Bridging from molecular simulation to biochemical networks

Matthias Stein[1], Razif R Gabdoulline[1,2] and Rebecca C Wade[1]

How can we make the connection between the three-dimensional structures of individual proteins and understanding how complex biological systems involving many proteins work? The modelling and simulation of protein structures can help to answer this question for systems ranging from multimacromolecular complexes to organelles and cells. On one hand, multiscale modelling and simulation techniques are advancing to permit the spatial and temporal properties of large systems to be simulated using atomic-detail structures. On the other hand, the estimation of kinetic parameters for the mathematical modelling of biochemical pathways using protein structure information provides a basis for iterative manipulation of biochemical pathways guided by protein structure. Recent advances include the structural modelling of protein complexes on the genomic level, novel coarse-graining strategies to increase the size of the system and the time span that can be simulated, and comparative molecular field analyses to estimate enzyme kinetic parameters.

Addresses
[1] Molecular and Cellular Modeling Group, EML Research, Schloss-Wolfsbrunnenweg 33, D-69118 Heidelberg, Germany
[2] BIOMS (Center for Modeling and Simulation in the Biosciences), University of Heidelberg, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany

Corresponding author: Wade, Rebecca C (rebecca.wade@eml-r.villa-bosch.de)

Current Opinion in Structural Biology 2007, **17**:166–172

This review comes from a themed issue on
Theory and simulation
Edited by Richard Lavery and Kim A Sharp

Available online 28th March 2007

0959-440X/$ – see front matter

DOI 10.1016/j.sbi.2007.03.014

Introduction

The structures of proteins are being solved in ever-increasing numbers, particularly as a result of structural genomics projects. The number of protein structures that can be modelled is rising concomitantly [1]. This structural information provides a basis for understanding protein function and for the design of modified proteins and ligands, including drugs [2]. Although it is important to study proteins as individual entities, a full understanding of their function requires treatment of their biological context, which may be considered as, for example, a multiprotein complex, a crowded cell compartment or a metabolic network. This task is very challenging because it requires bridging many orders of magnitude in spatial and temporal dimensions.

We consider two main strategies to bridge from atomic-detail structures of proteins to the complex biological systems they participate in (see Figure 1). The first approach is to construct spatial models by building up from models of multiprotein complexes to simulations of the dynamics of macromolecules in cellular compartments using multiscale methods. The models should permit zooming in and out in spatial and temporal resolution. The state of the art of this approach is summarized in the next section. The second approach is to feed properties, such as kinetic parameters, computed from protein structures into macroscopic biochemical network models. Simulation of such biochemical networks can permit critical steps in the networks to be identified. If these are dependent on parameters that can be estimated from protein structure, then they can be targeted for alteration by structure-based ligand design or protein engineering. Thus, an iterative cycle between protein-structure-based simulation and design, and mathematical modelling of biochemical networks can be established. This second approach, which provides a practical means of bridging from molecular modelling and simulation to biochemical networks, is the main focus of this review.

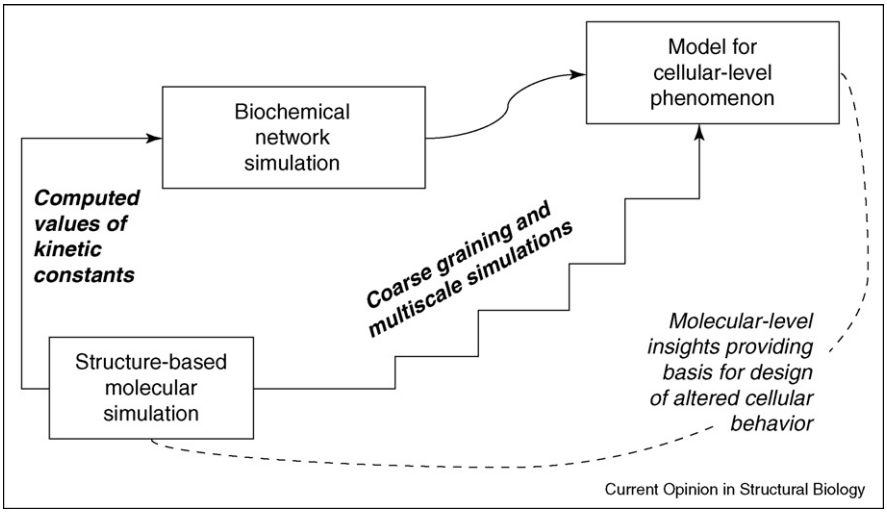
Towards dynamic structural models of the cell

Advances in imaging and fluorescence labelling techniques are providing insights into the location and dynamics of specific macromolecules in cells. For a recent overview of biological, biochemical and physical methods to characterize single cells, proteins and protein complexes, see [3]. The growing amount of spatial and temporal data needs to be put in a common framework [4]. This raises challenges in modelling the interactions of macromolecules at the cellular level. We summarize how this is being addressed from a bottom-up perspective.

Modelling multimacromolecular complexes

The modelling of multimacromolecular complexes is a first step in going from individual protein structures to a spatial model of a cell and all its molecular components. The methods and algorithms to elucidate the structures of protein–protein macromolecular assemblies have recently been reviewed [5]. Recent studies have shown that large complexes can be modelled using a combination of diverse types of experimental data and modelling techniques. For example, raw electron density maps from cryo-electron microscopy (cryo-EM) can be used to assist the modelling of multiprotein complexes (for a review, see [6]). The

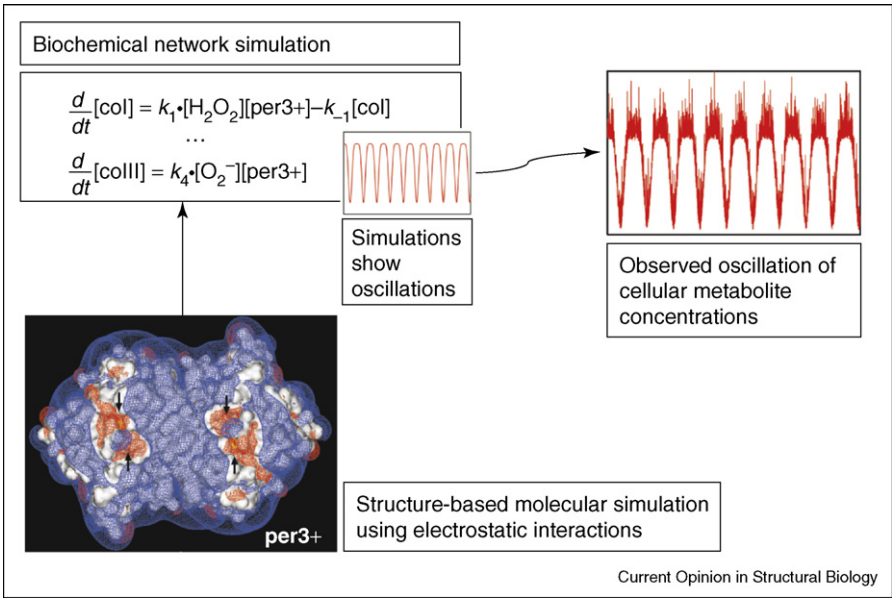
Figure 1



Schematic figure showing two different approaches to bridging from structure-based molecular simulation to complex cellular systems.

modelling of protein assemblies based on the structures of homologous protein–protein complexes has been attempted on the genomic scale [7,8]. When template structures of complexes are not available, protein–protein docking techniques have been applied to generate structures of binary and multiprotein complexes [9]. Larger systems can be modelled at a lower level of resolution. The nuclear pore complex was modelled by Devos *et al.* [10] using a modular fold recognition and assignment approach, followed by iterative sequence-structure alignment and model building. Geyer and Helms [11,12•] made spatial and kinetic reconstructions of chromatophore vesicles from *Rhodobacter sphaeroides* purple bacteria. These spherical vesicles of ~45 nm diameter contain the photosynthetic

Figure 2



Schematic figure illustrating the concerted use of protein-structure-based simulations and mathematical modelling to study a cellular phenomenon. This figure follows the layout of the scheme in Figure 1. The crystal structure of myeloperoxidase was used to simulate the diffusional association of the superoxide anion, and to compute the diffusional association rate constant and its dependence on pH. This is influenced by the electrostatic potential of the protein, shown by isosurfaces at −0.3 (red) and +0.3 (blue) kcal/mol/e. The kinetic parameter computed for this process (k_4) was used in a 19-differential equation model of the peroxidase-oxidase reaction, involving two cellular compartments and two enzymes, that reproduced experimentally observed oscillations in the levels of metabolites in activated neutrophils [41].

machinery and have been extensively characterized experimentally. The models provide a basis for stochastic dynamic simulations of the complete vesicle at the molecular level and assist in the resolution of experimental ambiguities.

Simulation of the dynamics of macromolecular assemblies

The simulation of the dynamics of macromolecular assemblies is the next step in building a dynamic model of the cell. The size and complexity of systems that can be simulated by atomic-detail molecular dynamics (MD) simulation techniques using molecular mechanics (MM) models is constantly increasing [13]. Examples of large systems simulated recently include the complete satellite tobacco virus [14], tRNA decoding in the ribosome [15] and cellular machines [16]. The rare occurrence of barrier-crossing events in biochemical processes on the MD timescale (of the order of tens of nanoseconds) means that alterations of the potential energy surfaces or special configurational sampling schemes are frequently necessary. An alternative is to reduce the level of detail in the simulations, that is, the number of particles simulated can be reduced by coarse graining (CG), enabling longer timescales and larger spatial scales to be accessed. CG methods for proteins have been reviewed recently [17]. A recent example of mesoscale simulations employing CG is the investigation of the role of histone tails in chromatin folding [18•,19].

Different parts of a system might require different treatments or different levels of detail. The rigid parts of a molecule can be identified based on properties such as the presence of a hydrogen-bonding network and then the links between them modelled as fully flexible [20]. By extending the timescale of simulations of proteins using CG, microseconds can be simulated with models comprising 1–2 beads per residue [21•,22], whereas milliseconds can be accessed when protein domains are treated as rigid bodies [23].

Simulation of the dynamics of transient interactions in large molecular systems

A range of techniques are required for the simulation of the dynamics of transient interactions in large molecular systems. Atomically detailed macromolecular structures can be used directly. This was recently demonstrated by McGuffee and Elcock [24•], who performed Brownian dynamics (BD) simulations of the diffusional motion of proteins in 1000-macromolecule systems of concentrated protein solutions. Each protein was treated as a rigid atomic-detail entity. The simulations reproduced the experimental diffusional properties of protein solutions and the same approach is being used to simulate a piece of cytoplasm crowded with a range of different macromolecular entities (AH Elcock, personal communication). Another way to use atomically detailed protein structures is in a multiscale modelling framework, in which parameters from more detailed models are used in less detailed models. For example, Pricl *et al.* [25,26•] used atomistic MD simulations of lysozyme–silicon surface interactions to compute molecular energetic and structural parameters for a macroscale continuum dynamic model of the diffusion of lysozyme through silicon membrane nanochannels.

Simulation of biochemical networks

The biochemical processes in biological cells are complex and interwoven (as can readily be seen from a glance at the 'Biochemical Pathways' wall chart; www.expasy.ch/tools/pathways). Pathway cross-talk by shared metabolites or enzymes, regulation, and positive and negative feedback all contribute to the complexity. Mathematical models are essential for understanding biochemical networks and predicting their behaviour under perturbation. Mathematical models permit the time dependence of metabolite concentrations to be simulated. The spatial variation of metabolite concentrations can also be simulated (see Box 1).

Although it is possible to attempt whole-cell simulations with all relevant metabolites included (e.g. the E-Cell project [27] or Project Cybercell [28], to name only two), it is more common to model individual pathways or

Box 1 Mathematical modelling and simulation of biochemical networks.

Models of biochemical networks are typically constructed as sets of differential equations describing the time dependence of compound concentrations. **Ordinary differential equations (ODEs)**, dependent on only one variable (e.g. time), can be used to describe the change of states (e.g. compound concentrations). The equations can describe changes due to reaction and diffusion or other transport processes. **Partial differential equations (PDEs)** depend on more variables and enable the investigation of spatial and temporal changes of molecular concentrations. The differential equations contain **kinetic parameters**, which need to be determined experimentally or estimated computationally.

When molecular concentrations fall below a threshold such that they can no longer be treated uniformly and deterministically, then the discrete nature of the molecules and their **stochastic** behaviour has to be taken into account in simulations with random fluctuations. This can be done using the Gillespie algorithm to compute the time evolution of the probability distribution of a stochastic system of reactions.

The dependence of the results on the choice of parameter values can be investigated by a sensitivity analysis. Metabolic control analysis is one tool commonly used to investigate the relationship between a steady-state system and individual reactions on a quantitative and qualitative level. The sensitivity of the steady-state solution to small parameter perturbations is analyzed and insight gained into the regulatory and feedback properties of the network.

A variety of software packages are available for modelling and simulating biochemical networks [51]. The models can be described and exchanged using systems biology mark-up language (SBML) or cell mark-up language (CellML) formats. An overview of concepts and applications of systems biology can be found in [52].

groups of pathways. A modular approach can then be taken to investigate cross-talk between pathways and reconstitute cell-like simulations from independent modules [29]. Pathways that have been extensively modelled and simulated recently include the mitogen-activated protein kinase (MAPK) pathway [30] and glycolysis [31,32].

Spatio-temporal simulations are reviewed in [33,34]. Recent applications include the simulation of ectopic neurotransmission with a model synapse using a Monte Carlo (MC) method with molecules represented as diffusing and reacting particles in a synaptic topology derived from electron tomography [35]. Fleire *et al.* [36] used BD simulations to model B-cell antigen collection through a spreading and contraction response, and Means *et al.* [37] applied a finite element reaction-diffusion model of mast cell calcium dynamics that incorporates the complex geometry of the endoplasmic reticulum.

Parameterization of models for the simulation of biochemical networks

When biochemical networks are simulated, kinetic parameters are required for the differential equations in the model (see Box 1). If a differential equation describes a reaction catalysed by an enzyme, the parameters will correspond to rate constants for an elementary reaction step or the overall reaction (e.g. K_m, V_{max} and k_{cat}). Ideally, all the parameters required for a given mathematical model will have been determined experimentally under conditions relevant to the model. Indeed, with advances in high-throughput experimentation, more quantitative and consistent enzyme kinetic data are becoming available [3]. This makes it feasible to obtain such parameters for a complete biochemical pathway in a systems biology project. However, the usual situation is that experimental measurements of the kinetic parameters for a pathway are distributed in different scientific publications and relate to different experimental conditions. Some of this kinetic information is stored in databases such as BRENDA [38] and SABIO-RK [39]. Curation of these databases and defined standards in reporting enzyme kinetic data (STRENDA, Standards for Reporting Enzymology Data) [40] will facilitate the task of locating and selecting relevant kinetic measurements. Often, experimental data are lacking and parameters must then be estimated and the effect of their variation assessed by sensitivity analysis of the model. Parameter estimation should make use of values available for homologous enzymes or for the same enzyme under different conditions (pH, temperature, buffers, etc.).

This is where protein structure information becomes useful. It has recently been shown that, when protein structure information is available for the enzymes involved, kinetic parameters can be estimated from the protein structures and used in macroscopic mathematical models [41] (see Figure 2). This provides a means for bridging between atomic-detail protein structures and cellular phenomena, such as the oscillation of the levels of metabolites in a cell.

Enzyme kinetic parameter estimation using protein structures

Enzyme specificity, ligand-binding affinity and catalytic turnover rate are all encoded in the three-dimensional macromolecular protein structure. The kinetic parameters are dependent on the free energy profile of the reaction. This free energy profile, and thus the rate-limiting step, varies according to the nature of the substrate and the environmental conditions (pH, temperature, etc.). The barrier to the reaction might be due to a bond breaking or making event, a conformational change or a diffusional process. Therefore, the type of computational method necessary to compute a kinetic parameter is dependent on the physicochemical properties of the rate-determining event.

When bond breaking or making is important, transition state theory and a quantum-mechanical (QM) treatment can be used to calculate enzyme kinetic parameters [42,43]. QM, QM/MM and empirical valence bond (EVB) methods are widely used to study enzyme catalysis [44]. Classical MD simulation methods are appropriate to compute free energy differences for enzyme reactions involving conformational changes [13]. Both QM and MD simulations are time and resource intensive, and therefore currently impractical in a biochemical pathway modelling project.

When the reaction is fast and subject to diffusion control, association rate constants can be calculated using BD simulations. These are much less computationally demanding than QM and MM calculations, as the solvent is treated as a continuum and the protein is typically kept rigid. Important dynamic motions in the protein can be accounted for using a CG model. BD simulations can be used to study enzyme–substrate or enzyme–inhibitor association, and the effects of point mutations and changes in pH, viscosity and ionic strength on bimolecular diffusional association rate constants. Concerted use of protein-structure-based BD simulations and mathematical network modelling has been demonstrated in the modelling of oscillatory behaviour in activated neutrophils [41] (see Figure 2).

Practically, for biochemical network modelling, methods are needed that can exploit protein structure information to produce approximate estimates of kinetic parameters with limited investment of human and computational time. One approach is to use similarity indices to compare the molecular interaction fields (MIFs), for example the molecular electrostatic potentials, of structurally related

proteins [45]. If the kinetic parameters are known for a few related proteins, then their MIFs can be compared with that of the protein for which a kinetic parameter is required and, based on similarity/dissimilarity, a value can be assigned to the missing kinetic parameter. This approach relies on the assumption that the MIF, or region of the MIF compared, is a major factor in determining the relative values of the kinetic parameters. Relative values of enzyme parameters can be correlated across a set of mutants or a set of orthologous enzymes from different species by comparing electrostatic potentials in the vicinity of the active site (e.g. for triose phosphate isomerase from 12 species; RR Gabdoulline, M Stein, RC Wade, unpublished). The approach is feasible for large-scale analysis and has been applied to ten enzymes of a biochemical metabolic pathway in eleven homologous species [46,47]. For estimation of parameter values, at least two experimentally determined rates for related systems are required. The accuracy of the estimation depends strongly on the quantity, quality and consistency of the experimental data. Even if no experimental values of kinetic parameters are available, the comparison of MIFs can assist in the identification of outliers or the classification of families of enzymes as regards enzymatic properties (RR Gabdoulline, M Stein, RC Wade, unpublished).

MIFs are a basic determinant of protein function [48] and therefore a generally applicable property for comparative functional analysis. The protein electrostatic potential in the vicinity of the active site has been shown to be a determinant of diffusional association rates [49]. Furthermore, the catalytic power of enzymes through stabilization of the transition state is classically thought to be due to the electrostatic properties [50]. No docking of ligands is required in a MIF analysis, but this also means that no direct ligand-specific information is obtained. Other comparative approaches are possible in a systems biology context to correlate enzyme kinetics with structural parameters such as ligand–active site distances, number of hydrogen bonds between substrate and active site, and so on. Further approaches are needed to develop the reliability and range of applicability of computationally viable procedures for kinetic parameter estimation in biochemical pathway simulation.

Concluding remarks

Only a small number of metabolic pathways have been thoroughly characterized experimentally and only for relatively simple organisms, such as yeast and *Trypanosoma brucei*. The identification of critical steps in pathways by a combination of quantitative biological assays and mathematical network modelling, augmented by detailed insights into processes at the protein level, opens up prospects for protein engineering, for example, to enhance the production of desired metabolites in biotechnology, and selective drug design, for example, against parasites. For *T. brucei*, metabolic control analysis has been used to suggest protein targets for the selective inhibition of glycolysis [32].

Signal transduction pathways have recently gained attention because of their intrinsic importance in controlling protein expression and activation in organisms from yeast to human. The increase in the availability of quantitative biological assays of signalling pathways has enabled the mathematical modelling of certain modules of these signalling pathways. The data, however, are available only for a few signalling pathways (modules) and currently cannot be combined to give a more complete picture of quantitative processes at the cellular level.

Databases [38,39], standards [40] and software tools are required to facilitate the concerted use of different modelling and simulation methods for studying biochemical networks. In addition, a common definition of ontologies and data-exchange formats is necessary. SYCAMORE (Systems Biology Computational Analysis and Modelling Research Environment [47]) is being developed to provide a software environment to guide the user through the steps of building a network model, locating information in databases, estimating kinetic parameters and choosing an appropriate simulation method. It integrates protein-structure-based modelling for kinetic parameter estimation and mathematical simulations of biochemical networks.

The modular building principle of biological modelling is apparent: different temporal and spatial scales need to be addressed. One has to move between the different aspects and capabilities of simulation techniques, and also between different levels of model representation. Results from more accurate simulations of smaller systems can be used to generate parameters for more approximate simulations of larger systems (CG; bottom-up principle). Likewise, once critical enzymatic steps have been identified in coarser or more abstract systems simulations, more detailed and accurate molecular simulations may be initiated (fine-graining) to elucidate the cause of the behaviour of the protein of interest (top-down principle). These two approaches complement to lead to an iterative modelling cycle.

Biological processes are complex. The investigation of these processes requires different tools and methodologies to study processes at atomistic, molecular, ensemble or abstract mathematical levels. Currently, the various approaches are fragmented and the mutual benefit of each level of representation has not been fully exploited. It is necessary to combine simulation methods at different levels of representation to bridge from molecular simulations to biochemical networks.

Acknowledgements

Financial support from the Klaus Tschira Foundation, the German Federal Ministry of Education and Research (BMBF) Systems Biology

Initiative 'Hepatosys' (grant number 0313076), the German Research Foundation (grant number WA-1381/1-1) and the Heidelberg Center for Modelling and Simulation in the Biosciences (BIOMS) is gratefully acknowledged. We thank Peter J Winn for critical reading of the manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Baker D, Sali A: **Protein structure prediction and structural genomics**. *Science* 2001, **294**:93-96.

2. Harrison SC: **Whither structural biology?** *Nat Struct Mol Biol* 2004, **11**:12-15.

3. Albeck JG, MacBeath G, White FM, Sorger PK, Lauffenburger DA, Gaudet S: **Collecting and organizing systematic sets of protein data**. *Nat Rev Mol Cell Biol* 2006, **7**:803-812.

4. Bork P, Serrano L: **Towards cellular systems in 4D**. *Cell* 2005, **121**:507-509.

5. Aloy P, Russell RB: **Structural systems biology: modelling protein interactions**. *Nat Rev Mol Cell Biol* 2006, **7**:188-197.

6. Topf M, Sali A: **Combining electron microscopy and comparative protein structure modeling**. *Curr Opin Struct Biol* 2005, **15**:578-585.

7. Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dumpelfeld B *et al.*: **Proteome survey reveals modularity of the yeast cell machinery**. *Nature* 2006, **440**:631-636.

8. Davis FP, Braberg H, Shen MY, Pieper U, Sali A, Madhusudhan MS: **Protein complex compositions predicted by structural similarity**. *Nucleic Acids Res* 2006, **34**:2943-2952.

9. Inbar Y, Benyamini H, Nussinov R, Wolfson HJ: **Prediction of multimolecular assemblies by multiple docking**. *J Mol Biol* 2005, **349**:435-447.

10. Devos D, Dokudovskaya S, Williams R, Alber F, Eswar N, Chait BT, Rout MP, Sali A: **Simple fold composition and modular architecture of the nuclear pore complex**. *Proc Natl Acad Sci USA* 2006, **103**:2172-2177.

11. Geyer T, Helms V: **Spatial model of the chromatophore vesicles of *Rhodobacter sphaeroides* and the position of the cytochrome bc1 complex**. *Biophys J* 2006, **91**:921-926.

12. • Geyer T, Helms V: **Reconstruction of a kinetic model of the chromatophore vesicles from *Rhodobacter sphaeroides***. *Biophys J* 2006, **91**:927-937.

Two papers [11,12•] describe how a kinetic model of chromatophore vesicles is coupled to the reconstruction of a spatial model. The model gives a steady-state picture of the photosynthetic apparatus and can be used as a basis for future simulations.

13. Karplus M, Kuriyan J: **Molecular dynamics and protein function**. *Proc Natl Acad Sci USA* 2005, **102**:6679-6685.

14. Freddolino PL, Arkhipov AS, Larson SB, McPherson A, Schulten K: **Molecular dynamics simulations of the complete satellite tobacco mosaic virus**. *Structure* 2006, **14**:437-449.

15. Sanbonmatsu KY, Joseph S, Tung CS: **Simulating movement of tRNA into the ribosome during decoding**. *Proc Natl Acad Sci USA* 2005, **102**:15854-15859.

16. Gao M, Sotomayor M, Villa E, Lee EH, Schulten K: **Molecular mechanisms of cellular mechanics**. *Phys Chem Chem Phys* 2006, **8**:3692-3706.

17. Tozzini V: **Coarse-grained models for proteins**. *Curr Opin Struct Biol* 2005, **15**:144-150.

18. • Arya G, Schlick T: **Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model**. *Proc Natl Acad Sci USA* 2006, **103**:16236-16241.

In [18•,19], a mesoscopic oligonucleosome model was constructed with the flexible histone tails modelled using a two-step CG procedure. A five-residue per protein bead model was parameterized to reproduce the dynamic and configurational properties of a one-residue per protein bead chain model simulated by BD.

19. Arya G, Zhang Q, Schlick T: **Flexible histone tails in a new mesoscopic oligonucleosome model**. *Biophys J* 2006, **91**:133-150.

20. Ahmed A, Gohlke H: **Multiscale modeling of macromolecular conformational changes combining concepts from rigidity and elastic network theory**. *Proteins* 2006, **63**:1038-1051.

21. • Chang CE, Shen T, Trylska J, Tozzini V, McCammon JA: **Gated binding of ligands to HIV-1 protease: Brownian dynamics simulations in a coarse-grained model**. *Biophys J* 2006, **90**:3880-3885.

The authors describe BD simulations of drug binding to HIV-1 protease in a CG model. The developed methodology is an efficient way to calculate the association rate constants gated by the microsecond dynamics of a protease.

22. Shih AY, Arkhipov A, Freddolino PL, Schulten K: **Coarse grained protein-lipid model with application to lipoprotein particles**. *J Phys Chem B* 2006, **110**:3674-3684.

23. Chu JW, Voth GA: **Coarse-grained modeling of the actin filament derived from atomistic-scale simulations**. *Biophys J* 2006, **90**:1572-1582.

24. • McGuffee SR, Elcock AH: **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* 2006, **128**:12098-12110.

BD simulations of thousands of rigid protein molecules treated in atomic detail are reported. The interaction and diffusional properties of proteins in concentrated solution are derived from microsecond-long simulations as a function of both solvent conditions and amino acid substitutions.

25. Pricl S, Ferrone M, Fermeglia M, Amato F, Cosentino C, Cheng MM-C, Walczak R, Ferrari M: **Multiscale modeling of protein transport in silicon membrane nanochannels. Part 1. Derivation of molecular parameters from computer simulations**. *Biomed Microdevices* 2006, **8**:277-290.

26. • Pricl S, Ferrone M, Fermeglia M, Amato F, Cosentino C, Cheng MM-C, Walczak R, Ferrari M: **Multiscale modeling of protein transport in silicon membrane nanochannels. Part 2. From molecular parameters to a predictive continuum diffusion model**. *Biomed Microdevices* 2006, **8**:291-298.

Two papers [25,26•] describe MD simulations that were used to derive parameters for protein–silicon membrane interactions. These were then used to model protein diffusion in nanochannels and reproduce non-Fickian diffusion on a timescale of days.

27. Tomita M, Hashimoto K, Takahashi K, Matsuzaki Y, Matsishima R, Saito K, Yugi K, Miyoshi F, Nakano H, Tanida S *et al.*: **The E-cell project: towards integrative simulation of cellular processes**. *New Generation Computing* 2000, **18**:1-12.

28. Broderick G, Ru'aini M, Chan E, Ellison MJ: **A life-like virtual cell membrane using discrete automata**. *In Silico Biol* 2004, **5**:0016.

29. Snoep JL, Bruggeman F, Olivier BG, Westerhoff HV: **Towards building the silicon cell: a modular approach**. *Biosystems* 2006, **83**:207-216.

30. Orton RJ, Sturm OE, Yvshermirsky V, Calder M, Gilbert DR, Kolch W: **Computational modelling of receptor-tyrosine-kinase-activated MAPK pathway**. *Biochem J* 2005, **392**:249-261.

31. Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV *et al.*: **Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry**. *Eur J Biochem* 2000, **267**:5313-5329.

32. Albert MA, Haanstra JR, Hannaert V, Van Roy J, Opperdoes FR, Bakker BM, Michels PAM: **Experimental and *in silico* analyses of glycolytic flux control in bloodstream form *Trypanosoma brucei***. *J Biol Chem* 2005, **280**:28306-28315.

33. Lemerle C, Di Ventura B, Serrano L: **Space as the final frontier in stochastic simulations of biological systems**. *FEBS Lett* 2005, **579**:1789-1794.

34. Takahashi K, Arjunan SNV, Tomita M: **Space in systems biology of signaling pathways - towards intracellular molecular crowding *in silico***. *FEBS Lett* 2005, **579**:1783-1788.

35. Coggan JS, Bartol TM, Esquenazi E, Stiles JR, Lamont S, Martone ME, Berg DK, Ellisman MH, Sejnowski TJ: **Evidence for ectopic neurotransmission at a neuronal synapse**. *Science* 2005, **309**:446-451.

36. Fleire SJ, Goldman JP, Carrasco YR, Weber M, Bray D, Batista FD: **B-cell ligand discrimination through a spreading and contraction response**. *Science* 2006, **312**:738-741.

37. Means S, Smith AJ, Shepherd J, Shadid J, Fowler J, Wojcikiewicz RJH, Mazel T, Smith GD, Wilson BS: **Reaction diffusion modeling of calcium dynamics with realistic ER geometry**. *Biophys J* 2006, **91**:537-557.

38. Schomburg I, Chang A, Ebeling C, Gremse M, Heldt C, Huhn G, Schomburg D: **BRENDA, the enzyme database: updates and major new developments**. *Nucleic Acids Res* 2004, **32**:D431-D433.

39. Wittig U, Golebiewski M, Kania R, Krebs O, Mir S, Weidemann A, Anstein S, Saric J, Rojas I: **SABIO-RK: integration and curation of reaction kinetics data**. *Lecture Notes in Computer Science* 2006, **4075**:94-103.

40. Apweiler R, Cornish-Bowden A, Hofmeyr J-HS, Kettner C, Leyh TS, Schomburg D, Tipton K: **The importance of uniformity in reporting protein-function data**. *Trends Biochem Sci* 2005, **30**:11-12.

41. Gabdoulline RR, Kummer U, Olsen LF, Wade RC: **Concerted simulations reveal how peroxidase compound iii formation results in cellular oscillations**. *Biophys J* 2003, **85**:1421-1428.

42. Gao J, Truhlar DG: **Quantum mechanical methods for enzyme kinetics**. *Annu Rev Phys Chem* 2002, **53**:467-505.

43. Gao J, Ma S, Major DT, Nam K, Pu J, Truhlar DG: **Mechanisms and free energies of enzymatic reactions**. *Chem Rev* 2006, **106**:3188-3209.

44. Friesner RA, Guallar V: ***Ab initio* quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis**. *Annu Rev Phys Chem* 2005, **56**:389-427.

45. Wade RC, Gabdoulline RR, Rienzo FD: **Protein interaction property similarity analysis**. *Int J Quant Chem* 2001, **83**:122-127.

46. Stein M, Gabdoulline RR, Wade RC: **Integrating structural and kinetic enzymatic information in systems biology.** In *NIC Workshop 2006: From Computational Biophysics to Systems Biology*, vol 34. Edited by Meinke J, Zimmermann O, Mohanty S, Hansmann UHE. John von Neumann Institute for Computing; 2006: 129–132.

47. Stein M, Gabdoulline RR, Besson B, Wade RC: **The estimation of kinetic parameters in systems biology by comparing molecular interaction fields of enzymes.** In *Experimental Standard Conditions of Enzyme Characterizations. Proceedings of the 2nd International Beilstein Symposium; 19–23 March 2006. Rudedesheim, Germany*. Edited by Hicks MG, Kettner C. Beilstein Institute; 2007:in press.

48. Cruciani G (Ed): *Molecular Interaction Fields*. Weinheim: WILEY-VCH; 2006.

49. Wade RC, Gabdoulline RR, Ludemann SK, Lounnas V: **Electrostatic steering and ionic tethering in enzyme-ligand binding: insights from simulations**. *Proc Natl Acad Sci USA* 1998, **95**:5942-5949.

50. Warshel A, Sharma PK, Kato M, Xiang Y, Liu H, Olsson MHM: **Electrostatic basis for enzyme catalysis**. *Chem Rev* 2006, **106**:3210-3235.

51. Alves R, Antunes F, Salvador A: **Tools for kinetic modeling of biochemical networks**. *Nat Biotechnol* 2006, **24**:667-672.

52. Klipp E, Herwig R, Kowald A, Wierling C, Lehrach H: *Systems Biology in Practice*. Weinheim: Wiley-VCH; 2005.