding
Confinement
Reaction-diffusion
Coarse-graining
Implicit solvent
Brownian dynamics

ABSTRACT

An increasing number of studies are aimed at modeling cellular environments in a comprehensive and realistic fashion. A major challenge in these efforts is how to bridge spatial and temporal scales over many orders of magnitude. Furthermore, there are additional challenges in integrating different aspects ranging from questions about biomolecular stability in crowded environments to the description of reactive processes on cellular scales. In this review, recent studies with models of biomolecules in cellular environments at different levels of detail are discussed in terms of their strengths and weaknesses. In particular, atomistic models, implicit representations of cellular environments, coarse-grained and spheroidal models of biomolecules, as well as the inclusion of reactive processes via reaction-diffusion models are described. Furthermore, strategies for integrating the different models into a comprehensive description of cellular environments are discussed.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Realistic models of cellular environments have long captured the imagination of biologists and physical scientists alike. A major driving force is that such models hold the promise to eventually make the genotype–phenotype connection by predicting a cell's behavior from its genomic DNA and resulting physical constituents. A particular exciting application of virtual cell models would be the rational design of new drugs to predict the effects of potential drug candidates on an entire cell instead of focusing on single targets at a time, which is the current paradigm [1,2]. A whole-cell focused drug design strategy would be able to identify potential unintended side effects early on and is likely to be much more effective for developing treatments of complex diseases such as cancer [3].

A defining step in capturing the fully complexities of biological cells were Goodsell's inspiring renderings of cellular environments based on knowledge available more than twenty years ago [4]. Since then, increasingly realistic simulations of bacterial cytoplasmic models have begun to appear [5–8] that suggest that we are rapidly moving toward comprehensive, physically and biologically

accurate cellular models. A major challenge in developing such models is the need to cover a wide range of scales. Atomistic details of molecular processes occur on length scales of 0.1 nm while cellular dimensions are between 300 nm for the smallest bacterial cells and 100 μm for large eukaryotic cells. Biologically relevant time scales range from nanosecond to microsecond time scales for the internal dynamics of individual molecules to time scales of seconds to hours for entire biological processes. While a model that represents an entire cell in full atomistic detail is conceivable, it would be extremely costly to reach nanosecond time scales with today's most advanced computers and impossible to cover seconds for such a system in the foreseeable future. Instead, the development of reduced and multi-scale models is the only viable strategy today for developing physically accurate models that at the same time cover both cellular length scales and time scales of cellular-level biological processes. Such models trade reduced levels of detail for increased computational speed and the key issue is what approximations one can tolerate without comprising the overall level of realism so much that a given model loses its predictive ability. In this review a variety of recently proposed models for biomolecules in cellular environments are discussed and compared in terms of their strengths and weaknesses.

An additional issue is that most past and present research aimed at understanding biological dynamics and function tends to fall

* Corresponding author. Tel.: +1 517 432 7439.

E-mail address: feig@msu.edu (M. Feig).

into two camps, concentrating on either extreme of the spatial and temporal scales. On one hand, structural biology has cultivated a single-molecule, or single-complex, focused view that has led to highly predictive models. They have provided a wealth of insight into the sequence–structure–function paradigm for a large variety of biological macromolecules but with only limited consideration of the cellular environment. On the other hand, the still growing field of systems biology seeks to quantitatively describe macroscopic phenomenological observations at the cellular level with often mathematically complex but physically relatively simple models that rarely go beyond representing biomolecules as single spherical objects. There are intellectual and cultural gaps between these two communities and enhancing mutual interactions is likely to be just as important as devising computationally efficient multi-scale modeling strategies. A further aim of this review is to outline how an integration of these different views could be accomplished.

In the following, we will first touch upon the different aspects to be considered when modeling biological processes at different scales in the context of cellular environments. We will then present an overview of recently applied models and discuss strategies for integration into comprehensive models of cellular environments that bridge between scales and in particular between structural and systems biology. The focus here is primarily on the modeling of cellular environments, but we will also briefly mention how to make important connections with experiments as appropriate.

2. Complexity of cellular environments

Depending on the perspective of the reader, the notion of cellular environments is likely to evoke very different ideas. To the biologist, the complexity of metabolic cycles and regulatory networks may first come to mind while structural biologists and biophysicists may think about the effects of crowding and confinement.

If we begin to think about cellular environments from a *chemical perspective*, the simplest cells, bacterial cells, consist in the interior of proteins and nucleic acids at a volume fraction of typically 20–30%, a diverse array of metabolites at different concentrations, various monoatomic ions, and water filling the remaining space [4,9,10]. Lipid membranes form the cell perimeter with a significant fraction of embedded membrane proteins [11]. In addition, cytoskeletal elements may be present to provide mechanical stability, facilitate cell motility, or to aid with developmental processes such as cell division [12]. Eukaryotic cells are complicated further by compartmentalization and a more extensive use of cytoskeletal elements [13].

From a *physical perspective*, the most defining feature of cellular environments is the effect of macromolecular crowding and confinement. Much attention has focused on how crowding and confinement affect protein folding landscapes [14–23] and protein–protein association rates [15,24–27]. The general picture from a combination of experimental, theoretical, and simulation studies is that both crowding and confinement tend to disfavor extended non-native states vs. compact, native states. The main effect at play is volume exclusion, which introduces an entropic penalty for extended states or makes them simply impossible under conditions of rigid confinement [28]. Furthermore, crowding may alter the internal flexibility or functionally relevant conformational dynamics [29–31]. This view is modulated by recent studies that indicate that native state destabilization may be possible when the crowder molecules are proteins as well so that folding and favorable protein–protein interactions can become competing processes [17,18,32,33].

From a *structural perspective*, we are approaching a time when complete structuromes for a given organism will likely become

available [34], at least at the level of individual macromolecules. If one combines experimentally resolved structures and structures that can be modeled with reasonable certainty through homology [35], structures are now available for the large majority of gene products for a given organism [34,36]. This is especially true for well-ordered soluble proteins. Membrane proteins are more problematic because much fewer experimental structures are available [37,38], but as that number increases, the fraction of membrane protein structures that can be modeled will also increase [37]. Disordered regions or entire proteins that are intrinsically disordered pose another challenge that is difficult to address experimentally [39], but insight into dynamic ensembles for such systems could be obtained at least in principle using computational tools [40–42].

Nucleic acid structure is another challenge. Certain types of RNA, in particular ribosomal and transfer RNA, are well resolved but much less is known about the detailed structural organization of genomic DNA [43]. It may be ironic that although the structure of short oligomeric DNA was resolved a long time [44] and only shortly after the first protein was resolved in atomic detail [45,46], the highly complex and dynamics structure of genomic DNA with millions of base pairs is now turning out to be one of the last frontiers of structural biology [47–50].

Finally, complete cellular models have to also consider the surrounding biological membranes. Biological membrane should be modeled as mixed systems containing a variety of lipid molecules and a high concentration of membrane proteins. Membranes are highly dynamic and the lipid composition can change the elasticity of biological membranes and varies for the membrane location in a cell. The modeling of membranes is further complicated by the formation of lipid rafts that form subdomains in cellular membranes, which contain cholesterol and sphingomyelins at high concentrations. Although relatively small phospholipid bilayer systems are modeled routinely, realistic models of entire cellular membranes have not yet been established.

In the *biological context*, interactions between different macromolecules are often the key to their biological function but also the source of undesirable aggregation [51]. This results in supramolecular complex formation with varying degrees of stability. There are many examples of well-resolved complexes such as homooligomeric structures or multimeric functional complexes like the ribosome or RNA polymerase [52,53]. However, presumably these complexes are highly stable with limited internal dynamics since otherwise successful crystallization would have been unlikely. Other, more dynamic complexes may only be accessible via lower resolution techniques such as electron microscopy in combination with modeling [54–56]. Finally, at the level of whole cells, tomography can determine not just the structures of complexes but their locations within a cellular context [57,58]. This was demonstrated recently for *Mycoplasma pneumoniae* [59].

From a *biochemical perspective*, cells are exquisitely tuned reactors that carry out an enormous variety of chemical reactions in parallel within a very dense environment. Essentially, these reactions serve two main objectives: (1) Continuous generation of energy from varying types of metabolites absorbed from the environment; and (2) Growth and propagation through cell division. The underlying reaction networks have been the subject of many studies [60]. We are now reaching the point of being able to generate complete reaction networks where all products and reactants are accounted for in metabolic cycles and where, at the same time, all of the enzymes involved in catalyzing the various reactions are mapped to gene products for a given organism. An example was recently given for the case of *Mycoplasma genitalium*, a bacterium with the smallest known genome [61]. As functional annotations of genomes and the application of proteomics and metabolomics studies continuously increase, it can be expected that similar complete network reconstructions will soon be reported for many other

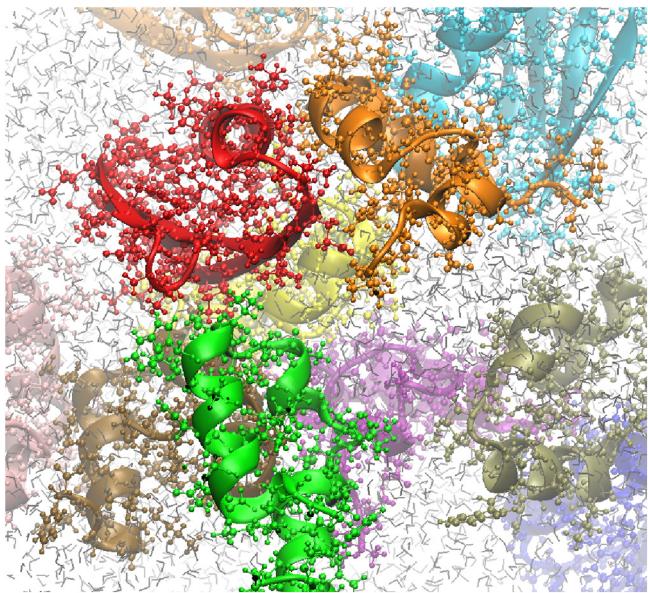


Fig. 1. Fully atomistic model of biomolecules in a crowded environment.

organisms as well. As a sign that this a rapidly moving field, a database for comprehensive whole-cell models has already been established [62] soon after the first complete reconstruction for *M. genitalium* was published.

A critical feature of the actual performance of reaction networks is the choice of kinetic parameters of both, the enzyme-catalyzed reactions and the diffusive motions of enzymes, reactants, and products. As a first approximation, catalytic rates may be viewed largely as intrinsic to a given enzyme. On the other hand, it is clear that diffusion rates depend strongly on crowding and confinement in the cellular environment [63,64], yet much remains to be learned about the kinetic and dynamic behavior of biomolecules in cellular environments.

Integrating all of these different aspects into a comprehensive and coherent picture that is consistent with experimental data is the essential conundrum of developing realistic models of cellular environments [65,66].

3. Models of cellular environments

3.1. Fully atomistic models

The most detailed models of cellular environments consist of fully atomistic descriptions of biological macromolecules, metabolites, and the surrounding solvent and/or membrane environment (see Fig. 1). Such models have 0.1 nm spatial resolution and fs time resolution if they are investigated via molecular dynamics (MD) simulations. Because of prohibitive computational costs, fully atomistic MD simulations were applied up until recently only to the study of single macromolecules at a time and rarely exceed microsecond time scales. In simulations of single biomolecules, cellular environments can be taken into account by creating confinement through artificial walls or pores [67–71] or, using a more biological environment, through reverse micelles [72]. Similarly, crowding effects can be included via spherical model crowder particles along with an otherwise fully explicit solute and solvent representation [73]. An alternative to direct inclusion of the crowders in the simulation of the biomolecule of interest is the post-processing approach proposed by Zhou et al. [74–77]. In this method the biomolecule is first simulated separately in conventional atomistic simulations and the generated ensemble is

subsequently reweighted by insertion into snapshots from simulations of just the model crowders to include the effect of crowding [74–76,78].

Another direction toward a more realistic account of cellular environments consists of simulations of biomolecules in the presence of various co-solvents to more closely mimic the chemical diversity of biological solution environments. The most complete simulation in that respect is a recent study by Cossins et al. where ubiquitin was simulated in the presence of a large number of representative metabolites from *Escherichia coli* [79]. Other examples of such simulations include simulations in the presence of trimethylamine-*N*-oxide (TMAO) [80–83] and denaturants such as urea and guanidinium chloride [84–92].

During the last few years, atomistic MD simulations of multiple interacting biomolecules in explicit solvent have also begun to appear including simulations of amino acid side chain analogs [93,94], peptides [95], and, most recently, proteins [32,33,96] at cellular concentrations. The latter probably represent the most realistic simulations of cellular environments to date.

The results of these studies have provided detailed insight into how biological macromolecules may be affected by cellular-like environments in terms of stabilizing native conformations vs. denatured and unfolded states, altered solvent interactions including preferential interactions with co-solvents, and altered diffusive properties. Many of these studies confirm the well-known general trend of favoring compact, native states over extended states upon crowding or confinement. However, there is also emerging evidence in particular from studies from our group that the presence of protein crowders may lead to structural perturbations that range from subtle changes to partial denaturation apparently as a consequence of non-specific protein–protein interactions in highly crowded environments [32,33]. In the case of chymotrypsin inhibitor 2 (CI2), increased structural fluctuations were observed at parts of the structure that was subject to extensive interactions with lysozyme crowder proteins [33] while such fluctuations were absent in the presence of bovine serum albumin (BSA) crowders that barely interacted with CI2. Interestingly, the site of increased structural fluctuations in the presence of lysozyme corresponded with increased hydrogen exchange in NMR experiment under the same conditions [18]. In a different system consisting of a mixture of protein G and villin head piece (referred to as villin from hereon) molecules at high concentrations, the native villin structure was destabilized in the presence of crowder proteins including partial loss of secondary and tertiary structure [32]. The denaturation was compensated by enthalpic gains due to protein–protein interactions between villin and protein G and overall favorable solvation free energies because of a reduced cost of cavity for the villin–protein G complex vs. individually solvated molecules [32].

Predictions of altered native state stability can be compared to experimental structural measures that either probe structure directly via NMR and other spectroscopic studies [97–99] or, indirectly, by probing for example changes in solvent-accessibility of natively buried residues as in hydrogen-exchange NMR measurements [100]. However, such experiments are best suited to detect significant structural changes and/or large shifts in conformational populations because of crowding while it is more difficult to detect the emergence of crowding-induced minor sub-populations of non-native states or alternate folding intermediates as suggested by some recent simulations [32,73].

Protein–protein interactions also affect diffusive properties in non-trivial ways. In the CI2 system in the presence of either lysozyme or BSA crowders at a volume fraction of about 7%, diffusion was reduced by a factor of two in the presence of BSA but by a factor of four in the presence of lysozyme [33], in rough agreement with experimental data [101]. Furthermore, there was a significant difference in short-term (<10 ns) and long-term (>10 ns) diffusion

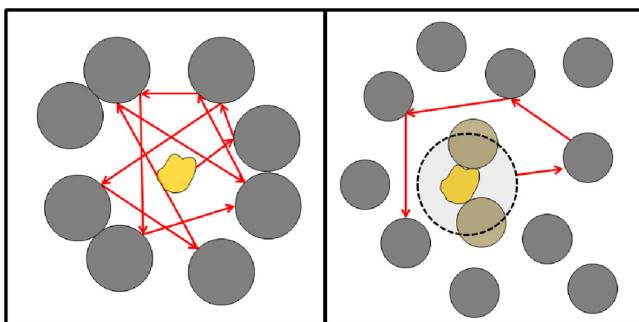


Fig. 2. Schematic model of diffusion of chymotrypsin inhibitor 2 (in yellow) in the presence of bovine serum albumin (left) and lysozyme (right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

rates in the presence of BSA but not with lysozyme indicating the presence of anomalous diffusive behavior. The interpretation of these results is that in the presence of BSA due to the lack of strong CI2–BSA interactions, CI2 is essentially free to move on short time scale but is slowed down when it eventually it encounters a BSA molecule leading to a type of cage effect (see Fig. 2). In contrast, stronger interactions with lysozyme result in CI2 effectively moving as a larger particle. Therefore, diffusion rates are reduced more strongly and with little difference between fast and slow time regimes (see Fig. 2). In villin/protein G and protein G only systems where higher protein volume fractions were reached diffusion of villin was reduced by a factor of 8 and that of protein G by a factor of 4–5 (see Fig. 3), somewhat less but comparable to experimental observations of reduced diffusion of GFP in cellular environments by about a factor of 10. Interestingly, the diffusion of protein G was reduced more in the mixed protein G/villin system than in the pure protein G system. Furthermore, protein G and villin diffusion rates were similar under crowding conditions while villin in dilute solution diffuses faster by 50% compared to protein G. Since extensive protein G–villin interactions were observed while protein G–protein G interactions were much weaker, this again suggests that protein–protein interactions affect diffusion more strongly if transient, non-specific complexes are formed upon crowding.

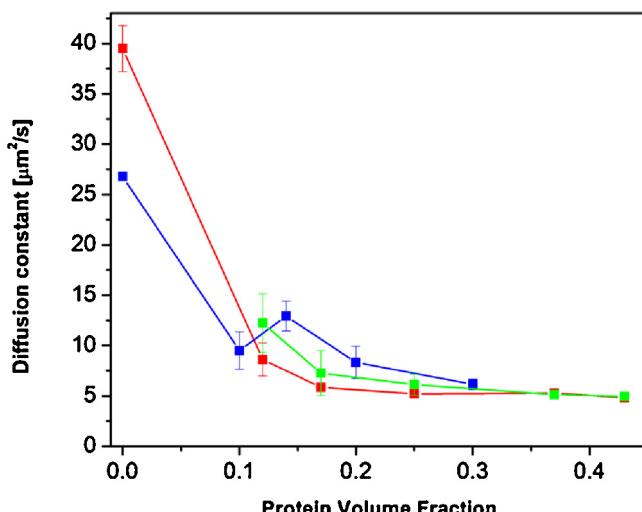


Fig. 3. Diffusion of villin (red) and protein G (green; blue) as a function of protein crowder volume fraction from simulations with error bars to indicating statistical errors from simulations of villin/protein G mixtures (red; green) or only protein G (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Diffusion rates predicted from simulations can be compared to experimentally obtained diffusion measurements from various fluorescence-based techniques such as single-particle tracking (SPT), fluorescence recovery after photobleaching (FRAP), or fluorescence correlation spectroscopy (FCS) [63]. While much insight about diffusion in crowded environments has been gained from such experiments, it is generally difficult to adequately describe spatially and temporally varying diffusive behavior and fully validate predictions of the presence of anomalous diffusion in crowded cellular environments [33,102–104]. Furthermore, the experimental methods rely on labeling with fluorophores, which presents a negligible perturbation when studying larger biomolecules but becomes an issue when studying smaller molecules such as metabolites.

The detailed insight described above would be difficult to obtain with simplified crowder models that focus on size exclusion but neglect the full spectrum of physicochemical properties of actual proteins. Additional simulations of atomistic biomolecules in cellular environments at increasing scales are likely to emerge in the near future as they provide the most reliable information into the effects of cellular environments apart from experiments. Such detailed information is of fundamental interest but also needed to inform the design of more approximate models discussed below.

3.2. Implicit solvent models of cellular environments

Biomolecules are present at high concentrations in cellular environments, but the majority (60–75%) of the molecular species is still water. Therefore, replacing explicit solvent with an implicit model significantly reduces system complexity and computational cost over models that include the solvent explicitly [105]. Solvent affects the conformations of biomolecules through enthalpic and entropic effects but also determines diffusional properties through hydrodynamic effects [106].

The thermodynamic effect of aqueous solvent is well represented by a dielectric continuum with the theoretical framework provided by Poisson–Boltzmann theory [107]. The default choice of the dielectric constant for representing aqueous solvent is 80, but recent studies indicate that macromolecular crowding may slow down water dynamics and reduce the dielectric constant of the solvent by as much as a factor of two [96,108,109]. The direct solution of the Poisson–Boltzmann equation to obtain the solvation free energy associated with the dielectric continuum model is straightforward [110] but a convenient alternative is the use of the numerically more efficient Generalized Born approximation [111]. Additional thermodynamic effects of the solvent environment arise from the entropic cost of creating a cavity in aqueous solvent and from solute–solvent van der Waals interactions. The cost of cavity can be simply related to the surface area [112] or molecular volume [113] of a given solute while approximate formalisms are available to describe solute–solvent van der Waals interactions [114]. Similar implicit solvent formalisms have also been proposed for modeling membrane environments [115–117]. Simulations of biological macromolecules using dielectric continuum based implicit solvent have found a wide range of applications in studies of the conformational sampling of biological macromolecules [111]. While most of these studies aim at modeling macromolecules in dilute solvent, simulations with reduced dielectric constants to model the effect of cellular conditions have also been carried out [118]. As mentioned above, a reduced dielectric constant (in the range of 40–60) reflects the altered properties of water in crowded environments [96]. However, a further reduced dielectric constant (20–40) may approximately capture the additional loss of environmental polarizability due to the displacement of water by proteins with a lower dielectric response than water [118–120]. On the other hand, the volume exclusion effect of crowding that leads to a shift toward

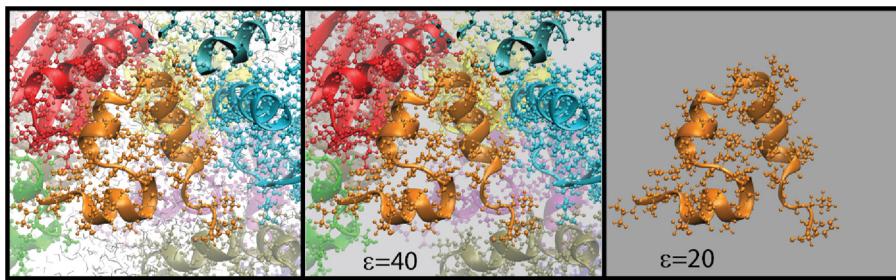


Fig. 4. Explicit solvent model of cellular environments (left) compared with intermediate dielectric model of water in crowded environments (center) and low dielectric model of cellular environments (right).

compact structures vs. extended structures can be modeled by increasing the solvent-accessible surface area based cost of cavity beyond that of dilute aqueous solvent. Thereby, low-dielectric continuum models with an increased cost of cavity may be viewed as a first order approximation of the thermodynamic effects of cellular environments that does not require the inclusion of crowder molecules (see Fig. 4).

Salt effects that are especially important for nucleic acids can be included through ionic concentrations in Poisson–Boltzmann theory [121] and, more approximately, also in Generalized Born-type formalisms [122]. However, this approach is strictly only valid for relatively low ionic concentrations. Alternatively, ions can be included explicitly along with the solute [123] although it tends to be difficult to reproduce the delicate balance of ion–ion, ion–solvent, and ion–solute interactions in the absence of explicit solvent [124].

The kinetic effects of the solvent are important for determining the diffusional properties of macromolecules in cellular environments [63]. There are essentially two effects: (1) Random collisions between water molecules and a given solute introduce stochasticity in the solute momenta. (2) The presence of solvent viscosity results in a velocity-dependent friction force. For a spherical solute at infinite dilution, these effects are isotropic but non-spherical shapes and the presence of other solutes at finite concentrations requires a friction tensor to fully capture hydrodynamic interactions.

The Langevin equation of motion extends Newtonian mechanics with stochastic collision and frictional terms to capture these effects [125]:

$$m_i \ddot{a}_i = -\nabla_i V - \sum_j \gamma_{ij} v_i + R_i \quad (1)$$

where m_i , a_i , and v_i are mass, acceleration, and velocity of a given particle i , V is the interaction potential, γ_{ij} is the friction tensor, and R_i is a random force to mimic stochastic collisions.

The Langevin formalism can accurately describe the kinetics of solutes in aqueous solution without explicitly including the solvent [126,127]. However, because the same short integration time-step as in Newtonian dynamics is required, it is still a computationally expensive approach when studying systems of interacting molecules over long time scales.

A further approximation is obtained in Brownian dynamics (BD) [128], which assumes that momenta relax much faster than positions thereby representing the high-inertia limit. This allows the Langevin equation to be simplified to the BD equation [129]:

$$\dot{r}_i(t + \Delta t) = r_i(t) + \sum_j \nabla_j D_{ij} \Delta t + \sum_j \frac{D_{ij}}{k_B T} M_j \Delta t + R_i \quad (2)$$

with the additional condition that:

$$\langle R_i R_j \rangle = 2D_{ij} \Delta t \quad (3)$$

where r_i are the propagated positions of particles i during a time step Δt , D_{ij} is the diffusion tensor, M_j are the combined force and torque on the rigid body solutes, and R_i are random displacements.

In Eq. (2), the diffusion tensor is used as input to propagate particle positions along random displacements that are added to the effect of forces due to intermolecular interactions to simulate Brownian motion. Brownian dynamics is applied to rigid body solutes and with sufficiently long time steps so that the assumption of the high-inertia limit remains valid. The time steps that are used in BD simulations are on the order of picoseconds. Furthermore, because solutes are represented as rigid bodies, the electrostatic interactions between different solutes in the presence of a continuum dielectric implicit solvent model can be pre-computed, thereby avoiding the computational cost of an implicit solvent formalism at each time step. As a result, it is possible to reach much longer time scales – well into tens of microseconds – for systems consisting of 100s of proteins with still reasonable computational resources [7,8,130–132].

Because the random displacements in Eq. (2) are coupled to the diffusion tensor via Eq. (3), factorization of D is needed to obtain correlated random displacements. This is typically accomplished via the computationally expensive Cholesky decomposition [129]. In the case of a 6×6 diffusion tensor that captures translational and rotational diffusion properties for each solute independent of other solutes, the factorization of D can be pre-computed once at the beginning of a simulation. Full consideration of hydrodynamic interactions requires, however, a $6N \times 6N$ diffusion tensor that couples the motions of different solutes through hydrodynamic interactions. A common approach is to use the Rotne–Prager–Yamakawa form of the diffusion tensor [133,134] which avoids the need to include the diffusion tensor gradient term in Eq. (2) and focuses on the low-order, long-range effect of hydrodynamic interactions. The higher-order, near-field effect that includes lubrication forces that become increasingly important at short particle separations can be included via additional terms [135,136]. In the case of a full $6N \times 6N$ diffusion tensor, its form depends on the positions of all solutes and therefore its factorization needs to be carried out at each step, thereby greatly increasing the computational cost of BD simulations. Consequently, full hydrodynamic interactions have only been included in a few studies of interacting biomolecules [8,132,137,138]. From the studies available so far, it has become clear that inclusion of hydrodynamic interactions significantly reduces the diffusion of macromolecular solutes over the case without hydrodynamic interactions and results in much better agreement with experimental data for the diffusion of proteins in crowded cellular environments [8,138]. Recent progress in the treatment of hydrodynamics in BD simulations is likely going to allow larger-scale cellular simulations over increasingly long time scales [139,140].

BD simulations with continuum dielectric based implicit solvent and hydrodynamic interactions applied to atomistic solutes can provide a physically very meaningful picture of interacting

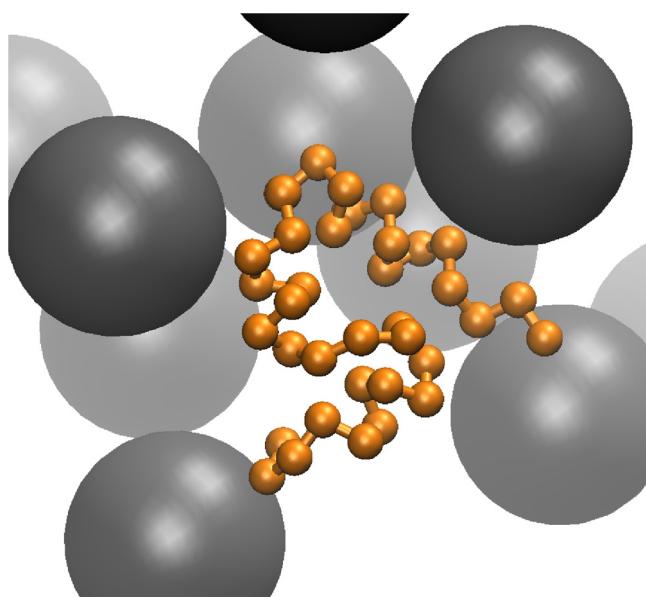


Fig. 5. Coarse-grained, $\text{C}\alpha$ -based model of a biomolecule in the presence of spherical crowders.

macromolecules in cellular environments. Results from such studies can provide insight into transient complex formation and potential organization of macromolecules involved in related biological function [141–145]. BD simulations can be used furthermore to generate realistic models of crowded cellular environments for thermodynamic analysis of macromolecular stability in crowded environments, e.g. via particle insertion methods [7,74]. The development of more efficient methods for including hydrodynamic interactions [132,146] will increase the scale of realistic BD simulations and it appears likely that such simulations will soon reach the scale of whole cells.

3.3. Molecular-shape preserving coarse-grained models of solutes in cellular environments

The next step in further model approximations involves coarse-graining biomolecular solutes (see Fig. 5). A large variety of coarse-grained models has been proposed ranging from near-atomistic resolution [147–149] to representations of a given molecule as a single sphere [150]. Increasing degrees of coarse-graining drastically reduce the computational cost so that biomolecular dynamics on the scale of whole cells and over biologically relevant time scales can be simulated. The trade-off is an increasing reliance on empirical models vs. physics-based models that limit transferability and the predictions that can be made. In this section, we will begin by discussing coarse-grained models that preserve the molecular shape, i.e. united-atom, residue-, or super-residue level models. Spherical models that do not preserve the molecular shape will be discussed in the next section.

A popular type of model employed in these studies for modeling peptides or proteins has a single bead per residue [31,151–160], connected with bond, angle, dihedral, and Lennard-Jones type potentials. Additional terms may be used to modulate attractive or repulsive interactions [153,156,157], add electrostatic interactions via a Debye–Hückel potential [155,159], or add certain residue–residue interactions with Go-like potentials to favor specific native-like folds of proteins that could otherwise not be maintained [152,154]. Some of these studies aim at modeling specific proteins or peptides, while other models describe polymers with general protein-like properties through a combination of hydrophilic and hydrophobic beads [31,151,157,158,161–163].

Slightly more detailed models include two beads per residue, one at the $\text{C}\alpha$ position and another one representing the amino acid side chain [29,164–174]. Other models that have been proposed reduce the core of a given protein sphere while maintaining residue-level beads for surface amino acids [159] or represent peptides at super-residue resolution [175,176]. In addition, coarse-grained models of nucleic acids have also been proposed to study crowding and confinement effects [177–179].

In the majority of studies using coarse-grained models of biomolecules, a single molecule, or a pair of molecules, is simulated in the presence of crowding or confinement conditions. Crowding is commonly modeled with repulsive crowder molecules with simple, often spherical, shapes although the effect of attractive protein–crowder interactions has also been considered [155] and in one study from our group crowder molecules were themselves represented with a coarse-grained model to provide protein-like behavior [119]. Alternatively, the effect of crowding can also be included implicitly via Asakura–Oosawa theory [152,180,181]. Confinement effects are typically included implicitly through appropriate boundary potentials. Most simulations of this type are carried out with Langevin dynamics or Monte Carlo sampling algorithms. They primarily aim at describing the thermodynamic effects of crowding or confinement on protein folding and aggregation rather than kinetic effects. In general, these studies emphasize native state stabilization under crowded conditions [168] and examine the modulation by a variety of more subtle effects as a function of crowder concentration [172], solute [171] and crowder shape [151,170], environmental anisotropy [173], and the presence of co-solvents [165]. As mentioned above, results from such studies can be compared to experimental stability measurements for specific systems, although such comparisons are more difficult when idealized model systems are used in the computational studies.

Maybe surprisingly, there are only relatively few examples of simulations of a larger numbers of interacting macromolecules with residue-level, or super-residue level coarse-grained molecular representations [119,159,175,176]. This is presumably a consequence of the commonly used coarse-grained models generally not being parameterized for representing protein–protein interactions. In particular, most coarse-grained models discussed in this section do not pay sufficient attention to electrostatic and solvation interactions that become important in balancing protein–protein interactions in crowded environments. The examples where protein–protein mixtures were simulated at the coarse-grained level included Debye–Hückel type treatments [159], attention to molecular dipoles [175], coarse-grained sites with effective charges interacting through a fast Poisson–Boltzmann implicit solvent representation [176] or the use of a higher-resolution coarse-grained model embedded in Generalized Born-type implicit solvent [119]. Given the success of approaching cellular scales with BD simulations where the biomolecules are typically represented in atomistic detail, there is clearly an opportunity to develop new coarse-grained models that focus on protein–protein interactions within a Langevin or BD framework to extend the scales of current simulations to cellular scales while maintaining a faithful representation of molecular shapes. First steps in that direction have begun to appear [182] and more efforts along those lines will likely follow.

3.4. Spherical coarse-grained models of solutes in cellular environments

Spherical models of macromolecules in cellular environments represent the next level of approximation. They lose information about specific molecular shapes, but are especially attractive in the context of BD simulations. Featureless spheres are computationally very effective because of their rotational invariance. However, more complex models where the spheres are decorated with attractive

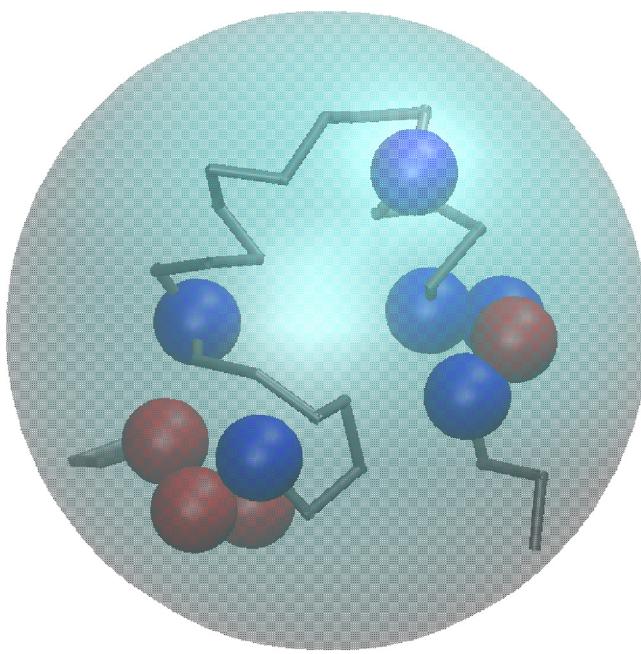


Fig. 6. Spherical model of a biomolecule with embedded charges for basic (blue) and acidic (red) amino acids. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and repulsive regions or embedded charges on the spherical surface, so-called patchy colloids [183] are more realistic (see Fig. 6). Interactions between spherical models of biomolecules generally involve three types of terms: repulsive interactions to avoid overlap of the spheres, attractive dispersion interactions, and electrostatic interactions.

Repulsive interactions may be modeled by treating the spheres either as hard spheres or by using a Lennard–Jones type potential. Dispersive interactions are typically modeled using $1/r^6$ potentials. In the simplest of such potentials proposed by Carlsson et al. hard-spheres may be combined with $1/r^6$ interactions [184,185]:

$$u_{ij}(r_{ij}) = \begin{cases} \infty, & r_{ij} < 2R \\ -\varepsilon \left(\frac{R}{r_{ij}}\right)^6, & r_{ij} > 2R \end{cases} \quad (4)$$

where i and j are different interacting particles, r_{ij} is the distance between their centers, R is their radius and ε scales the strength of the dispersive interactions. Because the hard-sphere repulsion causes numerical problems in MD simulations, softer forms are sometimes being used. Some studies use a standard Lennard–Jones potential with a $1/r^{12}$ potential [186,187], but because this may be too soft for preventing overlap of large molecular spheres, steeper soft-core potentials have been used in some studies to combine numerical convenience with sufficient repulsion [188–190].

An alternative formulation for the attractive dispersion interaction in a dielectric medium is given by the venerable Derjaguin–Landau–Verwey–Overbeek (DLVO) model [191] that continues to be applied in recent studies [188,192]:

$$u_{ij}(r_{ij}) = -\frac{A_H}{12} \left[\frac{R^2}{r_{ij}^2} + \frac{R^2}{r_{ij}^2 - R^2} + 2 \ln \frac{r_{ij}^2 - R^2}{r_{ij}^2} \right] \quad (5)$$

where r_{ij} is again the distance between two particles i and j , R is their effective radius and A_H is the so-called Hamaker constant to scale the strength of the dispersive interaction.

Finally, in the pioneering cytoplasm simulation by Bicout and Field, a somewhat different approach was taken [5]: Lennard–Jones

potentials were integrated over spherical volumes considering each atom type to be homogeneously distributed with average densities. Because such an explicit integration is computationally demanding, integrals were pre-calculated and a spline-interpolated potential was used during the actual simulations. This model implicitly considered different atom compositions of different particles in the simulated cytoplasmic model (proteins vs. tRNA and ribosomes).

Electrostatic interactions between spherical macromolecules are typically calculated in either one of two ways, both essentially resulting from considering charged particles in a dielectric continuum with additional screening due to the presence of mobile ionic species. Such a model is described by the Poisson–Boltzmann equation and gives rise to the Debye–Hückel model of electrostatic interactions in such environments.

In the first model, spheres are assumed to have a homogeneously distributed charge. In that case, the DLVO theory provides the following estimate for electrostatic interactions [191]:

$$u_{ij}(r_{ij}) = \frac{1}{4\pi\varepsilon_0\varepsilon} \left[\frac{Ze}{1 + \kappa R/2} \right]^2 \frac{\exp[-\kappa(r_{ij} - R)]}{r_{ij}} \quad (6)$$

where r_{ij} and R are the same quantities as in Eq. (5), ε is the dielectric constant of the environment, Z is the net charge of a given macromolecule, and κ is the inverse Debye screening length which depends on the salt concentration of the environment.

A computationally more expensive but somewhat more realistic alternative is to embed explicit charges within the sphere and calculate electrostatic interactions with a simple Debye–Hückel model [188–190,193]:

$$u_{ij} = \sum_{\alpha}^{N_i} \sum_{\beta}^{N_j} \frac{Z_{\alpha} Z_{\beta} e^2}{4\pi\varepsilon_0\varepsilon} \frac{\exp(-\kappa r_{\alpha\beta})}{r_{\alpha\beta}} \quad (7)$$

where the interaction between two spheres i and j now involves a double sum over the embedded charge pairs α and β in the respective spheres. In simpler models, some studies have also used Eq. (7) with just a single charge at the center of the spherical molecule [186,194].

Another approach pursued by Truskett et al. involves a multi-scale strategy where a given protein model is first simulated at a coarse-grained level to obtain information about the frequency and strength of interactions between folded and unfolded states of the system. The interactions at the coarse-grained level are then translated into an effective potential at the level of spherical particles using a modified Lennard–Jones type potential [195–198]. This is an interesting strategy because it suggests a path for connecting coarse-grained simulations described in the previous section with the spherical models described here.

Simulations using spherical models of biomolecules typically aim at describing phase behavior and diffusive properties of dense protein solutions [102,187–190,192,193,199]. Structure factors extracted from such simulations can be compared with small-angle neutron scattering experiments [189] or more general phase properties such as liquid–liquid coexistence curves or virial coefficients [193]. Most past studies focus on idealized systems with just one type of protein but it is certainly possible to examine more heterogeneous systems that approximate cellular environments more closely as in the early study by Bicout and Field [5]. Because the spherical models described here are computationally very efficient yet retain a physical foundation and include specific properties of different macromolecules through varying sizes and embedded charge distributions, it would be easy to imagine extending such studies to much more complex systems.

3.5. Reaction–diffusion models

The models described so far can describe biomolecules in cellular environments at different levels of complexity. Their focus is on conformational stability, interactions, and diffusive properties and the most coarse-grained models are clearly positioned to reach scales of whole cells and approach biologically most relevant time scales of milliseconds or longer. However, they lack in describing reactive processes without which comprehensive cellular models would be incomplete. The term ‘reactive processes’ is used here primarily to describe bond-breaking processes such as the chemical reactions catalyzed by biological enzymes, but may include any molecular process with distinct states that are separated by significant kinetic barriers, such as conformational switching of biomolecules, formation of complexes, binding of transcription factors to DNA etc.

At the most detailed, atomistic level, conformational switching and complex formation can be described well. Furthermore, chemical reactions via quantum mechanics (QM) or mixed quantum mechanics/molecular mechanics (QM/MM) formalisms [200]. However, the associated costs are prohibitive to extend QM or QM/MM studies to cellular scales other than by estimating kinetic parameters that can be used in more approximate models [201].

Reactions at cellular scales are typically modeled as kinetic processes. The simplest models involve ordinary differential equations (ODEs) to capture the kinetics of converting products to reactants described by their concentrations according to presumed rate laws with kinetic rate constants as empirical input. These models neglect space and assume that all of the system components are always well-mixed and homogeneously distributed [202]. They also assume that finite-size effects, that arise when reactants, products, or enzymes are present at very small concentrations, are negligible. The latter concern can be addressed by using a variety of stochastic approaches [203], but a remaining assumption of overall homogeneity makes it difficult to connect such models to the particle-based, spatially explicit models discussed in the previous sections of this review.

More complex models of reactive processes in cellular environments involve particle-based models and couple diffusion with reaction. One approach is to model diffusing particles via BD simulations and to allow reactions with a certain probability once suitable reaction partners collide [194,204–206]. This approach has been used, for example, in a recent simulation of a model cytoplasm of *E. coli* [207]. A main challenge with such simulations is the efficient and accurate determination of when exactly a reaction-capable collision occurs and what the probability is for the reaction to go forward at that point [204,205]. Another challenge is that many reactions occur on seconds or longer time scales so that only very simple spherical, featureless models of interaction molecules are typically considered within a framework of continuous space BD simulations, or further approximations are made by considering lattice-based representations [208–210].

Additional acceleration of reaction–diffusion simulation is gained by employing discrete event simulation techniques that are continuous in time, instead of using discrete time stepping, by essentially predicting which particles to move according to their likelihood of leading to a successful reaction [103,150,204,211–216]. This is the basic idea of the Gillespie algorithm [216] which generates an ordered list of reactions to occur next according to exponentially distributing waiting times for first order reactions. A more recent method is based on the Green’s function reaction dynamics framework [214,217–221] which solves the probability of reactions to occur for diffusion-limited reactions according to the Fokker–Planck equation. While event-driven methods can greatly accelerate reaction–diffusion simulations, one problem is that they become ineffective when

high concentrations of interacting particles are included because of a high rate of molecular collisions. Thus, these types of models are more typically applied to selected cellular components within an otherwise implicitly represented cellular environment, for example by choosing appropriate diffusion constants [204,214].

Reaction–diffusion simulations generally aim at matching and interpreting experimental data collected on macroscopic time scales, e.g. fluctuations of certain metabolites in response to certain stimuli or gene products resulting from variable transcriptional regulation. This necessitates models that are rather simplistic from a physical perspective and emphasizes empirical models that rely on measured or estimated data such as diffusion constants or kinetic parameters. Since such parameters are generally not available at sufficiently high accuracy for a given system, a significant part of building models for reaction–diffusion simulations consists of adjusting model parameters until agreement with experiment is reached. This presents an opportunity to integrate physically more detailed modeling into such studies in order to develop models that are less dependent on the data that they aim to interpret and are more robust for predicting the outcome of new experiments. This will be discussed in more detail in the next section.

4. Toward comprehensive cellular models

4.1. Dreaming the dream

Atomistic models provide the highest level of detail when studying biomolecules in cellular environments. If one considers the smallest known bacterium, *M. genitalium*, the exploration of a cytoplasmic model could be possible with the ever-increasing computer power at large supercomputer installations. The setup of such a system requires knowledge of what components are present at which concentrations and structural information for all of the involved macromolecules. Which components to include can be deduced from gene annotations, metabolic network reconstructions, and proteomics and metabolomics. Indeed, a complete network reconstruction was recently published for *M. genitalium* [61]. Furthermore, there are extensive data from proteomics and metabolomics studies of a closely related organism, *M. pneumoniae*, that provide qualitative and quantitative information about what molecular components are present at what concentrations [59,222,223].

An additional challenge is how to obtain detailed structural information for all of the molecular constituents. Experimental structures are only available for a few gene products of *M. genitalium*, but most other (soluble) protein structures can be obtained via homology modeling. It is possible to integrate all of this information into a cytoplasmic model as shown in Fig. 7. This model contains proteins, nucleic acids, metabolites, ions, and explicit water. The dimensions of this system are 100 nm × 100 nm × 100 nm covering about a tenth of the volume of an entire *M. genitalium* bacterium with a diameter of about 300–400 nm. The number of atoms contained in such a volume is about 100 M. While such a system greatly exceeds the size of typical MD simulations, sufficiently large allocations on the latest generation of highly parallel supercomputers would allow for such a system to be simulated for tens to hundreds of nanoseconds while microseconds may be possible with the next generation of supercomputers. Even larger systems up to an entire cell may be conceivable as well, however, apart from the even higher computational costs, there are other challenges with respect to modeling the structures of DNA and membrane proteins, both of which would have to be included at that point.

Such large-scale simulations would yield the most realistic information about the structure and dynamics of biomolecules in cellular environments but would still not be able to cover many

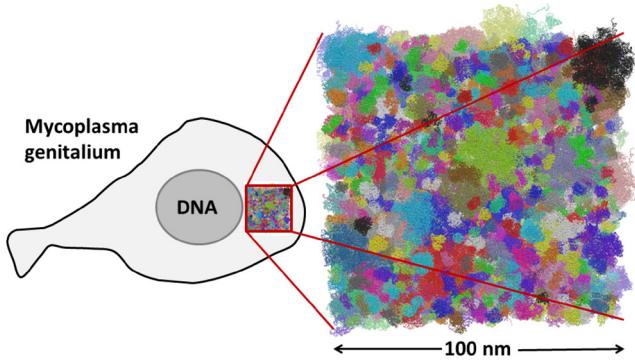


Fig. 7. Detailed cytoplasmic model of *Mycoplasma genitalium*. Proteins and nucleic acids are shown with each complex in a different color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

biological processes because of limited simulation time scales and miss entirely any reactive processes.

4.2. Multi-scale models

A more realistic proposition to reach biological time scales in simulations of cellular environments while maintaining detailed physical models is to employ multi-scale approaches. Multi-scale methods may either include different parts of a system at different levels of resolution simultaneously or use information from higher resolution simulations to inform the construction of models at lower resolution [224–227]. The latter strategy is commonly used in the development of interaction potentials for coarse-grained models [228]. Much of the past focus has been on intramolecular interactions but it is straightforward to extend such an approach to intermolecular interactions to develop new coarse-grained models that represent the energetics of protein–protein interactions better. Furthermore, we have given examples above where new results from atomistic can inform the model building at lower resolution levels. For example, recent results from our group suggest that a value between 20 and 60 rather than 80 [96] may be a more appropriate choice for the dielectric constant in implicit solvent models of the cellular aqueous environment [119,229]. Other detailed insight from atomistic simulations involves the variation of diffusive motion in the presence of different molecular crowders and on different time scales. While still anecdotal at this point, such data could lead to a better model of diffusive motion in BD simulations and especially reaction–diffusion type models that depend on accurate estimates of the diffusive behavior of biomolecules to be able to make reliable quantitative predictions.

The simultaneous representation of different parts at different levels of detail is used in essentially all studies where the stability of biomolecules in the presence of crowders has been studied. A common strategy is to represent the molecules of interest in relatively high level of detail but model the crowders at a much lower level. Typical combinations are coarse-grained solutes with spherical crowders [168] or atomistic solutes with coarse-grained crowders [119]. Such a strategy could be expanded to larger-scale simulations to include selected biomolecules or selected regions of a cellular model in higher level of detail than other regions. A further possibility is the implicit modeling of not just the solvent but also the entire cellular environment. Reduced dielectric constants may be a first order approximation of the crowded environment in cases or regions where detailed protein–protein interactions are not critical [118,119]. Furthermore, implicit models could also be used to model membrane environments with little additional effort [115,230].

4.3. Reactions

The real challenge is how to maintain a physically realistic model and include an ability to model reactive processes on biologically relevant time scales. Reaction–diffusion models provide a very successful framework for describing reactive processes but generally depend heavily on empirical parameters and fitting to experimental data. Increasing the level of physical detail in such models appears to be one strategy for combining the biophysical and systems biology views of cellular environments. An obvious first step would be to start with spatially explicit, reaction–diffusion models and replace simple spherical models with more sophisticated coarse-grained representations, either patchy colloid-type more shape-preserving models. This would lead to a more realistic account of biomolecular interactions and better predictions of diffusive behavior and probabilities of forming productive reaction complex encounters. Going to higher levels of detail, however, precludes efficient methods like GFRD to be used that rely on the simplicity of spherical particles. However, the loss of efficiency for highly crowded environments, where highly predictive reaction-centered methods become less effective would be reduced. The consequence of including biomolecules at a coarse-grained instead of simple spherical level would significantly increase the computational cost. However, given that BD simulations of atomistic proteins on cytoplasmic scales over tens of microseconds are already possible [7], simulations of simpler models over millisecond to second scales with coarse-grained models in a reaction–diffusion simulation framework appear to be with the realm of feasibility.

One specific issue with BD simulations is the need to maintain solutes as rigid bodies so that sufficiently long time steps can be used. This approximation makes it difficult to capture biomolecular conformational flexibility and in particular the sampling of alternate states that may be of functional relevance. One possibility to address this challenge is to allow conformations to switch stochastically [231]. This strategy could also be used in higher-resolution BD dynamics that do not include reactions to improve the level of realism. Insight into which states should be sampled stochastically could again come from detailed atomistic simulations that would be run first to generate a database of conformations with proper weights that is then drawn from during BD simulations.

Finally, another significant challenge is the estimate of kinetic rate constants in reaction–diffusion models. Kinetic rate constants are often the least reliable experimental data and the subject of most of the empirical adjustments in such models to match experimental data. Binding rate constants could also be estimated from atomistic simulations [201] if they do not result implicitly from the dynamics of interaction solutes. What are more difficult to obtain are rate constants for chemical reactions. That is in principle also possible using computational methods [232,233], but it would be rather tedious and expensive to carry out such calculations for a large number of different enzymes and reactions. There is hope, however, that this may not be necessary given recent efforts to obtain kinetic rate constants in a high-throughput and consistent manner that should result in more reliable data [234,235].

5. Summary

In this review, recent modeling efforts to study the structure and dynamics of biomolecules in the cellular context are contrasted and discussed in the context of providing a deeper understanding of cellular environments that includes a variety of different aspects ranging from crowding effects to biochemical reaction networks. It is clear that much progress has been made in modeling different levels of detail and processes at different

spatial or temporal scales. However, there is still a relative lack of integration between scales that is necessary to provide a truly comprehensive view. There appear to be many opportunities, however, to further develop methods that can bridge between scales and that, we believe, will ultimately lead to fully integrated cellular models that are both physically accurate and biologically relevant and thus can serve as a platform for in silico cellular models with broad practical applications. Another opportunity is better integration of modeling with emerging experimental techniques that are increasingly able to provide high-resolution imaging of cellular systems, such as those based on small-angle X-ray scattering (SAXS)[236], free-electron X-ray lasers (XFEL)[237] and super-resolution fluorescence microscopy [238].

Funding

NIH GM092949, NIH GM084953, RIKEN QBIC, MEXT SPIRE

Acknowledgment

Funding from RIKEN-QBIC, NIH GM092949 and NIH GM084953 (to MF) and by MEXT SPIRE Supercomputational Life Science (to YS) is acknowledged. Computer resources were used at RIKEN-RICC (RIKEN Integrated Cluster of Clusters).

References

- [1] C.A. Taft, V.B. Da Silva, C. Da Silva, Current topics in computer-aided drug design, *J. Pharm. Sci.*—US 97 (2008) 1089–1098.
- [2] G.R. Marshall, Computer-aided drug design, *Annu. Rev. Pharmacol. Toxicol.* 27 (1987) 193–213.
- [3] D.W. Borhani, D.E. Shaw, The future of molecular dynamics simulations in drug discovery, *J. Comput.-Aided Mol. Des.* 26 (2012) 15–26.
- [4] D.S. Goodsell, Inside a living cell, *Trends Biochem. Sci.* 16 (1991) 203–206.
- [5] D.J. Bicout, M.J. Field, Stochastic dynamics simulations of macromolecular diffusion in a model of the cytoplasm of *Escherichia coli*, *J. Phys. Chem.* 100 (1996) 2489–2497.
- [6] D. Ridgway, G. Broderick, A. Lopez-Campistrous, M. Ruaini, P. Winter, M. Hamilton, et al., Coarse-grained molecular simulation of diffusion and reaction kinetics in a crowded virtual cytoplasm, *Biophys. J.* 94 (2008) 3748–3759.
- [7] S.R. McGuffee, A.H. Elcock, Diffusion crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm, *PLoS Comput. Biol.* 6 (2010) e1000694.
- [8] T. Ando, J. Skolnick, Crowding and hydrodynamic interactions likely dominate *in vivo* macromolecular motion, *PNAS* 107 (2010) 18457–18462.
- [9] A. Vendeville, D. Lariviere, E. Fourmentin, An inventory of the bacterial macromolecular components and their spatial organization, *FEMS Microbiol. Rev.* 35 (2011) 395–414.
- [10] J.S. Clegg, Properties and metabolism of the aqueous cytoplasm and its boundaries, *Am. J. Physiol.—Reg. I.* 246 (1984) R133–R151.
- [11] C. Hunte, S. Richers, Lipids and membrane protein structures, *Curr. Opin. Struct. Biol.* 18 (2008) 406–411.
- [12] Z. Gitai, The new bacterial cell biology: moving parts and subcellular architecture, *Cell* 120 (2005) 577–586.
- [13] A. Hudler, L. Nathanson, M.P. Deutscher, Organization of mammalian cytoplasm, *Mol. Cell. Biol.* 23 (2003) 9318–9326.
- [14] S. Ebbinghaus, M. Gruebele, Protein folding landscapes in the living cell, *J. Phys. Chem. Lett.* 2 (2011) 314–319.
- [15] H.-X. Zhou, G. Rivas, A.P. Minton, Macromolecular crowding confinement: biochemical biophysical and potential physiological consequences, *Annu. Rev. Biophys.* 37 (2008) 375–397.
- [16] H.X. Zhou, Protein folding in confined and crowded environments, *Arch. Biochem. Biophys.* 469 (2008) 76–82.
- [17] H.-X. Zhou, Influences of crowded cellular environments on protein folding, binding, and oligomerization: biological consequences and potentials of atomistic modeling, *FEBS Lett.* 587 (2013) 1053–1061.
- [18] A.C. Miklos, M. Sarkar, Y. Wang, G.J. Pielak, Protein crowding tunes protein stability, *J. Am. Chem. Soc.* 133 (2011) 7116–7120.
- [19] M.M. Waegele, F. Gai, Power-law dependence of the melting temperature of ubiquitin on the volume fraction of macromolecular crowders, *J. Chem. Phys.* 134 (2011) 095104.
- [20] M. Hayter-Hartl, A.P. Minton, A simple semiempirical model for the effect of molecular confinement upon the rate of protein folding, *Biochemistry* 45 (2006) 13356–13360.
- [21] A.P. Minton, Excluded volume as a determinant of macromolecular structure and reactivity, *Biopolymers* 20 (1981) 2093–2120.
- [22] A. Christiansen, Q. Wang, M.S. Cheung, P. Wittung-Stafshede, Effects of macromolecular crowding agents on protein folding *in vitro* and *in silico*, *Biophys. Rev.* 5 (2013) 137–145.
- [23] F. Bergasa-Caceres, H.A. Rabitz, A simple quantitative model of macromolecular crowding effects on protein folding: application to the murine prion protein (121–231), *Chem. Phys. Lett.* 574 (2013) 112–115.
- [24] M.B. Gee, P.E. Smith, Kirkwood–Buff theory of molecular and protein association, aggregation, and cellular crowding, *J. Chem. Phys.* 131 (2009) 165101.
- [25] S.B. Qin, L. Cai, H.X. Zhou, A method for computing association rate constants of atomistically represented proteins under macromolecular crowding, *Phys. Biol.* 9 (2012) 066008.
- [26] D. Shah, A.L. Tan, V. Ramakrishnan, J.W. Jiang, R. Rajagopalan, Effects of polydisperse crowders on aggregation reactions: a molecular thermodynamic analysis, *J. Chem. Phys.* 134 (2011) 064704.
- [27] J. Wilf, A.P. Minton, Evidence for protein self-association induced by excluded volume myoglobin in the presence of globular-proteins, *Biochim. Biophys. Acta* 670 (1981) 316–322.
- [28] A.P. Minton, Confinement as a determinant of macromolecular structure and reactivity, *Biophys. J.* 63 (1992) 1090–1100.
- [29] Q. Wang, K.C. Liang, A. Czader, M.N. Waxham, M.S. Cheung, The effect of macromolecular crowding ionic strength and calcium binding on calmodulin dynamics, *PLoS Comput. Biol.* 7 (2011) e1002114.
- [30] E.A. Cino, M. Karttunen, W.Y. Choy, Effects of molecular crowding on the dynamics of intrinsically disordered proteins, *PLoS One* 7 (2012) e49876.
- [31] A. Kudlay, M.S. Cheung, D. Thirumalai, Crowding effects on the structural transitions in a flexible helical homopolymer, *Phys. Rev. Lett.* 102 (2009) 118101.
- [32] R. Harada, N. Tochio, T. Kigawa, Y. Sugita, M. Feig, Reduced native state stability in crowded cellular environment due to protein–protein interactions, *J. Am. Chem. Soc.* 135 (2013) 3696–3701.
- [33] M. Feig, Y. Sugita, Variable interactions between protein crowders and biomolecular solutes are important in understanding cellular crowding, *J. Phys. Chem. B* 116 (2012) 599–605.
- [34] Y. Zhang, J. Skolnick, The protein structure prediction problem could be solved using the current PDB library, *PNAS* 102 (2005) 1029–1034.
- [35] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779–815.
- [36] D. Kihara, Y. Zhang, H. Lu, A. Kolinski, J. Skolnick, Ab initio protein structure prediction to a genomic scale: application to the mycoplasma genitalium genome, *PNAS* 99 (2002) 5993–5998.
- [37] A. Elofsson, G. von Heijne, Membrane protein structure: prediction versus reality, *Annu. Rev. Biochem.* 76 (2007) 125–140.
- [38] A. Engel, H.E. Gaub, Structure and mechanics of membrane proteins, *Annu. Rev. Biochem.* 77 (2008) 127–148.
- [39] A.K. Dunker, C.J. Brown, J.D. Lawson, L.M. Iakoucheva, Z. Obradovic, Intrinsic disorder and protein function, *Biochemistry* 41 (2002) 6573–6582.
- [40] A. Vitalis, R.V. Pappu, ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions, *J. Comput. Chem.* 30 (2009) 673–699.
- [41] S. Rauscher, R. Pomes, Molecular simulations of protein disorder, *Biochem. Cell Biol.* 88 (2010) 269–290.
- [42] K. Lindorff-Larsen, N. Trbovic, P. Maragakis, S. Piana, D.E. Shaw, Structure dynamics of an unfolded protein examined by molecular dynamics simulation, *J. Am. Chem. Soc.* 134 (2012) 3787–3791.
- [43] M. Thanbichler, S.C. Wang, L. Shapiro, The bacterial nucleoid: a highly organized and dynamic structure, *J. Cell. Biochem.* 96 (2005) 506–521.
- [44] J.D. Watson, F.H. Crick, Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid, *Nature* 171 (1953) 737.
- [45] M.F. Perutz, M.G. Rossmann, A.F. Cullis, H. Muirhead, G. Will, A.C.T. North, Structure of haemoglobin—3-dimensional Fourier synthesis at 5.5-A resolution obtained by X-ray analysis, *Nature* 185 (1960) 416–422.
- [46] J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, D.C. Phillips, 3-dimensional model of the myoglobin molecule obtained by X-ray analysis, *Nature* 181 (1958) 662–666.
- [47] M. Macvanin, S. Adhya, Architectural organization in *E. coli* nucleoid, *Biochim. Biophys. Acta—Gene Regul. Mech.* 1819 (2012) 830–835.
- [48] M. Thanbichler, L. Shapiro, Chromosome organization and segregation in bacteria, *J. Struct. Biol.* 156 (2006) 292–303.
- [49] L. Postow, C.D. Hardy, J. Arsuaga, N.R. Cozzarelli, Topological domain structure of the *Escherichia coli* chromosome, *Genes Dev.* 18 (2004) 1766–1779.
- [50] K. van Holde, J. Zlatanova, Chromatin fiber structure: where is the problem now? *Semin. Cell Dev. Biol.* 18 (2007) 651–658.
- [51] A. Pastore, P.A. Temussi, The two faces of Janus: functional interactions and protein aggregation, *Curr. Opin. Struct. Biol.* 22 (2012) 30–37.
- [52] A.L. Gnatt, P. Cramer, J. Fu, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution, *Science* 292 (2001) 1876–1882.
- [53] N. Ban, P. Nissen, J. Hansen, P.B. Moore, T.A. Steitz, The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution, *Science* 289 (2000) 905–920.
- [54] F. Tama, O. Miyashita, C.L. Brooks, Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis, *J. Mol. Biol.* 337 (2004) 985–999.
- [55] P. Chacon, W. Wriggers, Multi-resolution contour-based fitting of macromolecular structures, *J. Mol. Biol.* 317 (2002) 375–384.

- [56] W. Jiang, M.L. Baker, S.J. Lutcke, W. Chiu, Bridging the information gap: computational tools for intermediate resolution structure interpretation, *J. Mol. Biol.* 308 (2001) 1033–1044.
- [57] M. Pilhofer, M.S. Ladinsky, A.W. McDowall, G.J. Jensen, Bacterial TEM: new insights from cryo-microscopy, in: T. MüllerReichert (Ed.), *Electron Microscopy of Model Systems*, 96, Elsevier Academic Press Inc, San Diego, 2010, pp. 21–45.
- [58] Z. Li, G.J. Jensen, Electron cryotomography: a new view into microbial ultrastructure, *Curr. Opin. Microbiol.* 12 (2009) 333–340.
- [59] S. Kuhnert, V. van Noort, M.J. Betts, A. Leo-Macias, C. Batisse, M. Rode, et al., Proteome organization in a genome-reduced bacterium, *Science* 326 (2009) 1235–1240.
- [60] J.J. Tyson, B. Novak, Functional motifs in biochemical reaction networks, *Annu. Rev. Phys. Chem.* 61 (2010) 219–240.
- [61] J.R. Karr, J.C. Sanghvi, D.N. Macklin, M.V. Gutschow, J.M. Jacobs, B. Bolival, et al., A whole-cell computational model predicts phenotype from genotype, *Cell* 150 (2012) 389–401.
- [62] J.R. Karr, J.C. Sanghvi, D.N. Macklin, A. Arora, M.W. Covert, WholeCellKB: model organism databases for comprehensive whole-cell models, *Nucleic Acids Res.* 41 (2013), D787–D92.
- [63] J.A. Dix, A.S. Verkman, Crowding effects on diffusion in solutions and cells, *Annu. Rev. Biophys.* 37 (2008) 247–263.
- [64] P.C. Bressloff, J.M. Newby, Stochastic models of intracellular transport, *Rev. Mod. Phys.* 85 (2013) 135–196.
- [65] A.H. Elcock, Models of macromolecular crowding effects and the need for quantitative comparisons with experiment, *Curr. Opin. Struct. Biol.* 20 (2010) 196–206.
- [66] T. Frembgen-Kesner, A.H. Elcock, Computer simulations of the bacterial cytoplasm, *Biophys. Rev.* 5 (2013) 109–119.
- [67] A. Bhattacharya, R.B. Best, J. Mittal, Smoothing of the GB1 hairpin folding landscape by interfacial confinement, *Biophys. J.* 103 (2012) 596–600.
- [68] J.S. Rao, L.R. Cruz, Effects of confinement on the structure and dynamics of an intrinsically disordered peptide: a molecular dynamics study, *J. Phys. Chem. B.* 117 (2013) 3707–3719.
- [69] J.H. Tian, A.E. Garcia, Simulation studies of protein folding/unfolding equilibrium under polar and nonpolar confinement, *J. Am. Chem. Soc.* 133 (2011) 15157–15164.
- [70] K.S. Harve, R. Lareu, R. Rajagopalan, M. Raghunath, Understanding how the crowded interior of cells stabilizes DNA/DNA and DNA/RNA hybrids-in silico predictions and in vitro evidence, *Nucleic Acids Res.* 38 (2010) 172–181.
- [71] Z.J. Tan, S.J. Chen, Ion-mediated RNA structural collapse: effect of spatial confinement, *Biophys. J.* 103 (2012) 827–836.
- [72] J.H. Tian, A.E. Garcia, Simulations of the confinement of ubiquitin in self-assembled reverse micelles, *J. Chem. Phys.* 134 (2011) 225101.
- [73] N.A. Kurniawan, S. Enemark, R. Rajagopalan, Crowding alters the folding kinetics of a beta-hairpin by modulating the stability of intermediates, *J. Am. Chem. Soc.* 134 (2012) 10200–10208.
- [74] S.B. Qin, D.D.L. Minh, J.A. McCammon, H.X. Zhou, Method to predict crowding effects by postprocessing molecular dynamics trajectories: application to the flap dynamics of HIV-1 protease, *J. Phys. Chem. Lett.* 1 (2010) 107–110.
- [75] H. Dong, S.B. Qin, H.X. Zhou, Effects of macromolecular crowding on protein conformational changes, *PLoS Comput. Biol.* 6 (2010) e1000833.
- [76] S.B. Qin, H.X. Zhou, Atomistic modeling of macromolecular crowding predicts modest increases in protein folding and binding stability, *Biophys. J.* 97 (2009) 12–19.
- [77] H. Tjong, H.X. Zhou, The folding transition-state ensemble of a four-helix bundle protein: helix propensity as a determinant and macromolecular crowding as a probe, *Biophys. J.* 98 (2010) 2273–2280.
- [78] J.S. Kim, A. Yethiraj, Crowding effects on protein association: effect of interactions between crowding agents, *J. Phys. Chem. B.* 115 (2011) 347–353.
- [79] B.P. Cossins, M.P. Jacobson, V. Guallar, A new view of the bacterial cytosol environment, *PLoS Comput. Biol.* 7 (2011) e1002066.
- [80] S.S. Cho, G. Reddy, J.E. Straub, D. Thirumalai, Entropic stabilization of proteins by TMAO, *J. Phys. Chem. B.* 115 (2011) 13401–13407.
- [81] E.P. O'Brien, G. Ziv, G. Haran, B.R. Brooks, D. Thirumalai, Effects of denaturants and osmolytes on proteins are accurately predicted by the molecular transfer model, *PNAS* 105 (2008) 13403–13408.
- [82] C.Y. Hu, H. Kokubo, G.C. Lynch, D.W. Bolen, B.M. Pettitt, Backbone additivity in the transfer model of protein solvation, *Protein Sci.* 19 (2010) 1011–1022.
- [83] C.Y. Hu, G.C. Lynch, H. Kokubo, B.M. Pettitt, Trimethylamine-N-oxide influence on the backbone of proteins: an oligoglycine model, *Proteins* 78 (2010) 695–704.
- [84] Z. Xia, P. Das, E.I. Shakhnovich, R.H. Zhou, Collapse of unfolded proteins in a mixture of denaturants, *J. Am. Chem. Soc.* 134 (2012) 18266–18274.
- [85] R. Zangi, R. Zhou, B.J. Berne, Urea's action on hydrophobic interactions, *J. Am. Chem. Soc.* 131 (2009) 1535–1541.
- [86] J. Heyda, M. Kozisek, L. Bednarova, G. Thompson, J. Konvalinka, J. Vondrasek, et al., Urea and guanidinium induced denaturation of a Trp-Cage miniprotein, *J. Phys. Chem. B.* 115 (2011) 8910–8924.
- [87] W.K. Lim, J. Rosgen, S.W. Englander, Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group, *PNAS* 106 (2009) 2595–2600.
- [88] M. Auton, D.W. Bolen, J. Rosgen, Structural thermodynamics of protein preferential solvation: osmolyte solvation of proteins, aminoacids, and peptides, *Proteins* 73 (2008) 802–813.
- [89] V. Pierce, M. Kang, M. Aburi, S. Weerasinghe, P.E. Smith, Recent applications of Kirkwood–Buff theory to biological systems, *Cell Biochem. Biophys.* 50 (2008) 1–22.
- [90] L. Hua, R. Zhou, D. Thirumalai, B.J. Berne, Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding, *PNAS* 105 (2008) 16928–16933.
- [91] H. Kokubo, B.M. Pettitt, Preferential solvation in urea solutions at different concentrations: properties from simulation studies, *J. Phys. Chem. B.* 111 (2007) 5233–5242.
- [92] M.E. Lee, N.F.A. van der Vegt, Does urea denature hydrophobic interactions? *J. Am. Chem. Soc.* 128 (2006) 4948–4949.
- [93] A.S. Thomas, A.H. Elcock, Direct observation of salt effects on molecular interactions through explicit-solvent molecular dynamics simulations: differential effects on electrostatic and hydrophobic interactions and comparisons to Poisson–Boltzmann theory, *J. Am. Chem. Soc.* 128 (2006) 7796–7806.
- [94] A.S. Thomas, A.H. Elcock, Direct measurement of the kinetics and thermodynamics of association of hydrophobic molecules from molecular dynamics simulations, *J. Phys. Chem. Lett.* 2 (2011) 19–24.
- [95] E. Rivera, J. Straub, D. Thirumalai, Sequence and crowding effects in the aggregation of a 10-residue fragment derived from islet amyloid polypeptide, *Biophys. J.* 96 (2009) 4552–4560.
- [96] R. Harada, Y. Sugita, M. Feig, Protein crowding affects hydration structure and dynamics, *J. Am. Chem. Soc.* 134 (2012) 4842–4849.
- [97] S.B. Zimmerman, A.P. Minton, Macromolecular crowding—biochemical and biophysical, and physiological consequences, *Annu. Rev. Biophys. Biomol. Struct.* 22 (1993) 27–65.
- [98] K. Inomata, A. Ohno, H. Tochio, S. Isogai, T. Tenno, I. Nakase, et al., High-resolution multi-dimensional NMR spectroscopy of proteins in human cells, *Nature* 458 (2009) 106–109.
- [99] F. Despa, D.P. Orgill, R.C. Lee, Molecular crowding effects on protein stability, *Ann. N.Y. Acad. Sci.* 1066 (2005) 54–66.
- [100] L.M. Charlton, C.O. Barnes, C. Li, J. Orans, G.B. Young, G.J. Pielak, Residue-level interrogation of macromolecular crowding effects on protein stability, *J. Am. Chem. Soc.* 130 (2008) 6826–6830.
- [101] Y. Wang, C. Li, G.J. Pielak, Effects of proteins on protein diffusion, *J. Am. Chem. Soc.* 132 (2010) 9392–9397.
- [102] T. Geyer, Mixing normal and anomalous diffusion, *J. Chem. Phys.* 137 (2012) 115101.
- [103] T.T. Marquez-Lago, A. Leier, K. Burrage, Anomalous diffusion and multi-fractional Brownian motion: simulating molecular crowding and physical obstacles in systems biology, *IET Syst. Biol.* 6 (2012) 134–142.
- [104] M.J. Saxton, Anomalous diffusion due to obstacles: a Monte Carlo study, *Biophys. J.* 66 (1994) 394–401.
- [105] M. Feig, J. Chocholousova, S. Tanizaki, Extending the horizon: towards the efficient modeling of large biomolecular complexes in atomic detail, *Theor. Chem. Acc.* 116 (2006) 194–205.
- [106] P. Ball, Water as an active constituent in cell biology, *Chem. Rev.* 108 (2008) 74–108.
- [107] M.K. Gilson, K.A. Sharp, B.H. Honig, Calculating the electrostatic potential of molecules in solution: method and error assessment, *J. Comput. Chem.* 9 (1987) 327–335.
- [108] F. Despa, A. Fernandez, R.S. Berry, Dielectric modulation of biological water, *Phys. Rev. Lett.* 93 (2004) 228104.
- [109] H. Tjong, H.X. Zhou, Prediction of protein solubility from calculation of transfer free energy, *Biophys. J.* 95 (2008) 2601–2609.
- [110] A. Nicholls, B. Honig, A rapid finite-difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation, *J. Comput. Chem.* 12 (1991) 435–445.
- [111] M. Feig, C.L. Brooks III, Recent advances in the development and application of implicit solvent models in biomolecule simulations, *Curr. Opin. Struct. Biol.* 14 (2004) 217–224.
- [112] D. Sitkoff, K.A. Sharp, B. Honig, Accurate calculation of hydration free-energies using macroscopic solvent models, *J. Phys. Chem.* 98 (1994) 1978–1988.
- [113] J.A. Waggoner, N.A. Baker, Assessing implicit models for nonpolar mean solvation forces: the importance of dispersion and volume terms *Proc. PNAS* 103 (2006) 8331–8336.
- [114] R.M. Levy, L.Y. Zhang, E. Gallicchio, A.K. Feltz, On the nonpolar hydration free energy of proteins: surface area and continuum solvent models for the solute–solvent interaction energy, *J. Am. Chem. Soc.* 125 (2003) 9523–9530.
- [115] S. Tanizaki, M. Feig, A generalized Born formalism for heterogeneous dielectric environments: application to the implicit modeling of biological membranes, *J. Chem. Phys.* 122 (2005) 124706.
- [116] W. Im, M. Feig, C.L. Brooks III, An implicit membrane generalized Born theory for the study of structure stability, and interactions of membrane proteins, *Biophys. J.* 85 (2003) 2900–2918.
- [117] T. Lazaridis, Effective energy function for proteins in lipid membranes, *Proteins* 52 (2003) 176–192.
- [118] S. Tanizaki, J.W. Clifford, B.D. Connelly, M. Feig, Conformational sampling of peptides in cellular environments, *Biophys. J.* 94 (2008) 747–759.
- [119] A.V. Predeus, S. Gul, S.M. Gopal, M. Feig, Conformational sampling of peptides in the presence of protein crowders from AA/CG-multiscale simulations, *J. Phys. Chem. B.* 116 (2012) 8610–8620.
- [120] M.K. Gilson, B.H. Honig, The dielectric-constant of a folded protein, *Biopolymers* 25 (1986) 2097–2119.
- [121] K.A. Sharp, B. Honig, Electrostatic interactions in macromolecules—theory and applications, *Annu. Rev. Biophys. Chem.* 19 (1990) 301–332.

- [122] J. Srinivasan, M.W. Trevathan, P. Beroza, D.A. Case, Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects, *Theor. Chim. Acta* 101 (1999) 426–434.
- [123] B.A. Persson, M. Lund, Association and electrostatic steering of alpha-lactalbumin-lysozyme heterodimers, *Phys. Chem. Chem. Phys.* 11 (2009) 8879–8885.
- [124] J.H. Chen, W.P. Im, C.L. Brooks, Balancing solvation and intramolecular interactions: toward a consistent generalized born force field, *J. Am. Chem. Soc.* 128 (2006) 3728–3736.
- [125] J.M. Deutch, I. Oppenhei, Molecular theory of Brownian motion for several particles, *J. Chem. Phys.* 54 (1971) 3547.
- [126] M. Feig, Kinetics from implicit solvent simulations of biomolecules as a function of viscosity, *J. Chem. Theory Comput.* 3 (2007) 1734–1748.
- [127] B. Zagrovic, V. Pande, Solvent viscosity dependence of the folding rate of a small protein: distributed computing study, *J. Comput. Chem.* 24 (2003) 1432–1436.
- [128] M. Drugosz, J. Trylska, Diffusion in crowded biological environments: applications of Brownian dynamics, *BMC Biophys.* (2011) 4.
- [129] D.L. Ermak, J.A. McCammon, Brownian dynamics with hydrodynamic interactions, *J. Chem. Phys.* 69 (1978) 1352–1360.
- [130] S.R. McGuffee, A.H. Elcock, Atomically detailed simulations of concentrated protein solutions: the effects of salt, pH, point mutations, and protein concentration in simulations of 1000-molecule systems, *J. Am. Chem. Soc.* 128 (2006) 12098–12110.
- [131] P. Mereghetti, R.R. Gabdoulline, R.C. Wade, Brownian dynamics simulation of protein solutions structural and dynamical properties, *Biophys. J.* 99 (2010) 3782–3791.
- [132] P. Mereghetti, R.C. Wade, Atomic detail Brownian dynamics simulations of concentrated protein solutions with a mean field treatment of hydrodynamic interactions, *J. Phys. Chem. B.* 116 (2012) 8523–8533.
- [133] J. Rotne, S. Prager, Variational treatment of hydrodynamic interaction in polymers, *J. Chem. Phys.* 50 (1969) 4831.
- [134] H. Yamakawa, Transport properties of polymer chains in dilute solution—hydrodynamic interaction, *J. Chem. Phys.* 53 (1970) 436.
- [135] J.F. Brady, G. Bossis, Stokesian dynamics, *Annu. Rev. Fluid. Mech.* 20 (1988) 111–157.
- [136] L. Durlofsky, J.F. Brady, G. Bossis, Dynamic simulation of hydrodynamically interacting particles, *J. Fluid Mech.* 180 (1987) 21–49.
- [137] T. Frembgen-Kesner, A.H. Elcock, Absolute protein–protein association rate constants from flexible coarse-grained Brownian dynamics simulations: the role of intermolecular hydrodynamic interactions in Barnase–Barstar association, *Biophys. J.* 99 (2010) L75–L77.
- [138] T. Frembgen-Kesner, A.H. Elcock, Striking effects of hydrodynamic interactions on the simulated diffusion and folding of proteins, *J. Chem. Theory Comput.* 5 (2009) 242–256.
- [139] T. Ando, J. Skolnick, On the importance of hydrodynamic interactions in lipid membrane formation, *Biophys. J.* 104 (2013) 96–105.
- [140] T. Ando, E. Chow, Y. Saad, J. Skolnick, Krylov subspace methods for computing hydrodynamic interactions in Brownian dynamics simulations, *J. Chem. Phys.* 137 (2012) 064106.
- [141] C.D. Keating, Aqueous phase separation as a possible route to compartmentalization of biological molecules, *Acc. Chem. Res.* 45 (2012) 2114–2124.
- [142] M. Beck, M. Topf, Z. Frazier, H. Tjong, M. Xu, S.H. Zhang, et al., Exploring the spatial and temporal organization of a cell's proteome, *J. Struct. Biol.* 173 (2011) 483–496.
- [143] P.L. Li, S. Banjade, H.C. Cheng, S. Kim, B. Chen, L. Guo, et al., Phase transitions in the assembly of multivalent signalling proteins, *Nature* 483 (2012) 336–340.
- [144] X. Huang, H.M. Holden, F.M. Raushel, Channeling of substrates and intermediates in enzyme-catalyzed reactions, *Annu. Rev. Biochem.* 70 (2001) 149–180.
- [145] T.O. Yeates, C.A. Kerfeld, S. Heinhorst, G.C. Cannon, J.M. Shively, Protein-based organelles in bacteria: carboxysomes and related microcompartments, *Nat. Rev. Microbiol.* 6 (2008) 681–691.
- [146] T. Geyer, U. Winter, An O(N-2) approximation for hydrodynamic interactions in Brownian dynamics simulations, *J. Chem. Phys.* 130 (2009) 114905.
- [147] S.M. Gopal, S. Mukherjee, Y.-M. Cheng, M. Feig, PRIMO/PRIMONA: a coarse-grained model for proteins and nucleic acids that preserves near-atomic accuracy, *Proteins* 78 (2010) 1266–1281.
- [148] A. Liwo, S. Oldziej, M.R. Pincus, R.J. Wawak, S. Rackovsky, H.A. Scheraga, A united-residue force field for off-lattice protein-structure simulations I. Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data, *J. Comput. Chem.* 18 (1997) 849–873.
- [149] A. Liwo, M.R. Pincus, R.J. Wawak, S. Rackovsky, S. Oldziej, H.A. Scheraga, A united-residue force field for off-lattice protein-structure simulations II. Parameterization of short-range interactions and determination of weights of energy terms by Z-score optimization, *J. Comput. Chem.* 18 (1997) 874–887.
- [150] M. Klann, H. Koeppl, Spatial simulations in systems biology: from molecules to cells, *Int. J. Mol. Sci.* 13 (2012) 7798–7827.
- [151] A. Kudlay, M.S. Cheung, D. Thirumalai, Influence of the shape of crowding particles on the structural transitions in a polymer, *J. Phys. Chem. B.* 116 (2012) 8513–8522.
- [152] D.L. Pincus, D. Thirumalai, Crowding effects on the mechanical stability and unfolding pathways of ubiquitin, *J. Phys. Chem. B.* 113 (2009) 359–368.
- [153] M. Wojciechowski, M. Cieplak, Effects of confinement and crowding on folding of model proteins, *Biosystems* 94 (2008) 248–252.
- [154] W. Wang, W.X. Xu, Y. Levy, E. Trizac, P.G. Wolynes, Confinement effects on the kinetics and thermodynamics of protein dimerization, *PNAS* 106 (2009) 5517–5522.
- [155] J. Rosen, Y.C. Kim, J. Mittal, Modest protein-crowder attractive interactions can counteract enhancement of protein association by intermolecular excluded volume interactions, *J. Phys. Chem. B.* 115 (2011) 2683–2689.
- [156] Y.C. Kim, R.B. Best, J. Mittal, Macromolecular crowding effects on protein–protein binding affinity and specificity, *J. Chem. Phys.* 133 (2010) 205101.
- [157] M.A. Griffin, M. Friedel, J.E. Shea, Effects of frustration, confinement, and surface interactions on the dimerization of an off-lattice beta-barrel protein, *J. Chem. Phys.* 123 (2005) 174707.
- [158] Y. Sliozberg, C.F. Abrams, Effects of confinement on the thermodynamics of a collapsing heteropolymer: an off-lattice Wang–Landau Monte Carlo simulation study, *Macromolecules* 38 (2005) 5321–5329.
- [159] A. Kurut, B.A. Persson, T. Akesson, J. Forsman, M. Lund, Anisotropic interactions in protein mixtures: self assembly and phase behavior in aqueous solution, *J. Phys. Chem. Lett.* 3 (2012) 731–734.
- [160] M. Wojciechowski, P. Szmyczak, M. Cieplak, The influence of hydrodynamic interactions on protein dynamics in confined and crowded spaces—assessment in simple models, *Phys. Biol.* 7 (2010) 046011.
- [161] A.R. Singh, D. Giri, S. Kumar, Force induced unfolding of biopolymers in a cellular environment: a model study, *J. Chem. Phys.* (2009) 131.
- [162] A.R. Singh, D. Giri, S. Kumar, Effects of molecular crowding on stretching of polymers in poor solvent, *Phys. Rev. E* 79 (2009) 051801.
- [163] F.X. Zheng, G.J. Chen, X.R. Zhang, W.C. Wang, A Monte Carlo study of crowding effects on the self-assembly of amphiphilic molecules, *J. Chem. Phys.* 130 (2009) 204701.
- [164] D. Tsao, N.V. Dokholyan, Macromolecular crowding induces polypeptide compaction and decreases folding cooperativity, *Phys. Chem. Chem. Phys.* 12 (2010) 3491–3500.
- [165] A. Samiotakis, M.S. Cheung, Folding dynamics of Trp-cage in the presence of chemical interference and macromolecular crowding I, *J. Chem. Phys.* 135 (2011) 175101.
- [166] D. Homouz, L. Stagg, P. Wittung-Stafshede, M.S. Cheung, Macromolecular crowding modulates folding mechanism of alpha/beta protein apoflavodoxin, *Biophys. J.* 96 (2009) 671–680.
- [167] D. Homouz, H. Sanabria, M.N. Waxham, M.S. Cheung, Modulation of calmodulin plasticity by the effect of macromolecular crowding, *J. Mol. Biol.* 391 (2009) 933–943.
- [168] M.S. Cheung, D. Klimov, D. Thirumalai, Molecular crowding enhances native state stability and refolding rates of globular proteins, *PNAS* 102 (2005) 4753–4758.
- [169] M.S. Cheung, D. Thirumalai, Effects of crowding and confinement on the structure of the transition state ensemble in proteins, *J. Phys. Chem. B.* 111 (2007) 8250–8257.
- [170] A. Christiansen, Q. Wang, A. Samiotakis, M.S. Cheung, P. Wittung-Stafshede, Factors defining effects of macromolecular crowding on protein stability: an *in vitro/in silico* case study using cytochrome c, *Biochemistry* 49 (2010) 6519–6530.
- [171] D. Homouz, M. Perham, A. Samiotakis, M.S. Cheung, P. Wittung-Stafshede, Crowded, cell-like environment induces shape changes in aspherical protein, *PNAS* 105 (2008) 11754–11759.
- [172] A. Dhar, A. Samiotakis, S. Ebbinghaus, L. Nienhaus, D. Homouz, M. Gruebele, et al., Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding, *PNAS* 107 (2010) 17586–17591.
- [173] S.Q. Zhang, M.S. Cheung, Manipulating biopolymer dynamics by anisotropic nanoconfinement, *Nano Lett.* 7 (2007) 3438–3442.
- [174] B.R. Jefferys, L.A. Kelley, M.J.E. Sternberg, Protein folding requires crowd control in a simulated cell, *J. Mol. Biol.* 397 (2010) 1329–1338.
- [175] A. Magno, A. Caflisch, R. Pellarin, Crowding effects on amyloid aggregation kinetics, *J. Phys. Chem. Lett.* 1 (2010) 3027–3032.
- [176] E.H. Yap, T. Head-Gordon, Calculating the bimolecular rate of protein–protein association with interacting crowders, *J. Chem. Theory Comput.* 9 (2013) 2481–2489.
- [177] N.A. Denesyuk, D. Thirumalai, Crowding promotes the switch from hairpin to pseudoknot conformation in human telomerase RNA, *J. Am. Chem. Soc.* 133 (2011) 11858–11861.
- [178] D.L. Pincus, C. Hyeon, D. Thirumalai, Effects of trimethylamine N-oxide (TMAO) and crowding agents on the stability of RNA hairpins, *J. Am. Chem. Soc.* 130 (2008) 7364–7372.
- [179] H.Q. Yu, X.B. Gu, S. Nakano, D. Miyoshi, N. Sugimoto, Beads-on-a-string structure of long telomeric DNAs under molecular crowding conditions, *J. Am. Chem. Soc.* 134 (2012) 20060–20069.
- [180] S. Asakura, F. Oosawa, Interaction between particles suspended in solutions of macromolecules, *J. Polym. Sci.* 33 (1958) 183–192.
- [181] S. Asakura, F. Oosawa, On interaction between 2 bodies immersed in a solution of macromolecules, *J. Chem. Phys.* 22 (1954) 1255–1256.
- [182] Q. Wang, M.S. Cheung, A physics-based approach of coarse-graining the cytoplasm of *Escherichia coli* (CCGYTO), *Biophys. J.* 102 (2012) 2353–2361.
- [183] E. Bianchi, R. Blaak, C.N. Likos, Patchy colloids: state of the art and perspectives, *Phys. Chem. Chem. Phys.* 13 (2011) 6397–6410.
- [184] F. Carlsson, M. Malmsten, P. Linse, Monte Carlo simulations of lysozyme self-association in aqueous solution, *J. Phys. Chem. B* 105 (2001) 12189–12195.

- [185] F. Carlsson, P. Linse, M. Malmsten, Monte Carlo simulations of polyelectrolyte–protein complexation, *J. Phys. Chem. B.* 105 (2001) 9040–9049.
- [186] Y.Q. Li, T.F. Shi, L.J. An, Q.R. Huang, Monte Carlo simulation on complex formation of proteins and polysaccharides, *J. Phys. Chem. B* 116 (2012) 3045–3053.
- [187] E.J. Cho, J.S. Kim, Crowding-induced phase separation of Lennard–Jones particles: implications to nuclear structures in a biological cell, *J. Phys. Chem. B.* 116 (2012) 3874–3879.
- [188] M.C. Abramo, C. Caccamo, D. Costa, G. Pellicane, R. Ruberto, U. Wanderlingh, Effective interactions in lysozyme aqueous solutions: a small-angle neutron scattering and computer simulation study, *J. Chem. Phys.* 136 (2012) 035103.
- [189] M.C. Abramo, C. Caccamo, M. Calvo, V.C. Nibali, D. Costa, R. Giordano, et al., Molecular dynamics and small-angle neutron scattering of lysozyme aqueous solutions, *Philos. Mag.* 91 (2011) 2066–2076.
- [190] M.C. Abramo, C. Caccamo, D. Costa, G. Pellicane, R. Ruberto, Molecular dynamics of an embedded-charge model of lysozyme aqueous solutions, *J. Phys. Chem. B* 114 (2010) 9109–9118.
- [191] E.J.W. Verwey, J.T.G. Overbeek, *Theory of the Stability of Lyophobic Colloids*, Elsevier, Amsterdam, 1948.
- [192] G. Pellicane, Colloidal model of lysozyme aqueous solutions: a computer simulation and theoretical study, *J. Phys. Chem. B* 116 (2012) 2114–2120.
- [193] T.W. Rosch, J.R. Errington, Investigation of the phase behavior of an embedded charge protein model through molecular simulation, *J. Phys. Chem. B.* 111 (2007) 12591–12598.
- [194] J. Sun, H. Weinstein, Toward realistic modeling of dynamic processes in cell signaling: quantification of macromolecular crowding effects, *J. Chem. Phys.* 127 (2007) 155105.
- [195] V.K. Shen, J.K. Cheung, J.R. Errington, T.M. Truskett, Insights into crowding effects on protein stability from a coarse-grained model, *J. Biomech. Eng.–T. ASME.* 131 (2009) 071002.
- [196] J.K. Cheung, V.K. Shen, J.R. Errington, T.M. Truskett, Coarse-grained strategy for modeling protein stability in concentrated solutions III: Directional protein interactions, *Biophys. J.* 92 (2007) 4316–4324.
- [197] V.K. Shen, J.K. Cheung, J.R. Errington, T.M. Truskett, Coarse-grained strategy for modeling protein stability in concentrated solutions II: Phase behavior, *Biophys. J.* 90 (2006) 1949–1960.
- [198] J.K. Cheung, T.M. Truskett, Coarse-grained strategy for modeling protein stability in concentrated solutions, *Biophys. J.* 89 (2005) 2372–2384.
- [199] M. Heinen, F. Zanini, F. Roosen-Runge, D. Fedunova, F.J. Zhang, M. Hennig, et al., Viscosity and diffusion: crowding and salt effects in protein solutions, *Soft Matter* 8 (2012) 1404–1419.
- [200] H.M. Senn, W. Thiel, QM/MM methods for biomolecular systems, *Angew. Chem. Int. Ed.* 48 (2009) 1198–1229.
- [201] M. Stein, R.R. Gabdoulline, R.C. Wade, Bridging from molecular simulation to biochemical networks, *Curr. Opin. Struct. Biol.* 17 (2007) 166–172.
- [202] H. Kitano, Computational systems biology, *Nature* 420 (2002) 206–210.
- [203] J. Pahle, Biochemical simulations: stochastic, approximate stochastic and hybrid approaches, *Brief. Bioinf.* 10 (2009) 53–64.
- [204] M.J. Morelli, R.J. Allen, P.R. ten Wolde, Effects of macromolecular crowding on genetic networks, *Biophys. J.* 101 (2011) 2882–2891.
- [205] Z. Frazier, F. Alber, A computational approach to increase time scales in Brownian dynamics-based reaction-diffusion modeling, *J. Comput. Biol.* 19 (2012) 606–618.
- [206] K. Aoki, K. Takahashi, K. Kaizu, M. Matsuda, A quantitative model of ERK MAP kinase phosphorylation in crowded media, *Sci. Rep.–UK* 3 (2013) 1541.
- [207] D. Ridgway, G. Broderick, A. Lopez-Campistrous, M. Ru'aini, P. Winter, M. Hamilton, et al., Coarse-grained molecular simulation of diffusion and reaction kinetics in a crowded virtual cytoplasm, *Biophys. J.* 94 (2008) 3748–3759.
- [208] E. Roberts, J.E. Stone, Z. Luthey-Schulten, Lattice Microbes:, High-performance stochastic simulation method for the reaction–diffusion master equation, *J. Comput. Chem.* 34 (2013) 245–255.
- [209] E. Roberts, A. Magis, J.O. Ortiz, W. Baumeister, Z. Luthey-Schulten, Noise contributions in an inducible genetic switch: a whole-cell simulation study, *PLoS Comput. Biol.* 7 (2011) e1002010.
- [210] K. Takahashi, S.N.V. Arjunan, M. Tomita, Space in systems biology of signaling pathways—towards intracellular molecular crowding in silico, *FEBS Lett.* 579 (2005) 1783–1788.
- [211] M.T. Klann, A. Lapin, M. Reuss, Agent-based simulation of reactions in the crowded and structured intracellular environment: influence of mobility and location of the reactants, *BMC Syst. Biol.* 5 (2011) 71.
- [212] A. Leier, T.T. Marquez-Lago, K. Burrage, Simulating diffusion in crowded environments with multifractional Brownian motion, *FEBS J.* 279 (2012) 524.
- [213] K. Burrage, P.M. Burrage, A. Leier, T. Marquez-Lago, D.V. Nicolau, Stochastic simulation for spatial modelling of dynamic processes in a living cell, in: *Design and Analysis of Biomolecular Circuits:Engineering Approaches to Systems and Synthetic Biology*, Springer, 2011, pp. 43–62.
- [214] K. Takahashi, S. Tanase-Nicola, P.R. ten Wolde, Spatio-temporal correlations can drastically change the response of a MAPK pathway, *PNAS* 107 (2010) 2473–2478.
- [215] A. Lapin, M. Klann, M. Reuss, Multi-scale spatio-temporal modeling: lifelines of microorganisms in bioreactors and tracking molecules in cells, *Adv. Biochem. Eng. Biotechnol.* 121 (2010) 23–43.
- [216] D.T. Gillespie, General method for numerically simulating stochastic time evolution of coupled chemical-reactions, *J. Comput. Phys.* 22 (1976) 403–434.
- [217] B. Lee, P.R. LeDuc, R. Schwartz, Three-dimensional stochastic off-lattice model of binding chemistry in crowded environments, *PLoS One* 7 (2012) e30131.
- [218] B. Lee, P.R. LeDuc, R. Schwartz, Unified regression model of binding equilibria in crowded environments, *Sci. Rep.–UK* 1 (2011) 97.
- [219] B. Lee, P.R. Leduc, R. Schwartz, Parameter effects on binding chemistry in crowded media using a two-dimensional stochastic off-lattice model, *Phys. Rev. E.* 80 (2009) 041918.
- [220] B. Lee, P.R. LeDuc, R. Schwartz, Stochastic off-lattice modeling of molecular self-assembly in crowded environments by Green's function reaction dynamics, *Phys. Rev. E.* 78 (2008) 031911.
- [221] J.S. van Zon, P.R. ten Wolde, Green's-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space, *J. Chem. Phys.* 123 (2005) 234910.
- [222] E. Yus, T. Maier, K. Michalodimitrakis, V. van Noort, T. Yamada, W.H. Chen, et al., Impact of genome reduction on bacterial metabolism and its regulation, *Science* 326 (2009) 1263–1268.
- [223] M. Guell, V. van Noort, E. Yus, W.H. Chen, J. Leigh-Bell, K. Michalodimitrakis, et al., Transcriptome complexity in a genome-reduced bacterium, *Science* 326 (2009) 1268–1271.
- [224] P. Sherwood, B.R. Brooks, M.S.P. Sansom, Multiscale methods for macromolecular simulations, *Curr. Opin. Struct. Biol.* 18 (2008) 630–640.
- [225] M. Praprotnik, L. Delle Site, K. Kremer, Multiscale simulation of soft matter: from scale bridging to adaptive resolution, *Annu. Rev. Phys. Chem.* 59 (2008) 545–571.
- [226] J.W. Chu, G.S. Ayton, S. Izvekov, G.A. Voth, Emerging methods for multiscale simulation of biomolecular systems, *Mol. Phys.* 105 (2007) 167–175.
- [227] A.P. Heath, L. Karvaki, C. Clementi, From coarse-grain to all-atom: toward multiscale analysis of protein landscapes, *Proteins* 68 (2007) 646–661.
- [228] S. Izvekov, G.A. Voth, A multiscale coarse-graining method for biomolecular systems, *J. Phys. Chem. B* 109 (2005) 2469–2473.
- [229] S. Tanizaki, J. Clifford, B.D. Connelly, M. Feig, Conformational sampling of peptides in cellular environments, *Biophys. J.* 94 (2008) 747–759.
- [230] M. Feig, S. Tanizaki, M. Sayadi, Implicit solvent simulations of biomolecules in cellular environments, in: *Annual Reviews in Computational Chemistry*, Elsevier, Oxford, 2008, pp. 4.
- [231] C. Echeverria, R. Kapral, Molecular crowding and protein enzymatic dynamics, *Phys. Chem. Chem. Phys.* 14 (2012) 6755–6763.
- [232] C. Dellago, P.G. Bolhuis, F.S. Csajka, D. Chandler, Transition path sampling and the calculation of rate constants, *J. Chem. Phys.* 108 (1998) 1964–1977.
- [233] M.J. Field, Simulating enzyme reactions: challenges and perspectives, *J. Comput. Chem.* 23 (2002) 48–58.
- [234] G. Louit, M. Hanedanian, F. Taran, H. Coffigny, J.-P. Renault, S. Pin, Determination of hydroxyl rate constants by a high-throughput fluorimetric assay: towards a unified reactivity scale for antioxidants, *Analyst* 134 (2009) 250–255.
- [235] C.T. Campbell, G. Kim, SPR microscopy and its applications to high-throughput analyses of biomolecular binding events and their kinetics, *Biomaterials* 28 (2007) 2380–2392.
- [236] S. Hihara, C.G. Pack, K. Kaizu, T. Tani, T. Hanafusa, T. Nozaki, et al., Local nucleosome dynamics facilitate chromatin accessibility in living mammalian cells, *Cell Rep.* 2 (2012) 1645–1656.
- [237] H.N. Chapman, X-ray imaging beyond the limits, *Nat. Mater.* 8 (2009) 299–301.
- [238] B. Huang, M. Bates, X. Zhuang, Super-resolution fluorescence microscopy, *Annu. Rev. Biochem.* 78 (2009) 993–1016.