Hyung-June Woo
Department of Biochemistry, Rm. W201
Weill Medical College of Cornell University
1300 York Ave., New York, NY 10021
hyw2001@med.cornell.edu

December 23, 2003

Biocomplexity Faculty Search Committee
c/o Professor Rob de Ruyter van Steveninck
Biocomplexity Institute
Indiana University
Swain Hall West 117
Bloomington, IN 47405-7105

Dear Professor van Steveninck:

In submitting an application for the faculty position at the Biocomplexity Institute, I would like to start by introducing myself with some details not listed specifically in the enclosed C.V: I have received a Ph.D. in 1997 in chemistry from U.C. Berkeley working under Professor David Chandler majoring in theoretical chemistry, and after completing 2 years of national service in Korea, have been working as a postdoctoral researcher at Iowa State University, University of Massachusetts at Amherst, and now at Cornell medical school in New York City. As a graduate student at Berkeley, my research centered on the applications of statistical mechanics and thermodynamics to physical chemistry, and more specifically, on theoretical studies of various self-assembly processes in complex fluids, and in particular, of surfactant mixtures such as micelle formation, lamellar phases, and bicontinuous microemulsions, and have acquired a solid background in analytical theories of phase transitions, liquid state theories, and complex fluids.

While at Iowa State, I have worked on applying and extending density functional theories along with molecular simulations to the freezing of anisotropic fluids. At UMass, I had the opportunity of gaining experience in some of the research areas of chemical engineering, and molecular thermodynamics in particular, by studying properties of gas adsorption and dynamics in mesoporous materials, and modeling the phase coexistence properties of network-forming fluids such as water. Currently at Cornell, I am expanding the breadth of my research experience to computational biology, by working on projects of developing computational methodologies for applications to structural biology; implementations of the grand canonical simulation methods for the solvation of proteins, and calculations of the free energy of binding of ligands to enzymes.

Over the postdoctoral years, I have also been independently pursuing a theoretical study of nonequilibrium dynamics and hydrodynamics, and in particular, the statistical mechanical link between thermodynamics and hydrodynamics far from equilibrium. Preliminary accounts of the work have been published last year, and in this month. The work, if pursued and developed further aggressively, will have practical applications to areas in physical chemistry, particularly in theories of reaction rates, and stochastic dynamical description of biomolecular processes as described in more detail in the research plan enclosed. I believe the topic is especially promising as a research area to be explored in the Biocomplexity Institute at the Indiana University, since it can potentially provide a well-founded physical basis for the thermodynamic driving force of biological organizations.

With a firm background and research experiences in various fields of theoretical chemistry, and in particular, statistical mechanics of simple and complex fluids, I believe I will be able to make worthwhile contributions to the overall research efforts of the Institute with the development and applications of the theory of nonequilibrium dynamics to biophysical processes and complex fluids, while effectively performing other duties of the University with an appointment at the Department of Chemistry.

Please find enclosed the C.V., Research Plan, and copies of a recent paper published. In addition, I have arranged for the recommendation letters to be sent from Professor David Chandler, Department of Chemistry, University of California, Berkeley; Professor Benoit Roux, Department of Biochemistry, Weill Medical College of Cornell University, and Professor Peter Monson, Department of Chemical Engineering, University of Massachusetts, Amherst. Please let me know if there is any further information which could help assess my qualifications. I look forward to the opportunity to discuss further my backgrounds and the possibility at the Institute.

Sincerely,

Hyung-June Woo

Research Plan

Dynamics of nonequilibrium processes: hydrodynamics, stochastic dynamics, and their applications in biological systems

H.-J. Woo

Introduction

The central theme of my research plan is the study of nonequilibrium processes using statistical mechanics, stochastic dynamics, and transport theory. Description of nonequilibrium processes, such as hydrodynamic pattern formations, networks of mutually catalyzing biochemical reactions, or the morphological development of embryos, remains as one of the main challenges of physical sciences. Many aspects of the phenomena observed in nonequilibrium dynamical processes closely parallel the phenomenologies of equilibrium phase transitions. The dynamical equations of motion governing the time-evolution of the relevant macroscopic variables are intrinsically nonlinear, and a plethora of patterns and self-organized structures arise as the instabilities of the nonlinear equations develop [1], as in phase transitions where the equation of state of one phase becomes unstable toward another.

Recent developments both in physical sciences and molecular biology have accumulated much of the necessary conceptual foundations as well as factual data of building blocks, to allow for attempts to provide global and quantitative descriptions of such complex, self-organizing systems. I plan to pursue systematic and quantitative studies of such systems, both on the fundamental level by extending currently available theoretical formulations, and on the more specific level by applying the tools to model systems of hydrodynamics, molecular dynamics, and biological systems.

Theoretical study of nonequilibrim dynamics and hydrodynamics

One of the most well-characterized classes of self-organizing phenomena far from equilibrium is the hydrodynamic pattern formation, such as the Rayleigh-Benard convection or the Taylor-Couette flow [1]. Currently remaining unsolved is the 'pattern selection problem,' the question of which of the newly-emerged multiple solutions would be most stable in reality. A more general formalism appears necessary, encompassing not only the deterministic hydrodynamic solutions, but also the stochastic deviations which become non-negligible near the instabilities.

As a possible candidate for such a theory, I plan to pursue further an idea that was proposed recently [2, 3], based on the Boltzmann entropy of hydrodynamic states in the microcanonical ensemble. An explicit expression can be derived for the probability distribution of the time-dependent trajectories of hydrodynamic variables, from which the conventional Navier-Stokes-type deterministic solution emerges as the extremal path [3]. An immediate goal in this direction will be to derive from the probability of trajectories the stationary probability distribution of hydrodynamic variables in nonequilibrium stationary states. The method for the derivation has in fact been developed and used previously in similar situations [4, 5, 6]. The results will be tested for the experimental data of pattern formation in the Rayleigh-Benard convection; in particular, regarding the selection and distribution of wave-numbers of the roll states [7].

Stochastic dynamics of chemical and biochemical processes

In molecular systems, it is often of interest to study the dynamical evolution of "reaction coordinates," collective degrees of freedom of interest undergoing stochastic evolutions due to the coupling to the bath [8]. The dynamical evolution equation, such as the Langevin or Fokker-Planck equations, involves the potential of mean force or free energy, which arises from the collective effects of the bath on the reaction coordinates, and is a highly nonlinear, multi-dimensional quantity in general. The computational techniques for calculations of such a potential of mean force are now well-established, and are being used extensively in computational biology of proteins [9]. The classic Kramer's theory of reaction rate [8], based on the analysis of the stochastic dynamics of a one-dimensional reaction coordinate on a symmetric bistable free energy surface, forms the theoretical foundation of much of the modern chemical applications of the rate theories. In many systems of interest, however, a minimal description of the system requires multi-dimensional reaction coordinates, whose study so far has only been made possible with direct numerical simulations of the stochastic dynamics. In particular, the dynamical evolution of the reaction coordinates in a multi-dimensional free energy surface is described by the equation of motion only locally, and the global statistics of the multiple possible pathways from one stable state to another and their associated rate constants cannot be obtained easily.

The nature of the problem is in fact closely related to the hydrodynamic case described in the previous section. A more general formulation for the global statistics of the dynamical evolution of reaction coordinates is desirable, which would yield the Langevin-type equation as its extremal Euler-Lagrange equation describing the most probable time-evolution. More specifically, a nearly-straightforward adaptation of the formalism of Ref. [3] should be possible, in which the reaction coordinates, potential of mean force, and the diffusion coefficients would replace the hydrodynamic variables, entropy production, and the transport coefficients, respectively. The formulation of a similar theory would yield an expression for the probability distribution of the trajectories of reaction coordinates, and the relative probabilities as well as rates of multiple pathways, each locally stable on the multi-dimensional free energy surface, could be obtained. A first step will be to derive known results of the Kramer's theory [8] using such a path integral-type formalism, and examine regimes in the parameter space inaccessible within the convectional theory.

Of also interest will be possible applications to simulations of dynamical processes using the method of the transition path sampling [10], whose recent development has had a great impact on studies of a variety of dynamical processes. In transition path sampling simulations, Monte Carlo samplings are performed on the trajectory space, allowing one to obtain information of the most relevant pathways and intermediates. With weights of arbitrary trajectories given theoretically, such Monte Carlo samplings could also be performed on coarse-grained reaction coordinates evolving non-deterministically.

As a first multi-dimensional application, I plan to consider the diffusion of potassium ions through the membrane protein, the KcsA channel, which plays a pivotal role in many physiological processes in biological cells. The potassium channels are proteins spanning the cell membranes, regulating the overall concentration gradient of potassium ions and the action potential between the cytoplasm and the exterior, and allows for the selective conduction of potassium ions [11]. An important advance in our understanding of the detailed mechanism of ion conduction was the calculation of the potential of mean force of ions on the channel with computer simulations [12]. It was revealed that 2 or 3 ions occupy an alternating series of well-defined local free energy minima along the selectivity filter, allowing the conduction via the "knock-on" mechanism. Since multiple ions are involved the potential of mean force is multi-dimensional, and multiple pathways exist within the surface. Currently one needs to perform Brownian dynamics simulations, where the Langevin equation with the potential of mean force and the diffusion coefficient as the inputs is integrated numerically [13], to obtain information regarding the rate as well as relative importance of various pathways. With the theory giving the global probability distribution of all possible trajectories, one would be able to obtain the relative populations of pathways as well as their rates analytically. In particular, a realistic and well-controllable model of the free energy landscape of the channel could be devised based on the published simulation data Ref. [12], and the results of the theory compared with those from direct simulations.

Molecular motors: mechanism of work production in biological cells

A potentially fruitful application of the studies of nonequilibrium dynamics on the fundamental levels described above is the study of the mechanism underlying the operations of the motor proteins within biological cells. Motor proteins are the biological equivalents of the macroscopic engines, driving a wide variety of cellular processes ranging from muscle contraction, cell division, and various cytoplasmic transport mechanisms of biochemical materials [15]. With recent developments in molecular and structural biology revealing in unprecedented detail the molecular mechanisms of the various constituent proteins involved in such processes, it should now be possible to combine theoretical studies of the foundations underlying the operation of such microscopic engines based on simplified models [16] with the realistic molecular structural information. It is likely that such studies of the principle of work production on microscopic scales should accompany the attempts to fabricate artificial machines on the nanoscale, which are being proposed in the recent drive for nanotechnology.

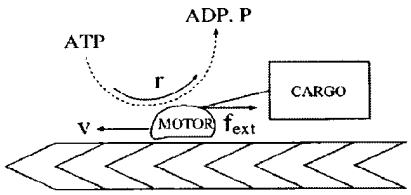

Figure 1: A cartoon of a molecular motor propelling itself along a filament. Reproduced from Ref. [14].

On the fundamental level, the switch from the macroscopic to microscopic scales entails a few changes in the characteristic physics of the operation of an engine: fluctuations become non-negligible [18], and the well-controlled temperature gradient, which forms the basis of heat engines in the macroscopic scale, becomes infeasible [14]. Instead, molecular motor proteins utilize the chemical energy released by the hydrolysis of ATP to ADP, to propel themselves uni-directionally. As prominent examples, myosin and kinesin, which power all muscle movements, propel themselves along the molecular "railway track" of actin or microtubule filaments [17]. Figure 1 shows a schematic arrangement of the constituents involved in such a movement.

In both cases, the "power stroke," a molecular equivalent of the expansion stroke of a piston in a heat engine, involves an elementary movement step of the motor protein along one unit of the polymerized filament, driven by the conformational change induced from the binding of ATP (ADP+P_i). Denoting the two conformational states as α and β connected by the interconversion process

$$\alpha + \mathrm{ATP} \underset{k_{\alpha\beta}}{\overset{k_{\beta\alpha}}{\rightleftharpoons}} \beta + \mathrm{ADP} + \mathrm{P}_i, \tag{1}$$

one can construct a minimal model and utilize the following stochastic equation:

$$\begin{aligned} \frac{\partial p_\alpha}{\partial t} &= \left[-\frac{\partial}{\partial x} v + \frac{\partial}{\partial v} \left[m^{-1} W'_\alpha(x) + \gamma v \right] + \frac{\gamma k_B T}{m} \frac{\partial^2}{\partial v^2} \right] p_\alpha \\ &+ k_{\alpha\beta} p_\beta C_{\mathrm{ADP}} C_{\mathrm{P}_i} - k_{\beta\alpha} p_\alpha C_{\mathrm{ATP}}, \end{aligned} \tag{2}$$

and an analogous equation for p_β, where $p_\alpha = p_\alpha(x, v)$ is the probability of finding a protein with the center of mass position x along the filament, velocity v, and in the conformational state α, m is the mass of the protein, and $k_B T$ is the Boltzmann constant times

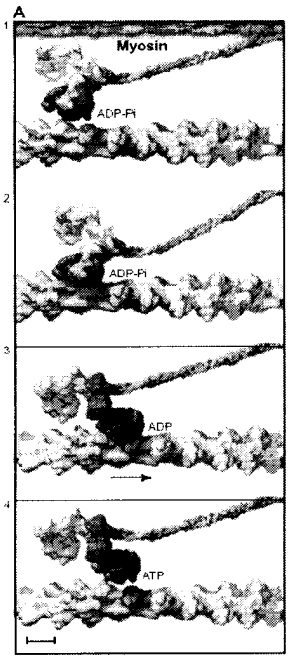

Figure 2: Molecular structures of the myosin complex moving along the actin filament. Reproduced from Ref. [17].

temperature. C_A denotes the concentration of a species A. Equation (2) is a Fokker-Planck equation, with source terms accounting for the inter-conversion process. Key parameters of Eq. (2) are the potentials of mean force $W_\alpha(x)$ and $W_\beta(x)$ of the two states along the coordinate, and the rate constants, $k_{\alpha\beta}$ and $k_{\beta\alpha}$.

Based on the formulation described above, a two-sided approach to the problem is possible: on the one hand, one can study the general physics of Eq. (2); the stochastic dynamics of a two-state system with reversible inter-conversion processes far from equilibrium, riding on the periodic potentials of mean force, for which simple model functional forms can be assumed. The properties of a Brownian ratchet along a periodic asymmetric potential had been considered previously as a general model for various physical systems [19]. Equation (2) adds a novel feature to such a model, the mechanism of the two-state inter-conversion with different potentials. Whereas the asymmetry of the periodic potential acts as the source of the net directed motion of the Brownian particle in the conventional models, it is the chemical potential of ATP, and the resultant asymmetry of the reaction (1), that drives the motion in Eq. (2).

Equally important in the overall efforts to understand the operations of the motor proteins will be the study of realistic molecular models based on the known crystal structures of the constituents [17], which would ultimately yield the appropriate values of the parameters such as $W_\alpha(x)$ and $k_{\alpha\beta}$ of the stochastic model, and allow for quantitative comparisons with experiments. The primary tool of the computational study will be the calculation of the potential of mean force, both for the conformational inter-conversion process of the protein, as well as the periodic free energy profiles $W_\alpha(x)$ and $W_\beta(x)$. For the conformational isomerization process, one can utilize the root-mean square deviation (RMSD) of the atomic coordinates from a fixed specified structure as the reaction coordinate, and perform free energy calculations via umbrella sampling methods [9]. The known X-ray crystal structures of the two isomers α and β of myosin [20, 21] can be defined as the target structures of the RMSD reaction coordinate. With the free energy profile connecting the two conformational states, one could estimate the rates using the transition state theory or its refinements, and ultimately the new theoretical approaches outlined in previous sections. The calculation of the periodic potentials along the filaments could be done in a number of ways: one possibility is to employ the continuum electrostatic treatment, the Poisson-Boltzmann method, where the interaction free energy between two proteins in solution is calculated by numerical solution of the Debye-Hückel-type Poisson equation of electrostatics [22].

References

[1] M. C. Cross and P. C. Hohenberg. Pattern formation outside of equilibrium. *Rev. Mod. Phys.*, 65:851, 1993.

[2] H.-J. Woo. Variational formulation of nonequilibrium thermodynamics for hydrodynamic pattern formations. *Phys. Rev. E*, 66:066104, 2002.

[3] H.-J. Woo. Statistics of nonequilibrium trajectories and pattern selection. *Europhys. Lett.*, 64:627–633, 2003.

[4] R. Kubo, K. Matsuo, and K. Kitahara. Fluctuation and relaxation of macrovariables. *J. Stat. Phys.*, 9:51–96, 1973.

[5] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Fluctuations in stationary nonequilibrium states of irreversible processes. *Phys. Rev. Lett.*, 87:040601, 2001.

[6] H. Goldstein. *Classical Mechanics.* Addison-Wesley, New York, 1980.

[7] M. S. heutmaker and J. P. Gollub. Wave-vector field of convective flow patterns. *Phys. Rev. A*, 35:242, 1987.

[8] P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.*, 62:251, 1990.

[9] T. Simonson, G. Archontis, and M. Karplus. Free energy simulations come of age: protein-ligand recognition. *Acc. Chem. Res.*, 35:430–437, 2002.

[10] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler. Transition path sampling: throwing ropes over rough mountaion passes, in the dark. *Ann. Rev. Phys. Chem.*, 53:291–318, 2002.

[11] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. L. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, and R. MacKinnon. The structure of the potassium channel: Molecular basis of K^+ conduction and selectivity. *Science*, 280:69–77, 1998.

[12] S. Bernèche and B. Roux. Energetics of ion conduction through the K^+ channel. *Nature*, 414:73–77, 2001.

[13] D. L. Ermak and J. A. McCammon. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.*, 69:1352–1360, 1978.

[14] F. Jülicher, A. Ajdari, and J. Prost. Modeling molecular motors. *Rev. Mod. Phys.*, 69:1269–1281, 1997.

[15] G. Oster. Darwin's motors. *Nature*, 417:25, 2002.

[16] M. E. Fisher and A. B. Kolomeisky. The force exerted by a molecular motor. *Proc. Nat. Acad. Sci.*, 96:6597–6602, 1999.

[17] R. D. Vale and R. A. Milligan. The way things move: Looking under the hood of molecular motor proteins. *Science*, 288:88–95, 2000.

[18] G. M. Wang, E. M. Sevick, Emil Mittag, Debra J. Searles, and Denis J. Evans. Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales. *Phys. Rev. Lett.*, 89:050601, 2002.

[19] P. Reimann. Brownian motors: noisy transport far from equilibrium. *Phys. Rep.*, 361:57–265, 2002.

[20] R. Dominguez, Y. Freyzon, K. M. Trybus, and C. Cohen. Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: visualization of the pre-power stroke state. *Cell*, 94:559–71, 1998.

[21] I. Rayment, W. R. Rypniewski, K. Schmidt-Base, R. Smith, D. R. Tomchick, M. M. Benning, D. A. Winkelmann, G. Wesenberg, and H. M. Holden. Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science*, 261:50–58, 1993.

[22] K. A. Sharp and B. Honig. Electrostatic interactions in macromolecules - theory and applications. *Ann. Rev. Biophys. Biophys. Chem.*, 19:301–332, 1990.