CECAM Workshop
Free energy calculations: From theory to applications
Titles and abstracts

Monday 4th

Ben Leimkuhler

The Concept of Accuracy in Molecular Sampling

I will consider the problem of computing thermodynamic averages for a molecular model. In the first part of my talk, I will discuss numerical methods for Brownian and Langevin dynamics. I will show that, in terms of the accuracy of configurational sampling, it is possible to identify a superconvergent scheme (whose order of accuracy is higher than would be expected). I will explore this using both asymptotic expansion and numerical experiments. I will also consider the role of accuracy in relation to the problem of conformational search, and the related issue of numerical stability. Finally, I will discuss numerical comparisons of monte-carlo and molecular dynamics methods on specific classes of models.

◆◆◆◆◆◆◆

Mathias Rousset

Langevin integrators with constraints and free energy calculations.

In this talk, I will give an overview of the formalism of Langevin processes with constraints and its application to molecular simulation.

◆◆◆◆◆◆◆

Martin Spichty

Conformational free energies from non-equilibrium transformations

Gaining insights into the thermodynamics of conformational changes in biomolecules is essential for the understanding of biological functions. Free-energy differences between subensembles of a Hamiltonian system, such as conformational basins, can be extracted from the conditional work data of non-equilibrium transformations, performed in forward and reverse direction of the conformational change.[1] In its previous applications, we applied this approach to relatively small model systems ; [1,2] here we present the application to larger systems, i.e., the Pin1 WW domain (33-residue) and the target T0624 of the CASP9 experiment (71 residues). We demonstrate that the estimate in the conformational free-energy difference is asymptotically unbiased. Remarkably, the error in the estimate (as obtained from a block analysis) is significantly lower than the analytical error of the maximum likelihood estimate.[2] Finally, we propose a method to extract conformational entropy differences from heat flow of such non-equilibrium transformations.

[1] Maragakis P, Spichty M, Karplus M. J Phys Chem B 2008, 112: 6168.
[2] Spichty M, Cecchini M, Karplus M. J Phys Chem Lett 2010, 1:1922.
[3] Shirts MR, Bair E, Hooker G, Pande VS. Phys Rev Lett 2003, 91:140601.

◆◆◆◆◆◆◆

Frédéric Legoll

Two recent works on molecular systems out of equilibrium

In this talk, we will describe two recent works on molecular systems out of equilibrium.

In the first part, we propose a derivation of a nonequilibrium Langevin dynamics for a large particle immersed in a background flow field. Motivation stems from multiscale simulations of liquids, where molecular dynamics models are coupled with continuum descriptions, using e.g. the Navier-Stokes equation. In such computations, one has to simulate a system at the atomistic scale, with a background flow field imposed by the continuum part of the simulation. The question arises on how to exactly do this. In our work, we consider a single large particle, placed in an ideal gas heat bath composed of point particles that are distributed consistently with the background flow field and that interact with the large particle through elastic collisions. In the limit of small bath atom mass, the large particle dynamics converges to a Langevin-type stochastic dynamics, which is parameterized by the background flow field. This derivation follows the ideas of Durr, Goldstein and Lebowitz. The limiting dynamics is found to be similar to the g-SLLOD dynamics. Some numerical experiments illustrate the use of the obtained dynamic to simulate homogeneous liquid materials under flow.

In the second part, we investigate numerically the energy transport properties of a one-dimensional chain of rotors, in a stationary state far from equilibrium. The system being far from equilibrium, linear response theory (which implies e.g. Fourier law) does not apply. The chain is subjected to both thermal and mechanical forcings in a nonequilibrium steady state. We observe that the temperature is maximal in the center of the system, which is an indication of the nonlocal behavior of the system. Our numerical results also show that when the mechanical forcing is strong enough, the energy current can be increased by an inverse temperature gradient.

Joint work with Matthew Dobson, Alessandra Iacobucci, Tony Lelièvre, Stefano Olla and Gabriel Stoltz.

◆◆◆◆◆◆◆

Manuel Athènes

Waste-recycling Monte Carlo : Applications to free energy computations.

In Monte Carlo simulations, waste recycling aims at reducing the variance of the estimates by including information from states rejected by the Metropolis algorithm within statistical averages [1]. The most fruitful applications of this technique are so far those involving free energy computations. We review these applications, both for the alchemical and reaction coordinate cases.

Waste-recycling Monte Carlo simulations often deal with multiple proposals and, for this reason, may be performed on parallel or sequential computer architectures [2]. Multiple proposal may for instance be constructed by exchanging replicas sampling multiple equilibrium distributions within parallel tempering [3], or may be generated through the propagation of nonequilibrium dynamics [4,5]. In the latter case, a multiple time-sliced estimator is to be built upon non-equilibrium work averages [4]. In all these examples, substantial variance reductions are observed numerically.

By formulating waste-recycling in a wider mathematical framework, Delmas and Jourdain proposed an estimator that achieves variance reduction optimally with respect to a control variate that can be evaluated directly using the simulation data [5]. We show in this presentation that the variance reduction obtained in practice by this estimator is close to the maximal theoretical one, despite the inaccuracy in the estimated control variate [6]. The performance of the approach for free energy calculations is also assessed by comparing the measured statistical variances to those of other estimators taken from the literature.

In a general perspective, we advocate that waste-recycling be used in combination with existing methods rather than as an alternative to them, as might be suggested by the overall good performance of the waste-recycling estimators described in this lecture. To illustrate this point, we discuss an implementation of waste-recycling in tandem with dynamical reweighting [7,8], an extended bridge estimator developed for multiple path distributions. We show how time-correlation functions and rate constants may be calculated directly in a rare event problem by both reweighting and recycling samples of dynamically activated trajectories.

[1] D. Frenkel, Waste-recycling Monte Carlo, in ``Computer Simulations in Condensed Matter Systems", Lect. Notes Phys. {\bf 703}, 127 (2006).
[2] J. Kim, J. Rodgers, M. Athènes and B. Smit, Molecular Monte Carlo Simulations Using Graphics Processing Units: To Waste Recycle or Not?, J. Chem. Theory Comput., 7, 3208-3222 (2011)
[3] M. Athènes and F. Calvo, Multiple-Replica Exchange with Information Retrieval, ChemPhysChem, 9, 2332 - 2339 (2008)
[4] M. Athènes, M.-C. Marinica, Free energy reconstruction without post-processing, J. Comput. Phys. 229 7129-7146 (2010)
[5] G. Adjanor, M. Athènes and J. Rodgers, Waste-recycling Monte Carlo with optimal estimates: application to free energy calculations in alloys, J. Chem. Phys.,135, 044127 (2011)
[6] G. Crooks, Path ensemble averages in system driven far from equilibrium, Phys. Rev. E {\bf 61}, 2361 (2000)
[7] J.-F. Delmas and B. Jourdain, Does waste recycling really improve the multi-proposal Metropolis-Hastings algorithm? An analysis based on control variates. J. Appl. Probab. 46, 938 (2009)
[8] Dynamical reweighting: Improved estimates of dynamical properties from simulations at multiple temperatures, J. Chodera, W. Swope, F. Noé, J.-H. Prinz, J. Chem. Phys. 134 244107 (2011).

◆◆◆◆◆◆◆

Ioan Andricioaei

Free energy profiles and time-correlation functions from re-weighted stochastic path integrals

◆◆◆◆◆◆◆

Emad Tajkhorshid

Optimizing the Transition Pathway of Large-Scale Conformational Changes in Membrane Transporters Using Non-Equilibrium Work

Tuesday 5th

Jasna Brujic

Reconstructing free energy profiles from single molecule force spectroscopy experiments

Reconstructing free energy profiles is an important problem in bimolecular reactions, protein folding or allosteric conformational changes. Nonequilibrium trajectories are readily measured experimentally, but their statistical significance and relation to equilibrium system properties still call for rigorous methods of assessment and interpretation. Here we show protein folding trajectories of tandem ubiquitin molecules obtained using single molecule force-clamp spectroscopy. A calibrated constant force is applied to a single protein and its end-to-end length is monitored over time. Exposing a me- chanically stable protein to a high pulling force leads to the unfolding and extension of the polypeptide chain. Quenching the force to a lower value triggers the collapse of the protein from a fully extended state back to a collapsed state with the same end-to-end length as the folded extended length. We then reconstruct the free energy of the end-to-end length at the low force from recordings of these collapsing traces. More generally, we address the question of how to calculate the free energy in situations where the data available is a set of nonequilibrium trajectories that, after initiation, evolve freely until they reach a specific location where they are terminated.

◆◆◆◆◆◆◆

Gabriel Stoltz

Free energy techniques in Bayesian Statistics

The success of free energy techniques to overcome sampling issues motivated the use of these methods in other fields than the ones traditionally covered by statistical physics, such as Bayesian statistics. In particular, the Wang-Landau dynamics attracted a lot of attention since it is very close in its structure to so-called stochastic approximation techniques. I present in this talk some recent mathematical convergence results on the Wang-Landau algorithm, as well as applications of free energy techniques in Bayesian Statistics.

◆◆◆◆◆◆◆

Wei Yang

Synchronous Sampling of Orthogonal Space in Free Energy Simulations

◆◆◆◆◆◆◆

David Kofke

On bias in free-energy calculations, and some free-energy methods for crystalline phases

◆◆◆◆◆◆◆

Lawrence Pratt

Statistical Theory and Learning from Molecular Simulations

This talk follows the idea that coarse graining schemes adopted in molecular simulations implicitly assert statistical mechanical assumptions or approximations. It is thus helpful to be explicit about those assumptions and to test them. We then turn that issue around to analyze progress in learning satisfactory statistical mechanical assumptions from raw simulation results. Molecular quasi-chemical theory (QCT) is discussed as a statistical mechanical structure for learning from simulations. The fundamental idea is definite separation of inner-shell features, that require high-resolution treatment, from outer-shell features that may be treated by proper statistical methods. Applications of QCT to the statistical thermodynamics of liquid water and of Rb+(aq) are discussed in detail. It is claimed that this QCT is the only available statistical thermodynamic theory that establishes a direct, practical role for in situ electronic structure computation of arbitrary accuracy.

◆◆◆◆◆◆◆

Simone Meloni

A rigorous observable for vacancy characterization and diffusion in crystals.

We introduce an observable field to describe the dynamics of a single vacancy in a crystal. This field is the density of a pseudo quantum wavefunction representing the vacancy, which, in turn, is the ground state eigenfunction of an Hamiltonian associated to the potential energy field generated by the atoms in the sample. In our description, the $\hbar^2/2 m$ coefficient of the kinetic energy term is a tunable parameter that makes the density localized in the regions of relevant minima of the potential energy field. Based on this description, we derive a set of collective variables that we use in rare event simulations to identify some of the vacancy diffusion paths in a 2D crystal. Our simulations reveal, in addition

to the simple and expected nearest neighbor hopping path, collective migration mechanisms of the vacancy. These mechanisms involve several lattice sites and produce a long range migration. Finally, we also observed a vacancy induced crystal reorientation process.

◆◆◆◆◆◆◆

Alek Aksimentiev

Modeling small solute transport using atomically precise 3D PMF maps

◆◆◆◆◆◆◆

Ron Elber

Calculation of Free Energies with Milestoning

Wednesday 6th

Eric Vanden-Eijnden

Some Recent Progresses in Free Energy Calculations and Reactive Events Problems.

A few recent techniques to calculate free energies in the context of molecular dynamics simulations and single molecule experiments are discussed: temperature-accelerated molecular dynamics, which is a method to explore fast the important regions in the free energy landscape associated with a set of continuous collective variables without having to know where these regions are beforehand; the single sweep method, which is a variational method to interpolate the free energy globally given a set of mean forces (i.e., a set of gradients of the free energy) calculated at specific points, or centers, on the free energy landscape; and a Voronoi-based free energy method for the calculation of the free energy of the Voronoi tessellation associated with a set of centers. I will also discuss how this last technique can be used in conjunction with the string method, and how kinetic information such as reaction rates can be calculated by milestoning using the edges of a Voronoi tessellation as milestones. Finally, I will discuss a method to reconstruct the free energy profile of a given variable from a set of short nonequilibrium trajectories, obtained by externally driving a system out of equilibrium and subsequently observing its free relaxation. This protocol is not suitable for the Jarzynski equality or the Crooks theorem since the irreversible work on the system is instantaneous, but it is relevant in the context of single molecule force-clamp spectroscopy applied to analyze the hydrophobic collapse of a protein.

◆◆◆◆◆◆◆

Luca Maragliano

Temperature Accelerated Molecular Dynamics Simulations of Biological Molecules

I will present a method designed to explore the free energy landscape of a system with respect to a given set of collective variables. An extended system is introduced where the collective variables are treated as dynamical ones and evolved adiabatically with respect to the physical variables. This allows to explore directly the free energy landscape associated to the collective variables, and the sampling can be accelerated via an artificially high temperature acting on them. I will discuss the technique and its implementation, and illustrate its potential via applications to conformational changes and ligand diffusion in proteins.

◆◆◆◆◆◆◆

James Gumbart

Application of free-energy methods in NAMD to complex biological systems

The highly scalable molecular dynamics program NAMD is supporting an increasing number of free-energy methods, including free-energy perturbation (FEP), adaptive biasing forces (ABF) and most recently replica exchange for collective-variable based calculations, e.g., umbrella sampling (US). Each of these methods has known advantages and disadvantages, although it is not always clear which is better suited to a given problem. In this talk, I will highlight the application of the latter two methods to real biological questions and discuss the problems encountered in the process. In the first case, ABF was used to determine a 2D potential-of-mean force for the folding and translocation of a small oligopeptide inside the protein-conducting channel, SecY. Efficient sampling of the full 2D landscape was greatly hampered by slowly evolving orthogonal degrees of freedom, however, requiring increasing stratification of the reaction coordinates. In the second case, both ABF and US have been applied to determining the PMF for the RMSD of a bound poly-proline ligand to the SH3 domain of Abl kinase as well as the PMF for its extraction from the binding site. As before, sampling with ABF was extremely non-uniform, while, on the other hand, US converged to a final

PMF quickly. That being said, both methods were prone to missing a second minimum in the RMSD PMF, the discovery of which depended entirely on the initial conformation. It is hoped that by drawing attention to the issues arising for real biological systems, i.e., not small, idealized test systems, ideas for improving the practical applicability of the methods can be found.

◆◆◆◆◆◆◆

Jérôme Hénin

Extended-system formulation of the Adaptive Biasing Force method

The Adaptive Biasing Force method is a convenient tool for near-equilibrium sampling and free energy landscape reconstruction, yet its mathematical requirements may prevent its applications to some geometrically complex processes, as well as some combinations of non-orthogonal coordinates. eABF is an extended-system formalism that lifts many technical restrictions on the use of ABF for more complex generalized coordinates, and combinations thereof. I will discuss the principle and implementation of eABF, the choice of tunable parameters, and the convergence and accuracy of relevant free energy estimators.

Thursday 7th

Michael Shirts

Validation efforts and efficiency improvements in free energy calculations

I will discuss three efforts in our group to improve and validate free energy calculations:

I. Development of the multistate Bennett acceptance ratio: We have developed an extension of the Bennett acceptance ratio, the lowest variance method for estimating free energies between two states, to an arbitrary number of states (Shirts and Chodera, J. Chem. Phys, 129, 124105, (2008)). This multistate Bennett acceptance ratio (MBAR) offers several important advantages over the weighted histogram analysis (WHAM) method typically used in multistate free energy analysis. First, MBAR requires no histograms, and thus eliminates this source of bias from binning. Like WHAM, it can be extended to free energies of states not sampled and expectation values. Unlike WHAM, it includes well-behaved asymptotic uncertainty and correlation estimates for free energies and ensemble averages. A Python implementation, including a number of examples, is distributed at https://simtk.org/home/pymbar, and has been significantly improved in recent months from previous releases.

II. Improved sampling of configuration space through Gibbs sampling: The widespread popularity of replica exchange and expanded ensemble algorithms for simulating complex molecular systems has generated great interest in improving the efficiency of these protocols in ways that enhance phase space mixing and therefore increase sampling efficiency. Both classes of algorithms can be considered Gibbs sampling within a Markov chain Monte Carlo (MCMC) framework (Chodera and Shirts, J. Chem. Phys, 135, 194110 (2011)). Clever updating of the thermodynamic states associated with these configurations can substantially increase mixing in a way that still samples from the desired distributions. Simple alternatives to standard protocols can significantly improve mixing of the overall Markov chain, thus reducing simulation time required to converge. These improvements are demonstrated in several common applications, including expanded ensembles among Hamiltonians for estimation of free energies, parallel tempering, and multidimensional umbrella sampling.

III. Improved dissemination and validation of free energy methods: We have started the collaborative website Alchemistry.org as a way to disseminate information about free energy calculations, which includes a public database of free energy references with annotation available through CiteULike. We have also developed a benchmark set for comparison of free energies of molecular transformations in solution (Paliwal and Shirts, J. Chem. Theory Comp, 7, 4115-4134 (2011)). A vast number of new methods for free energy calculations have been developed over the past decade, and deciding which method is most appropriate can be challenging, especially for researchers unacquainted with the details of such computations. This set includes a minimal test, a test of water rearrangement around charges, and a test of solvation of large molecules. This set is distributed with GROMACS, AMBER, and DESMOND input files, and includes 100 uncorrelated solvated starting structures and their energies for each program. We have used this test set to demonstrate that TI and MBAR error estimates are accurate estimates of statistical uncertainty, and that the BAR error estimate underpredicts the true error. We have also shown that bootstrap error estimates are very close to the sample standard deviation estimates for all methods, making bootstrap analysis an attractive tool for estimate statistical uncertainty for molecular simulations.

◆◆◆◆◆◆◆

Edina Rosta

Coupling Hamiltonian replica exchange with the string method in QM/MM simulations

Enhanced sampling methods are essential for accurate free energy simulations of complex systems. Yet, their general efficiency in actual simulations is often not well understood. We developed a kinetic network model to analyze the error and efficiency of replica exchange, and to develop optimal simulation protocols. Guided by our kinetic model, we combined Hamiltonian replica exchange with the finite temperature string method for biomolecular QM/MM simulations of enzymatic catalysis. We use a histogram-free reweighting method to obtain the unbiased free energy surface from combined multidimensional string simulations. Our method allows us to search for the optimal reaction pathway in multiple dimensions and, therefore, to identify the detailed sequence of reaction steps. We use this method to study the cleavage of RNA catalyzed by

ribonuclease H. From our calculations, coupled proton transfer reactions emerge as central factors in catalysis. The proton transfer mechanism revealed by the free energy calculation is consistent with the kinetic effects of protein mutations and RNA backbone modifications. By combining enhanced sampling methods and string simulations with QM/MM calculations, we could elucidate the energetics and mechanism of a complex multistep biomolecular reaction.

◆◆◆◆◆◆◆

Yun Lyna Luo

Performing 2D Hamiltonian Replica-Exchange Umbrella Sampling (US/REMD) for Binding Free Energy Calculations on Petascale Supercomputer

The quantitative determination of thermodynamic free energies is critical in the theoretical biophysics. Alchemical free energy perturbation (FEP) and potential of mean force umbrella sampling (US) are attractive methods to overcome the sampling difficulties of the brute-force molecular dynamics (MD) simulations. Nowadays as the degrees of freedom in the ligand-protein-solvent system increase, even the traditional umbrella sampling schemes and FEP become computationally expensive and often suffer from convergence problem. Here we introduce a novel computational strategy for multi-dimension umbrella sampling simulations of complex biological system on massively distributed supercomputers. The approach involves coupling thousands of windows along 2-dimension order parameters with Hamiltonian-Exchange algorithm to enhance sampling of each window. Hamiltonian-Exchanges are allowed to occur alternatively along the two axes corresponding to the two-dimension reaction coordinates. The method has been implemented in the program CHARMM and most recently in the greatly scalable MD package program NAMD2.9b.

The first application of 2D US/REMD method is calculating the absolute binding free energy of calcium ion to Calbindin D9k. 2D umbrella sampling MD simulation was performed on leadership supercomputer IBM Blue Gene/P. A large number of windows (4096 replicas) are employed to achieve meaningful acceptance ratio (~30%). The tests demonstrate that the 2D Hamiltonian-exchange scheme, compare with regular umbrella sampling, significantly accelerates the configurational sampling of the binding pocket, thereby improving the convergence of the free energy computations. Aided by the massively distributed computing power of leadership platform, 2D US/H-REMD can be extended to more dimensions with irregular shape to enhance the sampling and accelerate the convergence.

◆◆◆◆◆◆◆

Emilio Gallicchio

Order-disorder transitions in protein-ligand binding free energy calculations

Hamiltonian hopping in alchemical space is a powerful tool to accelerate sampling of conformational space and enhance convergence of binding free energy calculations. Conformational transitions that are not directly coupled to the protein-ligand interaction energy are however still poorly addressed and, in our experience, now constitute the next major bottleneck to rapid convergence of binding free energy calculations. In many instances these processes take the form of order-disorder pseudo phase transitions that can be tackled by non-linear alchemical pathway methods recently developed in similar contexts. We will review our work in this area highlighting the theory and numerical procedures, as well as the insights that they can provide for the understanding of the physical mechanisms of binding.

◆◆◆◆◆◆◆

Philippe Hünenberger

Parting from Newton and Boltzmann: Advanced approaches for enhanced sampling in molecular simulations.

Classical computer simulations are currently capable of emulating (a classical approximation of) the natural dynamics of (bio)molecular systems on the nanometric scale at a wallclock speed that is about 13 orders of magnitude slower than that of nature itself. This restricts the simulated timescales to the nanosecond-to-microsecond
regime, and represents a very severe limitation for the investigation of many relevant processes. However, when only thermodynamic (as opposed to dynamical) information is desired, the sampling rate can be considerably enhanced by a number of unphysical approaches, i.e. purely computational tricks without experimental counterpart. In this talk, two recently-developed approaches of this kind will be presented and discussed: the ball-and-stick local elevation umbrella sampling (B&S-LEUS) and the stretched probability extended ensemble dynamics (SPEED) methods.

The B&S-LEUS scheme relies on a memory-based biasing potential explicitly designed to focus and optimize the sampling in theneighborhood of user-defined conformational states (spheres), connected by user-defined transition paths (lines). In this way, the sampling is improved within a conformational space considered by the chemist as relevant (e.g. two known main conformations of a biomolecule), a high number of transitions is guaranteed (permitting a reliable estimation of the relative free energies), and irrelevant conformations are not visited (leading to a high statistical efficiency).

The SPEED approach relies on a new class of non-Hamiltonian equations of motion leading to the sampling of thermodynamic ensembles with arbitrary user-specified distributions of the kinetic and potential energies. For example, it is possible to use different effective temperatures for the kinetic and potential energies, or to stretch one or both distributions to a flat area ranging between two given temperatures. This approach opens up many possibilities for the design of unphysical ensembles and dynamics, specifically tailored for increased sampling efficiency compared to Boltzmann ensembles with Newtonian dynamics.

Peter Bolhuis

Free energy, rates and reaction coordinates of complex systems from reweighted path sampling

Numerical studies of rare events in high-dimensional complex systems, such as chemical reactions in solution, protein conformational changes, or crystal nucleation, often employ biasing along a predefined reaction coordinate to compute the free energy, compute the rate constants, and explore the mechanism. However, the choice of such a low-dimensional collective variable can be very delicate and often influences the results. If chosen poorly, the prediction of the transitions states and mechanisms might be completely wrong. Moreover, instead of imposing a reaction coordinate one would like to extract it from the simulation itself. The transition path sampling framework enables this, but does not immediately yield the free energy and rates. However, the transition interface sampling framework allows a reweighting of the path ensemble that gives rise to a full non-linear description of reaction coordinate, and yields projections of free energy, committor and transmission coefficient landscapes in arbitrary collective variable space. I will illustrate the methodology on a simple two-dimensional system, homogeneous crystal nucleation, as well as a conformational change in a signaling protein.

◆◆◆◆◆◆◆

Normand Mousseau

Constructing pathways for complex motion is slow systems.

The identification of the relevant pathways connecting two states in complex systems is central both for constructing mechanisms and as starting point for computing the free energy that controls its kinetics. For slow systems, it is often difficult to even identify a physical and relevant mechanism for moving from one state - say, the apo state of a protein - to another - the holo state -, let alone evaluation the free-energy associated with this pathway. In this talk, I will present a number of applications of activated methods based on ART nouveau, the activation-relaxation technique, to construct such pathways and identify relevant states for flexible proteins and loops. I will also briefly discuss kinetic ART, an off-lattice kinetic Monte Carlo method with on-the-fly catalog building capabilities, and show a few applications to iron and silicon.

This work is performed in collaboration with L. K. Béland, P. Brommer, L. Dupuis, J.-F. Joly and J.-F. St-Pierre at Université de Montréal, as well as E. Machado-Charry at CEA Saclay and F. El-Mellouhi at Texas A&M in Qatar.

◆◆◆◆◆◆◆

Thomas Simonson

Free energy calculations: approximate methods for biological macromolecules

To compute the thermodynamic properties of biomolecules is often too demanding for fully atomistic models, so that simpler, approximate methods are needed. One main route is to integrate out selected degrees of freedom, particularly those of the solvent, and approximate the resulting potential of mean force by an implicit solvent model. A second route is to model the fluctuations of the system, using linear response theory for example, to obtain extrapolation or perturbation formulas for the free energy and other properties. We are using continuum dielectric implicit solvent models to study ligand binding, acid/base reactions, protein stability, and to perform computational protein design (CPD), which attempts to optimize protein stability through large-scale mutagenesis. The solvent model typically includes a generalized Born electrostatic term, a surface area term, and, more recently, a dispersion term obtained as an integral over the solvent volume. Selected methodology and applications will be described, including constant-pH Monte Carlo simulations of several proteins and CPD calculations to modify the binding specificity of an enzyme.

Friday 8th

Helmut Grubmüller

Computional specificity screening of DNA-binding proteins by exhaustive free energy calculations

Recent advances in free energy calculations have enabled efficient and accurate computation of free energies even involving larger structural changes [1-5]. Here we exploit these advances to exhaustively explore and quantify sequence specific binding of DNA-binding proteins [6]. Chimeric proteins composed of DNA-binding domains and DNA modifying domains allow for precise genome manipulation. A key prerequisite is the specific recognition of a particular nucleotide sequence. A computational framework is presented to automatically set up in silico screening assays and estimate free energy differences using two independent procedures, based on equilibrium and non-equlibrium transformation pathways. The influence of simulation times on the accuracy of both procedures is presented. The binding specificity of a zinc-finger transcription factor to several sequences is calculated, and agreement with experimental data is shown. Finally we propose an in silico screening strategy aimin
g at the derivation of full specificity profiles for DNA-binding proteins.

[1] Goette M, Grubmuller H. Accuracy and convergence of free energy differences calculated from nonequilibrium switching processes. J. Comp. Chem.Vol 30, 447-456 (2009)

[2] Seeliger D, Soeroes S, Klingberg R, Schwarzer D, Grubmüller H, Fischle W. Quantitative assessment of protein interaction with methyl-lysine analogues by hybrid computational and experimental approaches. ACS Chem Biol. 7: 150-4 (2011)
[3] Benedix A, Becker CM, de Groot BL, Caflisch A, Böckmann RA. Predicting Free Energy Changes Using Structural Ensembles. Nature Methods. 6: 3-4 (2009)
[4] Buelens FP, Grubmüller H. Linear-scaling soft-core scheme for alchemical free energy calculations. J. Comp. Chem. 33: 25-33 (2011)
[5] de Jong DH, Schäfer LV, de Vries AH, Marrink SJ, Berendsen HJC, and Grubmüller H. Determining Equilibrium Constants for Dimerization Reactions from Molecular Dynamics Simulations. J. Comput. Chem. 32: 1919-1928 (2011)
[6] Seeliger D, Buelens FP, Goette M, de Groot BL, and Grubmüller H. Towards computional specificity screening of DNA-binding proteins. Nucl.Acids Res. July: 1-10 (2011)

◆◆◆◆◆◆◆

Jose Faraldo-Gomez

Conformational exchange in ATP:Mg2+ studied with a string method

◆◆◆◆◆◆◆

Niels Hansen

Methodological advances in the computation of relative free energies

Computer simulation of the dynamics of biomolecular systems by the molecular dynamics technique yields the possibility of describing structure-energy-function relationships of molecular processes in terms of interactions at the atomic level. Yet, the time and spatial scale of simulations is limited due to finite computing power. Recent advances in simulation methodology to compute relative free energies of different states of a system or of different systems, with respect to one-step perturbation or enveloping distribution sampling (EDS), which can be used to rapidly compute many free energies from a single simulation or to automatically find paths between different states of a system based on optimising the sampling of the relevant parts of configurational or conformational space will be discussed.

J. Chem. Phys. 135 (2011) 024105
J. Phys. Chem. B 115 (2011) 12984-12992 and 13570-13577
J. Chem. Theory Comput. 7 (2012) 3884-3897
J. Comput. Chem. 33 (2012) 640-651

◆◆◆◆◆◆◆

Robert Skeel

Efficient Calculation of Maximum Flux Transition Paths

The minimum resistance path of Berkowitz et al (1983) is a simple way of specifying a "representative" finite-temperature transition path between two given molecular conformations. (An implementation, MaxFlux, is proposed by Huo & Straub (1997).) Starting from the transition path theory of E & Vanden-Eijnden (2006), a generalization of the MaxFlux concept from Cartesian coordinates to generalized coordinates, especially coarse-grained collective variables, is derived in Zhao et (2010). The numerical calculation of involves 3 components: discretization, sampling, and the solution of a system of nonlinear equations, each presenting significant challenges. An effective method that addresses discretization and solution of equations is the semi-implicit simplified string method of Vanden-Eijnden, Ren, E, & Heymann. Its discretization employs upwinding to define the tangent, and its solution method is a novel nonlinear 2-stage process. Presented is an alternative discretization and solution method.

◆◆◆◆◆◆◆