HOWARD HUGHES MEDICAL INSTITUTE
UNIVERSITY OF CALIFORNIA, SAN DIEGO UCSD

TEL: 858-822-1469
FAX: 858-534-7042
EMAIL: rhenchma@mccammon.ucsd.edu
HTTP://mccammon.ucsd.edu/~rhenchma

Richard H. Henchman
Howard Hughes Medical Institute
University of California, San Diego
Department of Chemistry & Biochemistry
9500 Gilman Drive Dept. 0365
La Jolla, CA 92093-0365

November 7, 2003

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

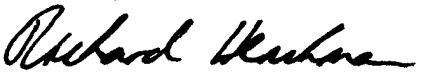
Dear Prof. de Ruyter van Steveninck,

I am excited to learn of the faculty position in Biocomplexity that you recently advertised in *Nature* and would like to be considered for the position at the Assistant Professor level. My preferred department is Chemistry and my focus would be on Computational Biology, Biochemistry, and Biophysics.

In my graduate work under Dr. Jonathan Essex I studied a model host-guest system using computer simulation both to understand factors contributing to molecular recognition and as a convenient system to test and develop methods relating to force field parametrization, free energy calculations, and Monte Carlo sampling. My postdoctoral studies under Prof. J. Andrew McCammon broadened my research to include many more problems of a biological nature. These included protein solvation, protein dynamics, and pharmaceutical drug design, and involved collaborations with researchers in pharmacology, organic and bioinorganic chemistry. My chief interest now lies in leading a research group to advance the development of computational methods relating to the calculation of thermodynamic properties and conformational sampling that are applicable to many problems encompassed by chemistry, biology and physics. I further wish to apply these new as well as existing computational methods in collaboration with experimental research groups. I also look forward to the opportunity to teach courses related to this area.

My curriculum vitae, statements of my research and teaching interests, summary of my research experience, and reprints of some of my recent papers are enclosed. I have also arranged for letters of recommendation to be sent to you from the three referees listed in my curriculum vitae. Thank you for your consideration.

Sincerely,



Richard Henchman

# Research Interests of Richard H. Henchman

It is an ideal time to be involved in computational biology, biochemistry and biophysics research. The field is relatively young and the prospects of applying new ideas and algorithms, increasing computer power, and effective collaborations to gain an understanding of a vastly expanding body of knowledge are tantalizing. My overall approach to research is twofold. The first part is to develop new computational methods that calculate the thermodynamic properties of a given system and methods that determine all the system's important conformations. I believe that the inadequacy of current methods in these areas is a major limitation facing their use in biology, biochemistry and biophysics, and addressing them is imperative. The second part of my research involves forming collaborations with fellow experimental researchers and applying computational methods to assist with their particular studies, gaining the benefits of multiple approaches, alternative viewpoints, and the productive transfer of successful ideas between different projects.

Being able to calculate the thermodynamic properties of a specified state of a system, in particular, its free energy, is a very desirable objective, since one may then predict which state the system adopts at equilibrium. Important molecular properties in biology that could be calculated in this way include the molecule's solubility (Orozco and Javier Luque, 2000), conformation, and strength of association with other molecules (Gohlke and Klebe, 2002). There are three general types of commonly used free energy methods. Knowledge based methods are generally simple, fast, and rely on generic descriptors rather than purely an atomic force field but suffer from inaccuracy and are limited to the types of system and thermodynamic conditions for which they are parameterized, reducing their predictive power for new systems. The perturbation class of methods measure the change in the atomistic force field energy landscape between two states and are exact in principle. However, they are computationally expensive, only give free energy differences, and convergence problems limit the size of the transformation. Single point free energy methods attempt to calculate the absolute free energy of a given

state based on the energy landscape sampled for that state, again based purely on an atomistic force field, a feature that should render them applicable to any system within the accuracy of the force field. There are a variety of different formulations of this kind in the literature, usually with a continuum treatment of the solvent, with moderate success. My first purpose is to build on the single point free energy methods and devise a new means to calculate the free energy of the solvent component directly from an explicit solvent simulation. I have already demonstrated how this may be done for the simplest test case, a simple liquid, using the forces that the molecule feels rather than its displacement as a means to determine the shape of the energy landscape (Henchman, 2003). I further plan to extend the approach to pure water which has rotational degrees of freedom and multiple rotational states, although other solvents will be possible. Second, I will account for the anisotropic nature of solvent around a solute to provide a direct means of evaluating the solute's free energy of solvation, an important determinant of its solubility. Solvation plays a particularly important role in molecular association via the hydrophobic effect and less stable water molecules can also reveal possible ligand binding sites or protein interaction surfaces. The final step is to account for molecules with multiple conformations. An important note about free energy calculations is that they require extensive sampling to obtain a representative sample of the whole energy landscape, hence the importance of my other planned area of research, described in the following section.

Being able to generate all the important conformations of a molecule is another desirable objective. This would improve the ability to study protein and nucleic acid folding, stability, conformational change, and molecular association. Knowing the multiple states of a molecule should also reveal aspects of its function. Because the energy landscapes of these systems are so vast and rugged, established sampling techniques such as Monte Carlo and molecular dynamics, which yield sufficient sampling for simpler systems, are limited in their ability to generate all the important conformations of typical biomolecules. While many new sampling techniques have recently been proposed, the only practical and more effective addition to the simulation toolbox so far has been the replica exchange method (Gnanakaran et al., 2003).

Even then, replica exchange is expensive and remains to be made workable for large systems. I believe that a key limiting feature, which many methods possess, is their adherence to equilibrium sampling. Equilibrium sampling allows thermodynamic averages to be calculated from trivial averaging but promotes poor sampling since most of the simulation is spent in stable states and transitions over high energy barriers or over large distances in conformational space rarely take place. Accelerating transitions or introducing biases away from stable states may alleviate this effect but the formulation of the correction factors during the simulation becomes increasingly problematic for large, complex systems. Therefore, the strategy that I propose to adopt is to sacrifice the equilibrium sampling feature altogether and instead use new and existing sampling techniques which are non-equilibrium in nature so that they may explore the energy landscape significantly faster while keeping close to the important, equilibrium configurations. This type of method would be useful in its own right as a means of generating a collection of reasonably stable configurations that a system may adopt. The power of the method, my main objective, comes from combining the sampled configurations with the free energy methods discussed earlier so that each local state may be appropriately weighted so as to recover the crucial equilibrium thermodynamic information.

There are other benefits to improving free energy and sampling methods. Indirectly, they would improve the viability of modeling systems with more realistic but expensive molecular representations that make use of polarization, semi-empirical quantum, and ab initio quantum methods. They would lead to much greater overlap with experimental results to provide further tests of modeling approaches, help elucidate physical principles, formulate new, simple models, assist in the interpretation of experimental data, and improve the ability to design pharmaceutical therapeutics and novel biomolecules. Given my more theoretical background, I enjoy studying many different aspects of biology, biochemistry and biophysics and have applied computational methods to a wide variety of problems and have established a number of fruitful collaborations with experimental groups. Therefore, I would like the opportunity to apply computational methods in collaboration with experimental researchers on a variety of cutting edge

problems and reap the benefits of combining my skills with theirs. In the process, I will also learn the most important types of new methods that are needed, gain inspiration for ways to develop them, and apply approaches that were successful in one system to the next. One particular system I have taken a great interest in from my postdoctoral work involves the ligand-gated ion channels, in particular the nicotinic receptor (Karlin, 2002). Ligand-gated ion channels make up an important kind of signal transduction pathway. Understanding their ligand binding, conformational states, and ion conductance properties at a molecular level would greatly facilitate in clarifying their role at a more macroscopic level.

To summarize, the research that I plan to undertake myself is mainly of a methodological nature focused on free energy calculations and sampling methods. Improving these methods would have a wide and beneficial impact on the usefulness of computational methods in biology, biochemistry and biophysics. I would also like to collaborate extensively with other researchers in applying new and existing computational methods to a variety of challenges.


1. Orozco, M. and F. Javier Luque. Theoretical methods for the description of the solvent effect in biomolecular systems. Chem. Rev. 2000, 100, 4187-4225.

2. Gohlke, H. and G. Klebe. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. Angew. Chem. Int. Ed. 2002, 41, 2644-2676.

3. Henchman, R. H. Partition function for a simple liquid using cell theory parametrized by computer simulation. J. Chem. Phys. 2003, 119, 400-406.

4. Gnanakaran, S., H. Nymeyer, J. Portman, K. Y. Sanbonmatsu, and A. E. Garcia. Peptide folding simulations. Curr. Opin. Str. Biol. 2003, 13, 168-174.

5. Karlin, A. Emerging structure of the nicotinic acetylcholine receptors. Nat. Rev. Neurosci. 2002, 3, 102-114.

# Research Experience of Richard H. Henchman

My current area of research is computational chemistry and biophysics. I have been engaged in a number of different research topics. These include understanding the function of the nicotinic receptor and acetylcholinesterase (AChE), examining protein solvation, calculating the solvent free energy and the association free energy, docking and designing drugs for HIV integrase and stromelysin, determining the binding of amino acids to a synthetic host, devising a model for the diffusion of small molecules in glassy polymers, and various other undergraduate projects. Studying these systems has familiarized me with a wide range of simulation techniques: molecular dynamics (MD), Monte Carlo (MC), and Brownian dynamics simulation methods; free energy perturbation (FEP), umbrella sampling, and single point free energy calculations; Poisson-Boltzmann and Generalized Born (GB) solvation models; the derivations of charge and dihedral force field parameters; quantum mechanics calculations; small molecule docking and drug design; kinetic rates from transition state theory; extensive simulation analysis. I am familiar with a wide range of simulation, analysis, and visualization software, program my own simulation and analysis routines in Fortran, C, Perl and Python, and am well-versed in Linux and Windows operating systems. A more detailed description of each project now follows. The references refer to the publication list in my curriculum vitae.

My most recent project has made use of extensive MD simulations on the $\alpha 7$ nicotinic receptor ligand binding domain to study how bound ligands affect the receptor's shape. Simulations were run in the apo form, with the antagonist d-tubocurarine, agonist acetylcholine, and potentiator calcium ion with acetylcholine. In the apo form or with antagonist bound, the homomeric pentamer was found to relax toward an asymmetric state[4] that has also been observed experimentally for the neuromuscular receptor from electron microscopy structures. However, with agonist bound, the receptor largely maintained the fivefold symmetry and with agonist and potentiator, the receptor even expanded a small amount. These results, still to be published, imply that the receptor has to assume the symmetric, expanded state to conduct ions.

Working with Kaihsu Tai, we examined the means by which the snake toxin fasciculin

II binds to and inhibits AChE using an MD simulation of the complex.[11] Fasciculin not only blocks and shrinks the entrance to the AChE active site, but also allosterically disrupts the position of the active site catalytic residues. We developed a new analysis technique termed the "porcupine plot" which reveals all the residues that contribute to the opening of each passage leading to the active site. Indeed, the opening of each passage involved the concerted motion of many residues. To further understand the process of ligand binding, I assisted Jennifer Bui in calculating the potential of mean force for moving a tetramethylammonium ion into the active site of AChE.[5] Even though the passage in the relaxed state is much narrower than the ligand size, the ligand easily pushed through the passage with a barrier of only 8 kJ $mol^{-1}$, too small to be rate-limiting to binding and consistent with the binding being diffusion controlled.

The AChE simulation was also used to examine the structure and dynamics of water molecules in the enzyme's active site gorge[13] and around the entire protein.[9] To do this I examined three different ways to determine the water density from the simulation and extract specific hydration sites where water molecules prefer to reside.[12] Two factors aided in resolving better densities. The first was to construct the water density from the time-averaged positions of water molecules rather than their instantaneous positions. The second was to calculate the water positions in the frame of the nearby residues rather the total protein frame. The resulting collection of water sites provided a convenient way to characterize water structure and track water motion around AChE at an unprecedented level of detail and demonstrate the real time motion of molecules entering and leaving the active site. The 10 ns length of the simulation proved sufficient to determine the time scales of all but the most buried water molecules.

To calculate the free energy of solvent from an explicit solvent simulation, I devised a method based on cell theory that achieves this goal competently for a simple liquid.[7] I plan to extend this model to water, water around solutes, and molecules with internal degrees of freedom. With Jessica Swanson we have derived a formulation of the association free energy when two molecules bind using quasi-harmonic analysis and an implicit solvent model.[1]

I have worked with Julie Schames on two drug design projects. The first project ex-

amined the binding mode of an existing inhibitor 5CITEP to HIV integrase. Docking of this inhibitor to multiple snapshots of the complex from a MD simulation revealed the presence of a new binding pocket adjacent to the active site that only appeared during the simulation. Consequently, we designed inhibitors that could take advantage of both this site and the active site binding position seen in the crystal structure.[2] The second drug design project aims to design inhibitors for stromelysin, a matrix metalloproteinase, in collaboration with the bioinorganic research group of Prof. Seth Cohen. In this work we developed a new docking method that combines small molecule crystal structures of the difficult-to-model metal ion complex with multiple force field minimizations of the remainder of the ligand.[6]

My graduate work was focused on understanding the binding of amino acid derivatives to a host molecule, macrobicycle 12.[3] FEP calculations using MC simulations supported the experimental finding that L isomers bound preferentially over D inside the host and that the amide bond for this binding mode was stabilized in the *cis* conformation. Part of the stabilization was attributed to a hydrogen bond and the remainder came from an internal stabilization that occurred when the $\chi_1$ side chain dihedral of the amino acid derivative was bound in a higher energy conformation. Further, the *cis* stabilization was predicted to be reduced for glycine and indeed this result was supported by a later experiment. In the course of the work I parametrized a GB model for chloroform and developed a charge derivation method based on the RESP method. Called the REPD method, it fits charges to a molecule's electrostatic potential with a restraint chosen to make the charges reproduce OPLS charges, the force field used in the rest of the system.[15] The suitability of these charges was assessed by calculating the free energies of hydration for many small molecules and comparing with experiment.[14]

In my main thesis during my undergraduate degree, I developed a model to predict the diffusion of small molecules in glassy polymers.[16] The model was inspired by MC simulation data of methane diffusing in atactic polypropylene, was based on transition state theory, and used the isothermal compressibility of the polymer as a measure of how easily a molecule diffuses through it. My other undergraduate projects are summarized in my curriculum vitae.

# Teaching Interests of Richard H. Henchman

The path of learning in any field of endeavor toward a position where one may contribute or apply new knowledge and understanding is long. Each area of science is amassing a huge body of knowledge, making it demanding for students to master them. This problem is invariably compounded for areas that are interdisciplinary. Hence the ever-present need for effective and evolving education. There are three aspects in particular that I feel contribute significantly to achieving this goal. These are providing encouragement, early exposure to research, and concept learning. These points are elaborated below, together with my preferred teaching courses.

An essential aspect of teaching is to provide encouragement. The learning environment should be welcoming and enjoyable and the course material interesting and varied with plenty of real examples that captivate the student and instill in them a desire to learn more. This is because learning is most effective when the student wishes to learn, and encouragement gives them greater confidence to explore new ideas and expand their abilities. Courses must appeal to and benefit a diversity of students both academically and culturally. Bright students must be pushed, ideally to the point of being able to learn for themselves, while struggling students should be provided with additional resources so they do not get left behind. As a teacher, I prefer having an open door and a willingness to help all students, particularly those from disadvantaged backgrounds.

A vital component of teaching is to introduce undergraduate students to research by giving them projects related to my own research, ideally leading to publication. Conducting frontier research motivates students to want to learn more, provides them with an opportunity to decide if they wish to pursue a research career, and develops their abilities to conduct research if they do. Graduate students should be offered challenging projects and allowed to explore them increasingly in their own way, learning how to conduct research on their own and being offered guidance where necessary. Presenting and being exposed to other research at national conferences would be an integral part of their education.

On a more philosophical note, an important way to address the difficulty of learning the tremendous amount of knowledge is to emphasize the learning of concepts at all levels of education. Facts are required to get started but later, developing concepts is necessary for the following reasons. First, concepts are an efficient means to understand knowledge. A vastly smaller number of concepts is required to gain proficiency in an area than a knowledge of facts. Second, memorized facts tend to be forgotten, while cross-linked concepts are more reliably retained. Third, concepts form a useful framework around which to store facts. Fourth, they are more easily tailored to new situations, giving greater independence. Fifth, they are more adaptable to change when old concepts become outmoded. Sixth, a powerful means of creating new ideas is by evolving or combining concepts, and even applying a single concept to a new problem may be productive. Seventh, more satisfaction and enjoyment is usually derived from an understanding of concepts. The difficulty with concepts is that they are much harder to teach than facts. Facts make up the information that is communicated directly and inevitably make up much of the teaching content. However, concepts are realizations which students have to make for themselves from the facts. Teaching should play a key role in this area, making quite clear to the students that they should actively engage in conceptualization. It may be possible to teach concepts directly, but care should be taken not to make courses too abstract and consequently incomprehensible. Therefore, concepts are best acquired by processing facts in a multi-faceted manner through a variety of activities such as worked examples, research projects, reading, analogies, discussion, reflection, and learning how important discoveries were made.

Finally, the subjects that I would prefer to teach are bio-, physical, theoretical, computational, and general chemistry. I also have a solid knowledge in a range of other subjects learned from coursework, self study and reading, seminars and conferences, and research experience. These subjects include biophysics, computer programming, mathematics, molecular biology, neuroscience, pharmacology, and physics, and I would be able to incorporate these into appropriate interdisciplinary courses. By staying abreast of these fields, I aim to keep the syllabus and my teaching methods up to date.